AN ANALYSIS OF THE PERFORMANCE OF PUBLIC ELEMENTARY SCHOOLS IN NEW YORK CITY DURING 2001–2005 FROM A GEOGRAPHICAL PERSPECTIVE

by

YLLI KELLICI

A dissertation submitted to the Graduate Faculty in Earth and Environmental Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2009

This manuscript has been read and accepted for the
Graduate Faculty in Earth and Environmental Sciences in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Sean Ahearn, Ph.D.

Date	Chair of Examining Committee

Yehuda Klein, Ph.D.

Date	Executive Officer

Sean Ahearn, Ph.D.

Edward Binkowski, Ph.D.

Yehuda Klein, Ph.D.

Hongmian Gong, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

AN ANALYSIS OF THE PERFORMANCE OF PUBLIC ELEMENTARY SCHOOLS IN NEW YORK CITY DURING 2001–2005 FROM A GEOGRAPHICAL PERSPECTIVE

by

Ylli Kellici

Advisor: Professor Sean Ahearn

This study examines from a geographical perspective the factors that impact the performance of public elementary schools in New York City during 2001-2005, a period when its schools were undergoing major reforms at both the local and national level. Education reforms have focused their attention on schools by increasing their responsibility and autonomy concurrent with an increase in accountability. This increased focus on schools as the main agent of change presents a challenge for geographers to investigate the impact of the geographical context on school performance. Although school achievement has a geographical component, the contribution of geographers on this issue has been limited.

There are several major findings of this research. Despite a general increase in educational achievement in the public elementary schools in New York City during the period of this study, the race/ethnicity achievement gap remains substantial and does not change in space or over time. Findings indicate that there is spatial clustering in school

performance the majority of which is explained by the spatial pattern of students' socioeconomic characteristics. There is no major spatial variation in the relationship between school performance and student and school characteristics and such relationships do not change over time. The presence of spatial dependence in school performance, not accounted for by student and school characteristics, necessitates the use of spatial regression models. The spatial model estimation indicates that in the first three years of the study period the spatial error model is a better fit of the data whereas in the last two years the spatial lag model is a better one. The switch from an error to a lag model, occurring in the first academic year (2003-2004) when education reforms were implemented in New York City, is an indication of a global change which can be identified with the effect of reforms throughout the City's public schools. Regression analysis shows that some school districts have an impact on school performance, after accounting for the student and school characteristics. From a policy perspective, the concentration of social disadvantage in space and over time should be taken into consideration in policies regarding allocation of resources that should be spatially focused. Furthermore, the education reforms should consider not only schools but also school districts in their accountability system.

In addition to its findings, this research contributes to the geographic literature by introducing a robust framework to explore the impact of the geographic context. This framework, with the Square Combining Table method at its core, includes also Bisquare Weights and Multiple Comparison procedures. Another important methodological contribution is the introduction to educational literature of the Jackknife technique to examine the spatial variation of the relationship between school performance and student and school characteristics. The methods applied on the jackknifed coefficient values to study their variation in space and over time can similarly be used in other geographical phenomena.

This dissertation is dedicated to the memory of my parents

my mother, Fiqerete Kellici

and my father, Abdyrrahman Kellici

Acknowledgments

I would like to thank many people that have helped me at various stages of my graduate studies. First of all, I would like to thank my advisor, Dr. Sean Ahearn, for his intellectual support, encouragement and enthusiasm during the process of this dissertation and for the collaborative and stimulating environment that the Center for Advanced Research of Spatial Analysis (CARSI) Lab has become under his leadership. Like most research, this dissertation owes a great deal to the environment in which it was conceived. I have learned a lot during my tenure at CARSI, which has been an exceptional support during all the years of this research. I thank my fellow workers at CARSI with whom I had the pleasure to work over the past few years, particularly Karen Rutberg and Loretta Ford, for their support and encouragement.

I thank Dr. Yehuda Klein and Dr. Hongmian Gong, my dissertation committee members, for their challenging questions, valuable comments and opinions.

My thanks go to the Earth and Environmental Science Program at the Graduate Center of the City University of New York for its financial support, to the former and current Executive Directors Dr. Jeffrey Osleeb and Dr. Yehuda Klein, their assistant Lina McClain, for their support and encouragement. I would also like to thank the faculty and staff at the Geography Department at Hunter College for their teachings, stimulating discussions and friendly environment. Thanks also go to the Division of Assessment and Accountability at the Department of Education of New York City for providing me the data used in this study.

I am also grateful to my committee member, Dr. Edward Binkowski, who through his teachings made statistics an interesting subject for me. His courses and advice have been a very important influence on this study.

On a personal level, I would like to acknowledge the support of my family and friends, particularly Agron Cece, who have helped keep my spirit alive and whose support and confidence never wavered. Above all, my thanks go to my loving wife, Irena, for her support since the beginning of this endeavor until fruition. This research belongs as much to her as it does to me.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction

This study examines the performance of public elementary schools in New York City from 2001 to 2005, a period when the public education system in the country was undergoing major reforms. At the national level, No Child Left Behind (NCLB) Act of 2001, signed into law in January of 2002, was a major change in the public education system that expanded the role of federal government. In New York City, Michael Bloomberg, after winning the mayoral elections in 2001, was granted control of the biggest public education system in the country by the New York State Legislature. In January 2003, Mayor Bloomberg launched Children First, a reform to radically change the public education system in the city (Children First History, 2003).

Reforms, at both the national and local level, focus their attention on schools by increasing their responsibility and autonomy concurrent with an increase in accountability. They also promote transparency by informing the public about the school performance in standardized tests and characteristics of the school and its students. As a consequence, this period of education reforms has been associated with new information that provide an opportunity to research the complex and multifaceted public education system of New York City. Furthermore, it is an opportunity not only to evaluate the impact of the reforms, but also to assess the different factors shaping educational achievement.

The remainder of this chapter is organized as follows. First, it provides the context of this study with an overview of both NCLB and New York City reforms, their goals and the changes that they brought to the public education system. Then, a discussion of the purpose of the study and the research questions it is trying to answer follows. The chapter concludes with the research significance of this study.

1.1. National Reform on Public Education

The NCLB was the latest revision of the Elementary and Secondary Education Act (ESEA), the federal law on elementary and secondary education, first enacted in 1965. NCLB, passed with strong bipartisan support, represented a significant shift in federal education policy by making the federal government a major force in shaping the goals and outcomes of education. Prior to this law, the traditional role of the federal government in education has been limited to provide aid to disadvantaged pupils and fund research and development (Stiefel et al, 2004). NCLB expanded this role by establishing a comprehensive framework of standards, testing and accountability to stimulate states to raise the achievement of low-performing students and reduce the achievement gap by race, ethnicity, language and special education status (NCLB, 2002).

ESEA, designed to address the problem of inequality in education, was the first major change in the role of federal government in education. Under Title I of this act, the federal government distributed funding to schools and school districts that had a high percentage of students from low income families. These funds were being allocated based on the number of eligible students, irrespective of whether they improved the performance of the target population (Fusarelli, 2004). Earlier authorizations of ESEA required standards and tests, but did not enforce accountability mechanisms (Hannaway and Hamilton, 2008). Fusarelli (2004) noted that since 1975, reading and math scores on the National Assessment of Educational Progress (NAEP), the nation's report card, showed little improvement even though there was a tenfold increase in federal spending on ESEA. The author suggested that the apparent ineffectiveness of Title I federal expenditures to eliminate the achievement gap and the embrace of accountability policy nationwide were the major reasons behind the latest revision of ESEA by NCLB.

The shift in education policy from a focus on school inputs to a focus on student outcomes and from minimum competency to high proficiency standards started with the national report in 1983 - A Nation at Risk - which stimulated a standards-based education reform across the nation (Lee and Wong, 2004; Fuller et al, 2007). Seven states (New York one of them) called "first-generation" accountability states, adopted accountability policies before NCLB and started holding schools accountable for their students' educational achievement (Lee, 2006). NCLB made accountability a national policy. The law was designed on the premise that if pressure is put on the low-performing schools receiving federal funds than better results in their performance will follow. The intent was that this external accountability and the imposition of sanctions will force schools to improve and motivate teachers to change their instructional practices, resulting in better school performance (Sunderman et al, 2006).

NCLB raised the expectations for higher achievement for poor and minority students in the hope of reducing the achievement gap, so no child is left behind. This goal was to be achieved through the establishment of standards and performance-based accountability mechanism. Standards are defined by subject and grade level and are comprised of important areas for students to learn, while performance-based accountability mechanisms establish incentives for teachers and schools to advance learning in those areas. Hannaway and Hamilton (2008) discuss the relationship between them emphasizing that the critical part of the reform is the alignment of curriculum and instructional practices with the standards and, moreover, that this alignment will increase from the response of educators to incentives and assistance provided by the law.

The objective of the law is to bring all students and defined subgroups such as: economically disadvantaged students, students from major racial/ethnic groups, students with disability and students with limited English ability, to proficiency in reading and math by the 2013-2014 academic year. NCLB does not create a national achievement standard but, in line with the principle of local control of schools, assigns each state the task to develop and implement the academic standards in reading and math. The states also develop annual tests that are linked to those standards for grades 3 through 8 in both subjects. These standardized tests are used in evaluating a school's progress toward the NCLB goal.

Each school has to make adequate yearly progress (AYP) to achieve the goal in each of the defined subgroups or otherwise will be identified as "in need of improvement". Schools that receive Title I funds and fail to make AYP for the second consecutive year, must offer to all students the option to transfer to another public school, either in or out of district, with transportation provided by the district. The school should also develop a two-year school improvement plan. In addition, these schools are eligible to receive technical assistance from the state. In schools that fail to make AYP for the third consecutive year, students are eligible for supplemental educational services. For schools that fail AYP for four consecutive years, the district must implement corrective actions to improve them, such as the replacement of staff members or the implementation of a new curriculum. Finally, if schools fail to make AYP for five consecutive years, they can be restructured, taken over by the state or a private management contractor, converted to charter schools or reconstituted with new staff (NCLB, 2002). NCLB's requirements that students in failing schools be given the option of obtaining supplemental services or transferring to successful schools links NCLB's standards-based accountability to market-based reforms. The objective of policy makers was to incorporate the assumed benefits of market-driven systems into public education, while preserving its unique character (Fusarelli, 2004).

NCLB emphasized the importance of teacher qualifications in the improvement of school performance. Research had shown that low-income students had the least prepared and experienced teachers and that the quality of teachers substantially affected student achievement. In response to such findings, the law required that all schools receiving NCLB aid must have qualified teachers. The underlying assumption was that principals and school districts were not trying hard enough to attract good teachers to impoverished schools and that they could rapidly fix this problem if required by law (Sunderman, 2006).

Another major change by the NCLB was the promotion of transparency. The law required that the public be informed about school characteristics and school performance in standardized tests. Test results on how each school is performing are made public in the Annual Report Card and, to help ensure that all groups of students are progressing at an adequate rate, the results are broken down and reported according to subgroups such as: race/ethnicity, special education, English language learners and economically disadvantaged. The disaggregation of the performance by subgroup brings more attention to low subgroup performance which was hidden before within school, district or statewide averages. Moreover, this new information allows public officials to look more closely at student performance and direct instructional resources toward groups of students whose academic needs the school system failed to address. The Annual Report Card also contains information about the school and its student characteristics such as: professional qualifications of teachers, school's enrollment, students' attendance, students that are recent immigrants, number of students eligible for free lunch, English language learners and special education students.

Several aspects of the NCLB have been criticized. Regarding NCLB's objective of having all students at the proficient level in 2014, Lin et al (2002), using a definition of proficient as in NAEP, concluded that such a goal is so high that it is completely out of reach. The authors emphasized that setting unobtainable goals can backfire by demoralizing the teachers instead of motivating them. Bryant et al (2008) reached the same conclusion for nearly all elementary schools in California, demonstrating that they will fail to meet the goal of proficiency by 2014. NCLB's accountability has also been subject to criticism. Kelly and Monczunski (2007) argued that NCLB unintentionally favors schools with fewer minority populations because schools with more subgroups are less likely to meet AYP. Stiefel et al (2007) share the same concern. The authors, using data on elementary and middle schools in New York State, note that the accountability burden will be higher on urban schools that are more racially and ethnically integrated compare to rural and suburban schools. Another point of criticism has been that the new requirements by NCLB were not supported with sufficient resources (Lee, 2006). In addition, Sunderman (2006) disagreed with NCLB's assumption that teachers will respond to being sanctioned and labeled as failing. The author argued that teachers can make a difference but they need external help whereas NCLB did not provide the policies, support or flexibility needed to meet its goals.

Another concern has been related with the NCLB's increased emphasis on the role of standardized testing as a performance indicator. Fusarelli (2004) noted that teachers will be forced to spend more time on preparation for the test than on authentic teaching and learning. Such focus on testing may narrow the curriculum by pushing out material not covered on the exams and, as a result, nonacademic areas such as sports, music and art may become eliminated because of the focus of attention and resources to preparation for tests. Furthermore, the author observed that even though NCLB explicitly assumes that state tests are aligned to the curriculum, according to research, this is seldom the case. Fuller et al (2002), like other researchers, were concerned that the narrow focus on teaching to a state test may produce inflated gains in scores whereas the fundamental concern should be with improved achievement not just higher test scores.

They also questioned the degree to which the gains on a state test can be generalized to gains on other measures of achievement such as for example NAEP. After four years since the NCLB was enacted, a study by Lee (2006) concluded that the use of state assessments as the basis of school accountability were misleading because tests administered by the states tend to significantly inflate proficiency levels as well as deflate racial and social achievement gaps in the state.

1.2. New York City's Education Reform

The New York City's public education system has undergone major reforms under Michael Bloomberg who pledged to fix the education system when he ran for mayor of New York City in 2001. Bloomberg made gaining control of the system one of his top priorities during the campaign. After being elected, he successfully pressed the state of New York to grant him control of the city schools, beginning in academic school year 2002-2003. The Mayor earned the power to appoint the Chancellor and assumed responsibility for the schools' performance. Since then, Bloomberg and his administration have left virtually no part of the system untouched (Schwartz and Stiefel, 2005; Hemphill et al, 2005). In January of 2003, the Mayor launched a plan to reform governance and curriculum in the New York City's public schools (Children First, 2003).

This was not the first time that New York City had a central management; the city had mayoral control from 1873 to 1969. By the end of the 1960s there was a campaign opposing the centralized system and advocating decentralization to better serve the local community needs (Hess, 1999). In the midst of the public concern about equal educational opportunity, New York State legislators solved the conflict by sanctioning decentralization and creating in 1970 a school system with 32 community school districts. These districts were governed by community school boards, selected in local popular elections to give local communities a significant voice in who run their schools and how (Wolff, 2002). Community school districts in New York City were wholly responsible for the operation of schools within their districts; they controlled the allocation of all resources to schools, the development and control of curricula, had the power to hire community superintendents, principals and assistant principals. Community superintendent were responsible for all the district's instructional policy and fiscal matters, including hiring, curriculum initiatives and professional development (Iatarola and Fruchte, 2004). In 1997, in an attempt to create greater system-wide accountability, the state legislature stripped local boards of their responsibilities, re-centralizing power with the Chancellor and superintendents and leaving the boards with policy setting duties but no real authority (Hess, 1999; Ravitch and Viteritti, 2000).

In her article in Gotham Gazette, a few weeks before the state legislature gave control to the Mayor, Wolff (2002) described the disagreement that existed at the time about the value of local boards. She noted that when the community boards worked they were a real avenue for participation, set sound policies, supported district efforts and, through them, there was a structured way for people to have a role. Even though voter turnout was always low, still thousands of people participated in school board elections. In addition, because parents who were not citizens could vote, it gave a voice to people who often struggled to have their views heard. The author, though, observed that while some boards have been avenues for greater community involvement in schools, many have not. For a long time school boards have ignited criticism and complaints; some had been accused of patronage and corruption and others for handing out jobs to political partners.

The legislation signed by Governor George Pataki on June, 2002 turned control of the school system over to the Mayor. The old and powerful Board of Education was replaced by the Panel for Educational Policy, which does not have the same power because the Mayor appoints the majority of its members. The Mayor appointed the Chancellor Joel Klein who quickly implemented re-organization of the governance of the schools. Under the new state legislation, the new Chancellor assumed the powers of the Board of Education and gained direct authority over the district superintendents and 32 community school districts. The new law changed the role of community school districts boards by replacing them with community district education councils. Each new district council has nine parent members selected by the parents associations from that district, two community members who are appointed by the borough presidents and a nonvoting student member, appointed by the district superintendent. As with the boards they replaced, the councils' duties are largely advisory. Baum (2005) described the activities of education councils over the period of few months after they replaced the 32 community school boards in fall of 2004. The education councils hold monthly meetings with the superintendent to discuss the district's educational progress, submit an annual performance evaluation of the district's leadership to the Chancellor, hold a hearing on the district's annual capacity needs, review the district's educational programs and their impact on student achievement, approve boundaries between schools, evaluate supervisors and make recommendations on building use. The author noted that, despite frustrations and setbacks, the community education councils were more effective in their first year than many observers expected.

Along with changes in the organization of the system, curriculum reform followed. In the academic year 2003-2004, public elementary schools had new and uniform math and reading curricula. Until then, New York City had no uniform education program for its more than a million students. Individual community school districts and schools selected their programs, often changing them from year to year. This confused students and placed a particular burden on low-income youngsters who tend to switch schools more frequently than their more affluent classmates (Robinson, 2004). The Mayor and the Chancellor instituted a new, uniform curriculum, an ambitious attempt to offer all schools in the city progressive techniques that had previously been applied in middle-class or wealthy neighborhoods such as Manhattan's Upper East Side. Hemphill et al (2005), who surveyed nearly 500 of the New York City's elementary schools, described in their book the new changes that had taken place in those schools such as: more inviting classroom libraries of children's literature and picture books that had replaced unappealing textbooks; teachers being encouraged to lead class discussions, rather than to offer lectures; children that had begun to write far more frequently and in more depth than had been common in elementary schools.

The quality of teaching was also another subject of the reforms in New York City. To comply with teacher requirements from NCLB act, the city toughened its requirements for teaching certificates. The Department of Education cut hiring of uncertified teachers and expanded its recruitment of alternatively certified teachers (Kane et al, 2007). The New York City Teaching Fellows program had accounted for most of the growth in alternative certification in the city, for the five year period since its creation in the summer of 2000. The program streamlined and subsidized the certification process for college graduates willing to teach in low-performing schools, mainly in Bronx and central Brooklyn (Biederman, 2004). Teachers that graduated from the program contributed in the improvement of school achievement in new small schools, even in high-poverty neighborhoods, that had a competent administration and encouraging environment. In contrast, fellow teachers were not successful in poorly run schools with weak leadership (Aspel, 2005).

Children First initiative instituted the parent coordinator program in order to help parents to negotiate the bureaucracy. To encourage involvement in individual schools, the Department of Education placed full-time paid parent coordinators in all but the smallest city schools. The parent coordinators help parents with paperwork and provide support for academic success (Robinson, 2004). Other changes that occurred on public education in New York City included: creation of charter schools, vouchers, a new training academy for school principals, hiring of reading and math coaches to support teachers and so on. Following the belief that small schools are more successful, the effort to build new schools continued, funded in part by the Gates Foundation.

The New York City's education reform has not gone through without criticism. Herszenhorn (2005), in an article on New York Times, noted that on much of the school system the main impact of the changes had been confusion, start-up difficulties, conflicting policy changes, high staff turnover and problems with the safety on schools. Hemphill et al, (2005) also discussed the administrative problems created by the change in the organization of the system. Special education students did not get some of the services they needed because thousands of special education records were misplaced. There has been criticism from principals and staff in some schools, particularly in Queens, about the micromanagement by the central administration. In those schools, the reforms were presented as a non-negotiable list of new rules to be followed, rather than a creative new approach to teaching. As a result, dozens of the city's best principals and administrators left, depriving the school system of their needed leadership.

The system of mayoral control has been criticized for shutting out parents and others from any meaningful involvement in school decision-making. Critics complain that parents too often are ignored and do not have elected board members to turn to for help. Since parent coordinators are hired by the principal, they are not always seen as responsive to parent's needs (Robinson, 2004). The impact of parent coordinator has varied because every school has just one parent coordinator regardless of the size of the school (Herszenhorn, 2005).

In its report at the end of the academic year 2004-2005, the New York City Council Education Committee rated the Department of Education "not proficient" in class size, teacher and principal morale, special education and operations (Saulny and Herszenhorn, 2005). Even though the administration claimed of having driven down the cost of school construction, critics said that the new schools and the state money to build them did not materialized fast enough to reduce class size (Herszenhorn, 2005). In addition, the new small schools aggravated overcrowding in the existing big schools. Another issue has been the concentration of inexperienced teachers and principals in schools with high-needs kids. Finally, while the administration considered the rising test scores in 2005 as signs of success, critics questioned whether the scores really have increased significantly. They raised concerns about the validity of the tests suggesting that the improvement in scores was partly attributable to the decline in the number of eligible students tested, especially among students with limited English skills, and to he less demanding nature of the 2005 exam.

1.3. Purpose and Research Questions

The central issues of public education reforms over the last decade have been the increase in school performance, the reduction of the achievement gap between different subgroups and the improvement of school factors affecting performance. These reforms have focused attention on schools by increasing their responsibility and autonomy concurrent with an increase in accountability. The increased focus on schools as the main agent of change presents an opportunity for geographers to investigate the impact of the geographical context on school performance.

School achievement has a geographical component because it is affected by the characteristics of the school catchment area, its surrounding areas, and by the management and policies of the school district to which the school belongs. NCLB requires schools to make available to the public, in their Annual Report Cards, data about exam results and different school and student characteristics. These data aggregated at school-level are at the appropriate scale for geographical analysis because geographers are interested in the spatial pattern of school performance and in the spatial social process generating its clustering in space at the neighborhood and urban scale. Moreover, the recent increased importance of school-level test results as an indicator of school effectiveness and as a measure of the impact of new policies makes empirical analysis that use schools as the unit of study noteworthy. Although school achievement has a geographical component, the contribution of geographers on this issue has been limited.

The purpose of this study is to examine from a geographical perspective the performance of public elementary schools in New York City and factors shaping it during 2001-2005, a period when its schools were undergoing major reforms. Being the biggest public education system in the country, reforms in New York City have been of national interest, with lessons learned being applied in other cities. New York is a very diverse and dynamic city which provides the opportunity to study different race/ethnic subgroups and changes in the system over time. Even though New York City is unique because of its size, in every other aspect its problems of urban education are similar to the problems of urban education in the country. Public elementary schools in the city were the first to have been affected by the reforms and, therefore, their results on state standardized tests have been used in evaluating the impact of reforms. The availability of annual state exam results for each race/ethnic subgroup in New York City makes it possible to implement a more in-depth analysis of the achievement gap over time and space. The low performance of students in the English Language Arts tests compare to the math tests has created the expectation for a more rapid improvement in this subject. Since most of the nation's public schools do operate under some form of school district control, this study considers as important the assessment of their impact on school performance. Even though the 32 community school districts operate within the public education system of New York City, they resemble in their functionality and diversity with other districts in the country that function under the state system. For three decades the New York City's community school districts have operated as independent units, with their own identities and serving different student populations. As such, this study's perspective, analysis and findings are relevant for other districts in the country.

From a geographical perspective, several research questions are raised by this study:

- Is there inequality in educational achievement in New York City across districts and does it change over time?
- Is there variation in the achievement gap between race/ethnicity subgroups across districts?
- What are the student and school characteristics that explain the spatial patterning of school performance and does this relationship change over time?
- Does the relationship between school performance and student and school characteristics vary in space?
- Is there a spatial dependence in school performance and what is the spatial process generating it?

The first two research questions are addressed in chapter three and the last three in chapter four.

1.4. Research Significance

This research investigates, to the best of my knowledge, for the first time the school performance and the achievement gap in New York City's public elementary schools from a spatial perspective. Although school achievement has a geographical component, the contribution of geographers on this issue has been limited. This study takes advantage of the new data available and performs a more in-depth analysis of the race/ethnicity achievement gap. It contributes to the educational literature by introducing a robust framework to examine the effects on school performance of three components: time, race/ethnicity and space. Within this framework, this study applies a robust technique, Square Combining Table, to estimate the effects of each component and their interactions. Then, the Multiple Comparison technique is used to assess the statistical significance of these effects. In addition, this research introduces to the educational literature the Jackknife technique which is used to explore the spatial heterogeneity of the relationship between school performance and independent variables.

From a policy perspective the main contribution of this research consists in its finding about the persistence over space and time of the race/ethnicity achievement gap between the low performing subgroups (Black and Hispanic) and high performing ones (White and Asian). After accounting for race/ethnicity effect, there is inequality in educational achievement between the lowest performing districts in Upper Manhattan and south Bronx and the highest performing ones in north-east Queens and Lower Manhattan. This study finds that districts have no impact in reduction of the race/ethnicity achievement gap. The increase in school performance over the time period of this study was not associated with a reduction in race/ethnicity achievement gap, indicating that reforms did not achieve their main goal of gap reduction. Findings indicate that there is spatial clustering in school performance the majority of which is explained by the spatial pattern of students' socioeconomic characteristics. There is no major spatial variation in the relationship between school performance and student and school characteristics and such relationships do not change over time. The presence of spatial dependence in school performance, not accounted for by student and school characteristics, necessitates the use of spatial regression models. The concentration of social disadvantage in space and over time should be taken into consideration in policies regarding allocation of resources that should be spatially focused. Furthermore, the education reforms should consider not only schools but also school districts in their accountability system.

Chapter Two

Literature Review

This chapter reviews the literature relevant to several aspects of this study including: assessment of the impact of NCLB on state performance, analysis of racie/ethnicity gap in school performance, the impact of community school districts in school attainment, school characteristics that effect school performance and the geographical analysis of school achievement.

2.1. The Impact of NCLB on School Performance at State Level

Previous research on the impact of NCLB on school performance has explored academic achievement over time at the state level. The law established state assessments as the basis for NCLB accountability, but it recommended corroborating those results using an independent national test, the National Assessment of Educational Progress (NAEP) which is the federal assessment system run by the Educational Testing Service, the so-called nation's report card of student achievement. Fuller et al. (2007) used NAEP data spanning over the 1992-2006 period for 4th graders in 12 diverse states, to study the effect of the law. The authors found that the consistent growth in reading test scores over a three year period before the enactment of the NCLB in 2002, largely faded and flattened out over the three years after the reform. The progress in math continued from 1999 to 2004, but the growth was slower after compared to before the enactment of NCLB. The authors admitted that the discrete effect of NCLB, beyond the momentum of state-led accountability reforms, was difficult to estimate. They found similar patterns regarding the narrowing of the achievement gap; there was no progress, on both reading and math subjects, in reducing the Black-White gap after compared to before the enactment of the law. In contrast, the Hispanic-White gap continued to narrow.

Lee (2006) analyzed NAEP reading and math results for the 4th and 8th graders at the national and state level. The author using data from pre-NCLB (1990-2001) and post-NCLB (2002-2005) periods found that NCLB did not have a significant impact on improving reading and math test results across the nation and states. The national achievement in reading remained the same as before the NCLB whereas in math it continued to grow at the same rate. Math test results for students in 4th grade showed a temporary improvement in the growth rate right after NCLB that was followed by a return to the pre-reform growth rate. The author stated that the continuation of the current trend will leave the nation far behind the NCLB target of 100 percent proficiency by 2014. Similar results were found with respect to the reduction of the racial achievement gap leading the author to conclude that the NCLB reform had not helped the nation and states to significantly reduce it.

In addition, Lee (2006) stated that contrary to NAEP results, the only independent national test, state assessment results showed improvements in reading and math. The study concluded that NCLB's reliance on state assessment as the basis of school accountability is misleading since state-administered tests tend to significantly inflate proficiency levels and proficiency gains as well as deflate racial and social achievement gaps within the states. The author found that the higher the stakes of state assessments, the greater the discrepancies between NAEP and state assessment results. These discrepancies were particularly large for poor, Black and Hispanic students. The author concluded that if the same policy course is maintained, academic proficiency is unlikely to improve significantly, even though it is possible that the state assessment will continue to give a false impression of progress which will encourage more investment into a failed test-driven accountability reform policy.

2.2. Race/Ethnicity Gap in School Performance

The persistence of the gap in school performance between racial and ethnic subgroups has been the focus of policy makers and researchers for a long time. The NCLB Act acknowledged that the growing achievement gap was left unaddressed for too long and made gap reduction its main goal. The racial/ethnic achievement gap has been documented and researched extensively using data at the national and state level. The Equality of Educational Opportunity Report, also known as Coleman Report (Coleman et al, 1966), one of the largest studies in history funded by the U.S. Department of Education, has been influential and has served as the base for numerous public policies in the area of education. The findings of this report strengthened the general understanding about the existence of the racial/ethnic gap in education achievement. The report demonstrated quantitatively the important fact that achievement differences were found at all grade levels and in approximately the same degree.

Several studies in Jencks and Philips (1998) examined the scope and causes of the racial/ethnic achievement gap. The evidence presented in their volume indicated that the gap emerged prior to kindergarten and persisted throughout the elementary and secondary school years. Reardon (2003) examined the causes of the gap in the earliest years of schooling. The author assessed the contribution of three processes in the development of test score gaps: out-of-school, between-school, and within-school processes using data from Early Childhood Longitudinal Study for kindergarteners through first graders. The author found that there was a large race/ethnic gap in reading and math tests as children entered the kindergarten and a substantial proportion of that initial gap was explained by socioeconomic status (SES) differences between race/ethnic groups. Moreover, after controlling for SES and school context, the gap increased during the kindergarten year, even between students enrolled in the same schools, indicating that within-school processes played an important role in the growth of the gap. The author also found that between-school processes accounted for 60 to 80 percent of the change in the gap in first grade.

Stiefel et al. (2006) examined the size and distribution of the gap in reading test scores across races within New York City's public schools and the factors that explained that gap. The authors utilized student-level data of 5th and 8th grades for 2000-2001 academic school year and found significant disparities between the test scores of White and Black students and those of White and Hispanic students. These test score gaps were partially explained by differences in student characteristics, such as poverty, and differences in limitations in English proficiency, at least for the Hispanic-White gap. The authors were not able to explain the entire within-schools gap, which remained even after controlling for differences in classrooms. They concluded that differences in schools attended by the students were also important and found evidence that school characteristics mattered. Based on their finding that racial gaps were negatively correlated with school size, the authors suggested that some reduction in these gaps could be achieved by redistributing children across schools.

2.3. The Role of School Districts on School Performance

Over the last two decades the focus of educational reforms has been on schools as the principal unit of change while the role of district has been overlooked. The NCLB is no exception. As a result, the focus of the research has been mainly on schools while the effect of districts on school performance has received much less attention. Recently, however, the accountability systems with schools at their focus, which ignored or diminished the role of districts in education achievement, have begun to be challenged by public policy makers and researchers (Iatarola and Fruchter, 2004).

Iatarola et al. (2002) addressed the question of whether and to what extent districts contribute to students' achievement, after controlling for students and school characteristics. Using data on New York City's public schools and districts from three academic years (1996 to 1999) the authors found that school districts had an important role in student performance. They implemented school-level analysis and measured the variation in school-level test score results, that was unaccounted for by variation in student and school characteristics, which was attributed to district effects. In their analysis, district effects measured the influence of unmeasured or unobservable factors as well as measurable district factors that contributed to school-level performance. Their findings of significant district effects provide strong statistical evidence for the effectiveness of districts, an aspect that previous research had not demonstrated.

Iatarola and Fruchter (2004) investigated the ways that districts affect the academic achievement of students and schools. The authors selected four districts, among the 32 in New York City, from both high and low performers, which varied in their characteristics. They found differences between high and low performing community school districts in terms of educational goals, instructional focus, leadership development, teacher recruitment and retention, and professional development. The authors suggested that although national reform efforts such as NCLB are directed primarily at schools, local school districts should be considered because they are responsible for a number of functions critical to schooling effectiveness such as: hiring, collective bargaining, curriculum development, assessment, fiscal operations and auxiliary functions.

With regard to districts, other than their role in school performance, researchers have examined a wide variety of topics including district administrative spending, district consolidation and district size (Iatarola et al, 2002). In their study, Driscoll et al (2003) found that controlling for characteristics of the student population and other environmental factors, including class and school size, district size appears to have a negative impact on educational achievement, having its biggest impact on middle school student performance.

2.4. School Characteristics and Their Effect on School Performance

The Coleman Report (Coleman et al, 1966) is credited with launching a substantial increase of studies assessing the relationships between educational achievement and school characteristics. The authors found that family socioeconomic status and peers were the most important predictors of student performance whereas school characteristics had little effect. While the subsequent studies have confirmed Coleman's findings with regard to the impact of family characteristics on educational achievement, the relative importance of school factors has been debated.

Powers (2003), using school as the primary unit of analysis, investigated the relationship between school performance and school characteristics. She examined the relationship across two large urban districts in California to control for unmeasured variations in district policies. The study covered two academic years and used the same sample of schools to allow for both cross-district and cross-year comparisons. School-level independent variables were combined in three groups: student, teacher and school characteristics. Student characteristics included students eligible for reduced-price or free lunch, mobility and English language learners. Teacher characteristics included measures of their credentials, experience and education. The author found a strong negative correlation between school performance and the variables measuring the socioeconomic status of the student population such as the percentage of students eligible for reduced-price or free lunch, mobility and percentage of English learners. The variables measuring teachers' credentials and experience were positively correlated with performance and were statistically significant. Regarding the policy implications of her findings, the author suggested policies with the goal of integrating schools by socioeconomic status, attracting and maintaining an experienced staff at schools and creation of pay or other incentives to encourage experienced teachers to work in high-poverty schools.

Barrow and Rose (2005) provided evidence that teacher quality mattered and is positively related with experience. The authors also stated that smaller class sizes can improve student outcomes. Gibbons (2002) found that after allowing for neighborhood composition, residential selection and prior school performance there was no evidence that average incomes influence school performance, although the well-known relationship between performance and students eligible for free meal persisted. The author stated that most of the relationship between school performance and catchment area characteristics could be attributed to differences in the underlying attributes of the people who live there, rather than the incomes these people receive. School size had a significant negative effect on assessment tests; an increase in school size by one hundred students decreased school performance by 0.5 percentage points.

2.5. Geographical Analysis of School Performance

Geographers, mostly in England, in their studies of the geography of school achievement using test scores at the school-level and their relationship with the socioeconomic characteristics of the school catchment areas, have emphasized the need to contextualize school performance (Brimicombe, 2000; Conduit et al. 1996; Gibson and Asthana, 1998). Bradford (1991) argued that focusing on the school neglects the fact that social geography matters and it affects some of the outcomes being attributed to school processes. Other studies using information on individual students have also provided evidence about neighborhood effects on students' educational achievement (Coombes and Raybould, 1997; Garner and Raudenbush, 1991).

In a workshop held in Lancaster in 1998, a group of geographers was asked to analyze the same dataset of primary school performance test scores in northern England. They carried out what they considered to be appropriate analyses of this dataset, with the overall aim of determining how much the test results reflect catchment area characteristics. Flowerdew and Pearce (2001) used different ways of defining catchment areas in their analyses. In this context, they pointed out that the modifiable area unit problem is also relevant. In other words, the catchment area definition affected the results of statistical analysis. Their results showed the importance of variables such as parent educational level, ethnicity, housing tenure and unemployment. Martin and Atkinson (2001) applied an autoregressive model, not previously used in modeling educational data. They used the Townsend index (which comprises home and car ownership, overcrowding and unemployment variables) scores as the independent variable. In this model school test results are regressed not only on Townsend scores in the school's catchment area, but also on school scores in surrounding areas. Fotheringham et al (2001) used Geographically Weighted Regression to determine if there were any interesting spatial variations in the relationship between school performance and environmental variables. They found a negative relationship between a school's overall performance on standardized tests and a high level of unemployment, public housing occupancy, and the proportion of single parent households in a school's catchment area and positive relationship with the high levels of professionals.

Gibbons (2002), analyzing school level data in England, found that the clustering of primary school performance remained even after controlling for the characteristics of school and catchment area residents, suggesting that school performance depended on the performance of neighboring schools. The author noted that this spatial dependence could be a neighborhood human capital spillover, operating through social interaction of pupils from neighboring schools or a knowledge spillover in terms of teaching technologies.

In a 2007 Special Issue of Urban Studies journal, several researchers considered the ways in which educational policies are operating in England and elsewhere in the developed world. In their introduction to the Special Issue, Butler and Hamnett (2007) noted that England's education system stands in between North-American choice-driven model of educational allocation and a more geographically driven allocation model traditionally favored in Europe. The authors discussed the connection between educational attainment and space, social structure and social processes. They emphasized that spatial variations in the educational achievement are due to spatial variations in educational provisions, social composition by the segregation of different social groups and local school catchment policies. It is the interplay between these different factors that brings in evidence the significance of geography; the existence of class and race/ethnicity based residential differentiation and segregation conditions the provision of education in those areas resulting in some areas having better school buildings, equipment or more teachers than others. The authors discussed other effects from the concentration of different groups in different areas, termed composition or area effects. One aspect of this process is related with the implications created by the concentration of social disadvantage which should be taken into consideration and condition spatially focused policies in the allocation of resources. The other aspect of the social exclusion is evident on the effect of social mix on school performance. Schools located in deprived areas have a lower performance compared to schools in more socially advantaged areas, which is a result of the social mix rather than issues of school management. The policy implication resulting from their finding is that the attempt of school accountability initiative with quasi-market policies that discipline failing schools does not take into account the social intake mix but attributes poor performance to poor management. The authors argued the limitations of this policy that is focused only in the role of school managers and teachers in revitalizing the ethos of the school which, while important, are not the most significant part of the problem.

In one of the articles in this Special Issue, Gordon and Monastiriotis (2007) used school-level data to examine the impact of location at neighborhood and sub-regional scales on dexam results in public secondary school. At the school and neighborhood level the authors found strong evidence of the impact from class and ethnic composition, school absence, students with special needs and the school intake quality. At the sub-regional level, the social class effects continued to be significant explaining the better performance of schools in the south of England. In another article in the issue, Hamnett et al (2007) found that while the ethnic origin of students affected their school attainment in state schools in inner London, it was not as important as their social background. In summarizing the findings from the work published in the issue, Butler and Hamnett (2007) noted that the biggest drivers for differential educational attainment were social background and the social mix of the school and that the social composition was more important than school leadership or teacher motivation. They suggested that these factors should be taken into consideration by policy-makers. The authors also emphasized that in England the publication of school results and the expectation that it would drive school performance were at best marginal.

Chapter Three

A Robust Framework to Assess the Impact of Time, Space and Race/Ethnicity on School Performance

Geographical analysis is the study of the impact of geographic context on the phenomena under investigation. The geographic context is examined at different levels from local scale such as the geographic location and the surrounding neighborhood, to more global geographical scale such as regional, state and national levels. In this study, the geographic context ranges from the school catchment area, to its surrounding areas and to the school district to which the school belongs. This chapter investigates the impact of school districts on the overall school performance and more specifically on the race/ethnicity achievement gap. The district effects indicate not only the impact of district management on school performance but also the contribution of the community within the district, unaccounted for by race/ethnicity. The recent data availability by race/ethnicity subgroups makes it possible for a more in-depth analysis at a finer spatial scale such as that of school district, instead of the larger scale such as cities or states that has been the case so far.

This chapter considers as important the assessment of the impact of district on school performance, since most of the nation's public schools do operate under some form of school district control. Even though the 32 community school districts operate within the public education system of New York City, they resemble in their functionality and diversity with other districts in the country that function under the state system. For three decades the New York City's community school districts have operated as independent units, with their own identities and serving different student populations. As such, this chapter's perspective, analysis and findings are relevant for other districts in the country.

From a geographical perspective, this chapter addressees two research questions. First, is there inequality in educational achievement in New York City across districts and does it change over time? Second, is there spatial variation in the achievement gap between race/ethnicity subgroups? This chapter answers the question concerning the existence of spatial inequality in education achievement in New York City by assessing the impact of school districts on the performance of schools within the district. From this perspective, this chapter also addresses the question of the impact of school districts on the reduction in the race/ethnicity gap. In other words, the objective is to find whether there are districts that have done a better job contributing to the reduction of the race/ethnicity gap. Additionally, since this study took place during a period of change in school policies at the City level, the overall school performance, the impact of school districts on school performance and on the race/ethnicity gap, are examined in the light of these education reforms.

3.1. Data

Since the advent of mayoral control in 2002, the New York City's Department of Education has increasingly relied on standardized exams as an indication of each school's performance and of the state of the nation's largest school system. Despite criticism, standardized tests have become shorthand for determining whether what is being done in the classroom really works. Test results are used as a way to hold principals and teachers accountable while providing information about school performance to parents. This study uses the performance of 4th grade students in English Language Arts (ELA) annual state test as a representation of the performance of each public elementary school. The data for the 4th grade state test results were the first to be broken down by race/ethnicity subgroups in the academic year 2000-2001 and are appropriate to be used for examining the questions that this chapter is addressing.

Research studies have shown that race/ethnicity gap widens beginning in the first year of elementary schools, so the assessment of the gap at this level becomes more important (Jencks and Philips, 1998; Reardon, 2003). The focus of reforms in New York City was initially on public elementary schools and was accompanied by changes in curriculum. ELA test results have been generally lower than those in math, which created the expectation for a more rapid improvement in this subject due to reforms. These are the reasons why this study selected the ELA test results on the 4th grade to represent performance of public elementary schools.

The data used for this study, span over the 2001 to 2005 period, were made available on request from the New York City's Department of Education. The year 2001 is selected as the starting year for this study because it was the first year that the results from the state exams in ELA for the 4th grade were broken down by defined subgroups such as those based on race/ethnicity, special education and English language learners.

This chapter takes advantage of this opportunity by using the results in the state exam disaggregated by race and ethnicity. The 2005 is the last year that the data were available from the Department of Education.

3.1.1. Public Elementary Schools

The academic year 2003–2004 is the first year when reforms introduced by Mayor Bloomberg were being applied fully in the public education system in New York City. The period of investigation under this study covers the period before mayoral control, during the transition and after the reforms were implemented throughout the city. The number of public elementary schools during this period varies by year because new schools were opened and old schools were closed in a dynamic education system as that of New York City (Table 3.1). There are 616 schools included in this analysis because they have been in activity during the entire five year period. Using the same schools throughout the analysis allows for both cross-ethnicity and cross-year comparisons.

Table 3.1. Number of Schools per Year

Year	2001	2002	2003	2004	2005	All Years
Number of Schools	712	717	676	686	694	616

3.1.2. Race/Ethnicity Subgroups

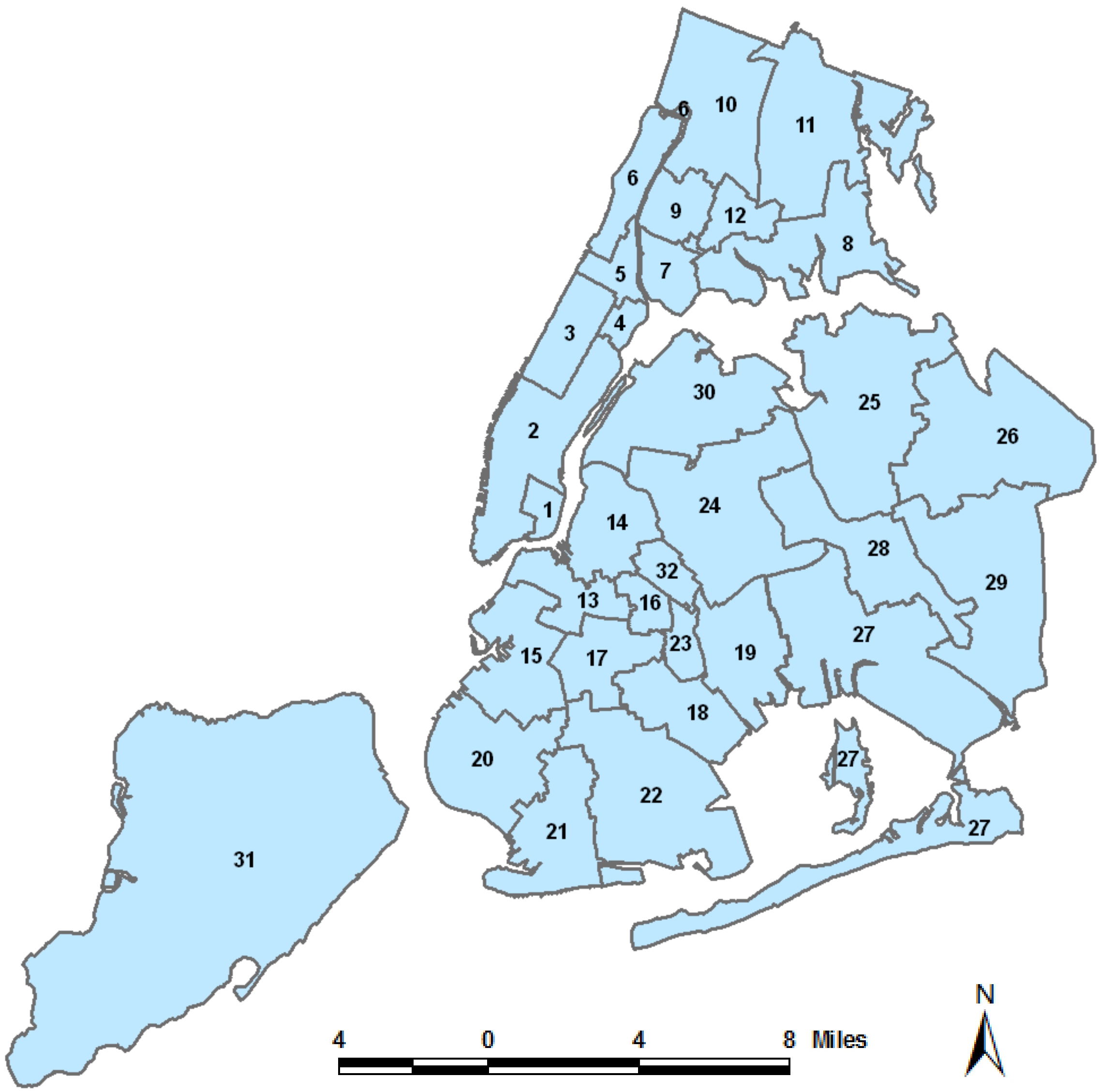
Public elementary schools are required under NCLB to report state test results by race/ethnicity subgroups such as: American Indian/Alaskan Native, Black, Hispanic, Asian/Pacific Islanders and White. To ensure students' privacy, when racial/ethnic groups with fewer than five students are tested, the numbers and percentages for the group are combined with the next smallest group and reported in a separate group called "Small Group Totals". This group is not included in the analysis. The other group excluded is American Indian/Alaskan Native due to their small number of students. School performance is assessed using the results achieved by 4th grade students in annual state test, aggregated at the school level for each race/ethnicity subgroup. Based on their performance, students in each school are grouped into four levels. Students that perform at level 1 have serious academic deficiencies and those at level 2 need extra help to meet the standards established by the state. Students performing at level 3 meet the standards and those performing at level 4 exceed the state standards. In this study, the percentage of 4th grade students performing at the level 3 and level 4 in the state exam in ELA subject is used to represent school performance.

3.1.3. Community School Districts

Each public elementary school in New York City is located within one of the 32 school districts (Figure 3.1). Local community school districts were in charge of managing public elementary schools in New York City for three decades. School districts, by serving different student populations and operating as independent units during this period, have created their own identities within these geographical areas. On June 2002, when the state legislation gave Mayor Bloomberg control over the education system in the city, it also changed their role from management to a more advisory one.

The old powerful community school districts were replaced by community educational councils in fall of 2004 which corresponds to the last academic year in our study period. While their role has substantially diminished after mayoral control, school districts retain some responsibilities as required by New York State.

Figure 3.1 New York City's 32 Community School Districts



3.2. Methods

To answer the research questions posed above, this chapter introduces a robust framework composed of three components: time, race/ethnicity and space. The school achievement over time and therefore the impact of educational reforms is assessed by the time component. The race/ethnicity component is comprised of four major racial and ethnicity groups in New York City: White, Black, Asian and Hispanic. The reduction of the gap in achievement between the low and high performing racial/ethnic students has been the major focus of the NCLB. The analysis of the spatial component makes use of the inherent spatial structure of the education system in NYC where each school belongs to one of the 32 school districts. The magnitude and the spatial pattern of district effects provide an indication of the spatial inequality on educational achievement. An advantage of this framework is that it allows estimating the interaction effects between the three components which would assess the race/ethnicity gap over space and time, and also the district effects over time. Within the framework, this study applies a robust technique, Square Combining Table, to estimate the above effects. Then, the statistical significance of these effects is assessed by using the Multiple Comparison technique.

3.2.1. Framework Layout

This section introduces the layout of the framework to examine the effects on school performance of three components: time, race/ethnicity and space. The structure of our data and research questions that this chapter is addressing lend themselves to the framework offered by Analysis of Variance (ANOVA) where time, race/ethnicity and district are factors with five, four and 32 levels respectively. Any level of each of the three factors occurs in combination with every level of any other factor. This structural relationship between these three factors is a fully-crossed three-way ANOVA (Singer, 1991). When two factors are crossed, the overall effects of each factor (called main effects), averaging over the other factors, can be assessed. The main effect of the year factor will show the pattern of school achievement over time for New York City providing evidence of the impact of education reforms. The race/ethnicity main effect will show the difference in performance between different race and ethnicity subgroups. The impact of the district on the performance of schools within it will be revealed by the district main effect demonstrating the spatial aspect of school achievement.

Crossed factors allow also for the exploration of school performance as one factor is changed while holding the other factor fixed. In this study's data there is more than one observation within every combination of crossed levels which creates the possibility to examine the interaction between different factors. Thus it can be examined whether the effects of one factor differ according to the level of the other factor and these interactions can be separated from chance fluctuations in the data. This allows, for example, to investigate the performance of each race/ethnicity subgroup over time, and thus make a comparison between years conditional on a given race/ethnicity level. Similarly, year factor can be held constant to study the achievement gap between race/ethnicity subgroups. Moreover, the interaction between race/ethnicity and district addresses the question about the impact of school districts on the reduction of the race/ethnicity gap. The study of interaction effects – which can reveal complex relationships between the response and factors – is a major goal of analysis of variance. "The possibility of examining interaction effects is the major payoff from having crossed factors in a data layout" (Singer, 1991, pg. 57). When two factors are crossed, we can investigate another type of interaction called bicomparisons (double differences). This interaction is estimated when the levels of two factors are changed. For example, the difference in performance between any two race/ethnicity subgroups over time will provide evidence of whether education reforms have had an impact in reducing the achievement gap. For the reasons explained above this study finds the framework offered by ANOVA as the most appropriate to explore the effects of the reforms, race/ethnicity and districts on school performance.

3.2.2. Square Combining Table

Table 3.2 shows the number of schools per year, by race/ethnicity and district. The first column contains the district number and the second column the number of schools located in that district. The third column contains the number of schools in the year 2001 that have students from Black race/ethnicity subgroup performing at level 3 and level 4. For example, in District 1 which has 14 public elementary schools there are only nine schools in year 2001 (row one, column three) that have students from Black race/ethnicity subgroup performing at level 3 and level 4. One reason for this discrepancy might be that five schools in District 1 have no 4^{th} grade students from Black

Table 3.2. Number of Schools per Year, by District and Race/Ethnicity

District	Number of Schools	Black					Hispanic					Asian					White				
		2001	2002	2003	2004	2005	2001	2002	2003	2004	2005	2001	2002	2003	2004	2005	2001	2002	2003	2004	2005
1	**14**	9	12	12	7	5	14	14	14	14	13	5	3	5	4	2	5	4	4	4	3
2	**25**	20	20	18	17	15	23	23	23	19	22	19	20	17	17	17	16	19	16	17	17
3	**14**	14	14	14	12	10	14	14	14	11	10	1	3	2	3	3	7	6	6	6	6
4	**13**	13	13	13	5	5	13	13	13	12	13	0	0	0	0	0	0	0	0	0	0
5	**8**	8	8	8	8	8	7	7	8	3	3	0	0	0	0	0	0	0	0	0	0
6	**17**	10	10	12	3	2	17	17	17	14	14	0	0	0	0	0	2	1	1	1	0
7	**13**	12	13	13	5	4	13	13	13	13	13	0	0	0	0	0	0	0	0	0	0
8	**18**	18	17	17	9	9	18	18	18	17	18	3	2	5	3	2	4	3	3	3	3
9	**19**	17	19	19	11	10	18	19	19	16	18	1	2	2	1	1	0	0	0	0	0
10	**30**	29	29	29	25	22	30	30	30	29	29	13	10	12	10	3	12	14	10	8	6
11	**23**	23	23	21	21	21	23	23	23	20	19	12	7	9	8	4	11	9	9	9	10
12	**11**	11	11	11	5	6	11	11	11	10	11	3	2	2	0	0	0	0	0	0	0
13	**18**	18	18	18	16	16	15	17	13	4	5	0	0	0	0	0	2	2	1	0	0
14	**19**	14	14	16	6	10	19	19	19	15	15	4	3	3	1	3	6	5	4	4	4
15	**23**	19	17	17	12	13	23	23	23	19	21	10	7	9	7	7	19	17	20	14	9
16	**8**	8	8	8	8	8	7	6	6	1	3	0	0	0	0	0	0	0	0	0	0
17	**17**	17	17	17	14	15	13	16	16	4	4	0	0	0	0	0	0	0	0	0	0
18	**12**	12	12	12	9	7	8	10	8	6	3	4	4	3	1	3	5	3	2	1	0
19	**22**	22	22	21	21	20	22	22	22	11	11	5	5	6	2	2	1	1	1	1	1
20	**22**	6	2	6	2	1	22	21	20	13	10	22	22	22	17	17	22	22	22	20	20
21	**22**	16	16	12	9	22	21	22	22	15	14	19	20	20	16	17	21	21	20	19	20
22	**24**	24	24	23	22	22	22	21	22	18	16	15	11	14	13	15	17	15	18	14	18
23	**14**	14	14	14	14	12	12	12	12	3	4	0	0	0	0	0	0	0	0	0	0
24	**20**	6	6	8	9	6	19	19	20	16	19	16	16	15	14	10	16	15	16	14	12
25	**22**	12	11	9	7	7	20	22	22	16	15	22	22	20	17	20	21	21	19	18	15
26	**20**	11	14	12	10	11	15	16	16	10	11	20	20	20	20	20	20	19	20	20	20
27	**30**	28	25	19	25	24	29	29	29	24	23	19	19	19	15	16	18	17	16	10	12
28	**22**	18	17	18	12	15	21	18	21	12	16	17	18	18	17	16	11	12	10	11	11
29	**22**	22	22	22	18	17	15	14	14	7	7	7	6	6	5	6	3	2	0	1	0
30	**22**	14	14	12	11	8	22	22	22	22	20	20	18	20	15	16	16	17	15	11	12
31	**40**	25	22	25	20	19	35	37	37	36	32	21	18	17	10	8	39	40	37	35	32
32	**12**	12	11	11	4	4	12	12	12	10	11	1	2	2	1	0	0	0	0	0	0

race/ethnicity subgroup. Another reason might be that even though there are students from Black race/ethnicity subgroup, the number of those tested is fewer than five and in this case, to ensure the privacy of students, their results are not reported separately but are combined with the next smallest group (another race/ethnicity subgroup with fewer than five students) and reported in a separate group called “Small Group Totals” (not included in the analysis). Table 3.2 reveals that there are districts whose schools in certain years have no data for White and Asian race/ethnicity subgroups. Thus, the number of schools with test results in each cell differs, resulting in an unbalanced dataset.

Square Combining Table (SCT) is a method of analysis that performs better than ANOVA with datasets that have missing data and are unbalanced (Godfrey, 1985). SCT analysis provides an additive fit to the data where the test result of each race/ethnicity subgroup is the sum of different components such as: an overall median which is constant for all data, a main effect for each factor (year, race/ethnicity and district), three two-way interaction effects (race/ethnicity-year, race/ethnicity-district and year–district), and residuals. There are two constraints imposed on the SCT fit, the sum of all levels in each main effect is zero and the median of the residuals is zero.

A separate SCT is generated for each one of the six effects and the steps to create one such table are shown using the year factor in the example below. Table 3.3 shows the SCT for the year factor which by having five levels (one for each year) results in a square table with five rows and five columns.

Table 3.3 Square Combining Table for the Year Factor

	2001	2002	2003	2004	2005
2001	0	2.4	7.7	5.5	15.0
2002	-2.4	0	4.7	3.0	11.7
2003	-7.7	-4.7	0	-2.4	7.0
2004	-5.5	-3.0	2.4	0	8.3
2005	-15.0	-11.7	-7	-8.3	0
Effects	**-6.1**	**-3.4**	**1.6**	**-0.4**	**8.4**

The SCT makes use of the medians of the differences between each level with the other levels. The year factor has five levels so ten differences should be calculated to account for all combinations. To calculate the difference between one level (e.g. year 2002) with another level (e.g. year 2001), the percentage of students performing at level 3 and level 4 for a given race/ethnicity subgroup at a given school for year 2001 is subtracted from the percentage of students performing at level 3 and level 4 for the same race/ethnicity subgroup from the same school but for year 2002. This difference contains 2,464 values (616 schools and four race/ethnicity subgroups for each school), but because of the missing values (815 in the year 2001 and 842 in the year 2002) 1,537 values are obtained (927 missing). The median of these values is 2.4 and represents the change in performance in the city from year 2001 to 2002. It is assigned to the cell in the first row and second column of Table 3.3. The other values on the first row are the median differences of the remaining years with the year 2001. The upper triangle of Table 3.3 contains the values for all ten combinations, diagonals are all zero (the difference of each year with itself) and the values in the lower triangle are symmetrical to the upper one but with the opposite signs. The main effects are calculated as the mean of each column in the SCT. The effects satisfy the constraint of their sum being zero. In the same manor, the main effects are calculated for the district and race/ethnicity factor and their SCT have 32 and four rows and columns respectively. The three main effects are subtracted from the raw values and the remaining differences are used to calculate the three two-way interaction effects: race/ethnicity-year, race/ethnicity-district and year–district by utilizing SCT. To calculate the overall median, first the three interaction effects are subtracted from the remaining differences mentioned above. Then, the median of the resulting values is calculated. This overall median represents the school performance over the entire period and the main and interaction effects are deviations from it. The residuals are obtained by subtracting the overall median from the resulting values above. By construction the median of the residuals is zero.

SCT method uses median in two instances: finding the median of the differences between levels of a factor and finding the overall median. In contrast to the mean, median is less affected by the outliers and therefore SCT is a robust method. In a balanced dataset, if means of the differences are calculated instead of medians, and an overall mean instead of median, an ANOVA fit is obtained with residuals having a mean zero. All the analysis possible with ANOVA can be done with the SCT method. Moreover, SCT is robust and more efficient with unbalanced datasets and datasets with missing data, as is the case with the dataset in this study, which makes it a better method to be used compared to ANOVA.

3.2.3. Bisquare Weights

The unit of analysis in the SCT fit is the percentage of students of a race/ethnicity subgroup performing at level 3 and level 4 in each school. The number of students from each race/ethnicity subgroup that took the test differs and to take that into account, that number is used to weight each test result. These weights are used in the first SCT fit and the residuals from this fit are used to calculate the weights in the successive fittings. The later weights are calculated as the ratio of the residual in the previous fit to some measure of the general size of the residuals in that fit, such as:

$$u_i = \frac{residual(i)}{c * S}$$

where *i* is the unit of analysis, *c* is a numerical constant and *S* is the median of the absolute deviations of the residuals and thus measures the spread the residuals. The result from the above ratio is used with the "Bisquare Weight" function to calculate the weights *w*:

$$w(u) = (1 - u^2)^2 \qquad u^2 < 1$$

$$w(u) = 0 \qquad u^2 >= 1$$

Test results that have small residuals get larger weights and as the deviations from the fit increase, their weights decrease up to the cutoff value of 1 after which they are zero. The value of *c* used in this study is 6 and it determines how soon the weight becomes zero. Unusual results that have residuals more than 6**S* from the last SCT fit will get zero weights. The value of *S* is used in this study as a measure of convergence which gets to a stable value after a few iterations. This weighting sequence which uses the residuals from the previous fit to calculate the weights of the next iteration "offers an effective fitting procedure that is resistant and robust of efficiency" (Mosteller and Tukey, 1977, pg. 378). The weights described above are also used in the analysis in chapter 4.

3.2.4. Multiple Comparisons

The next step in the analysis is to decide whether the differences between different effects estimated from SCT fit are real or random variations. The example with the year factor is used to illustrate testing the significance of the difference in performance between years. The year effects are -6.1 and -3.4 for year 2001 and year 2002, respectively, and their difference is 2.7 (Table 3.3). A t-test is used to conclude whether this difference is statistically significant at a selected error rate, such as 0.05, and a 95% confidence interval of the difference is constructed. Since the null hypothesis being tested with the t-test is that the difference in performance between the two years is zero, then the difference is significant if the confidence interval does not include the zero. Year factor has five levels and therefore to account for all year differences ten comparisons need to be performed. If each of the ten comparisons is made using a t-test with an error rate of 0.05 for each test, the probability of making a Type I error (rejecting a null hypothesis when it is true) for all comparisons becomes bigger than 0.05. It means that some comparisons will be declared as significant when in reality they are not at 0.05 error rate.

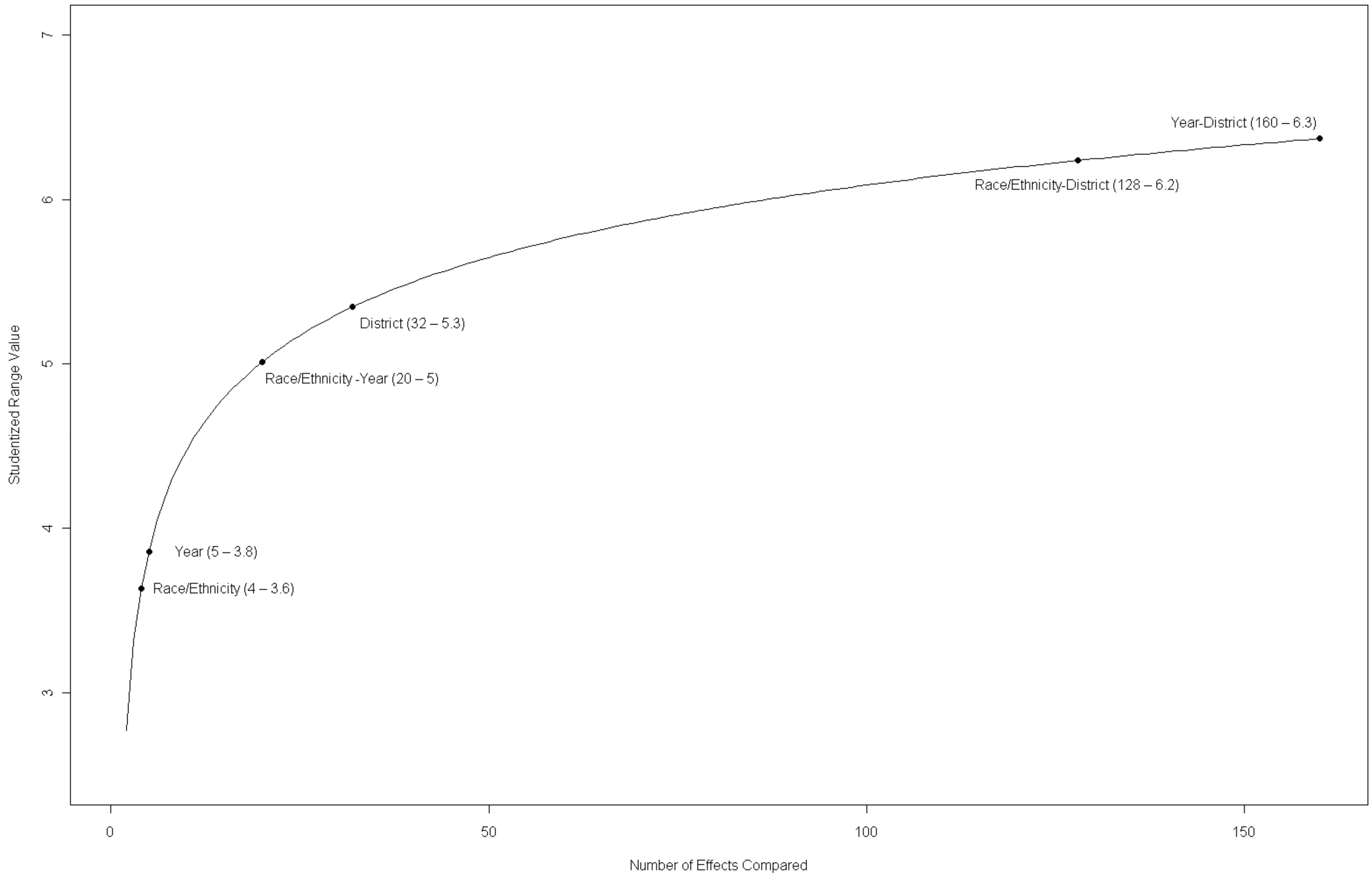
Studentized Range Distribution deals with the multiple comparison issue by increasing the critical value in accordance with the number of effects being compared (Tukey and Hoaglin, 1991). Figure 3.2 shows the Studentized Range Distribution that is used in this analysis which depends on the confidence level chosen (0.05 in this case) and the degrees of freedom of residuals (7,139 in this case). The horizontal axis shows the number of effects being compared and the vertical axis shows the Studentized Range value. The critical value for each of the main and interaction effects is positioned on the distribution and labeled in parenthesis with the number of effects being compared and the respective critical value.

As in a t–test, the actual difference between the effects is divided by the estimated standard error of their difference when the null hypothesis of no difference is true. The estimated standard error of the difference between the effects of years 2001 and 2002 is equal to the square root of the sum of the ratio of the mean square error (MSE) from the SCT fit with the number of values for year 2001 and 2002 as shown below:

$$SE = \sqrt{\frac{MSE}{n1} + \frac{MSE}{n2}}$$

where *n1* is the number of non missing data for 2001 and *n2* for 2002. The MSE is calculated by dividing the sum of the squared residuals (1,688,745) with the degrees of freedom of the residuals from the SCT fit.

Figure 3.2. Studentized Range Distribution



Degrees of freedom (df) are calculated as shown below:

- df = (total number of observations) – (df used) – (missing values)
    - total number of observations = 12,320
    - df used = grand median + df (year) + df (race/ethnicity) + df (district) + df (race/ethnicity-year) + df (race/ethnicity-district) + df (year-district)

      = 1 + 4 + 3 + 31 + 12 + 76 + 124 = 251
    - missing values = 4.930

  df = 12,320 – 251 – 4,930 = 7,139
- MSE = 1,688,745 / 7,139 = 237

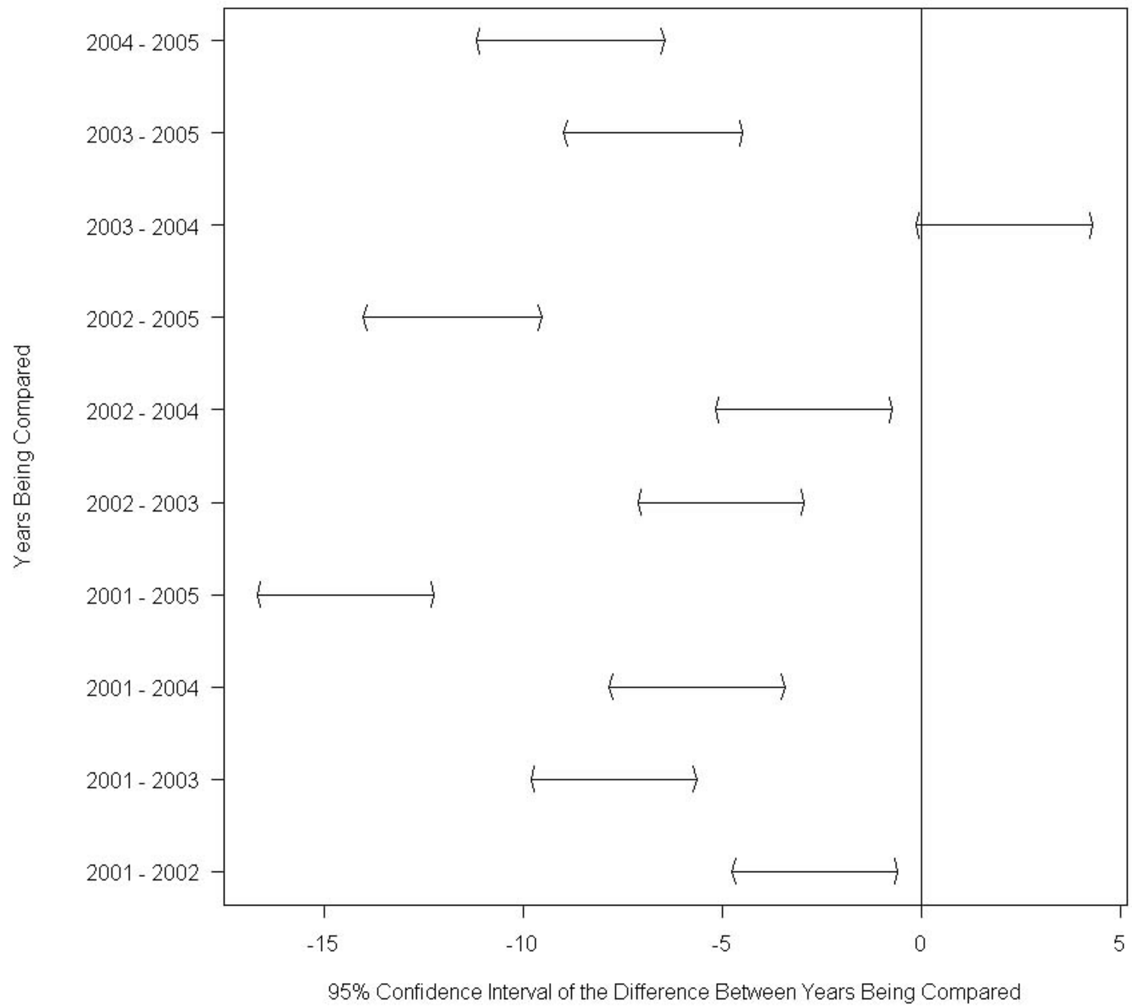
The standard error of the difference under the null hypothesis of no difference between years 2001 and 2002 is:

$$SE = \sqrt{\frac{237}{1{,}648} + \frac{237}{1{,}620}} = 0.54$$

The ratio between the actual difference of 2.7 with 0.54 is equal to 4.98 which is bigger than the critical value from the Studentized Range Distribution of 3.85 at 0.95 confidence and 7,139 degrees of freedom, concluding that there was a significant increase in the school performance from year 2001 to year 2002.

In Figure 3.3 the horizontal axis shows the 95% confidence interval of the difference between years and the vertical axis shows years being compared. The 95%

Figure 3.3. Multiple Comparisons Between Years



confidence interval of the difference between years 2003 and 2004 includes zero which means that the null hypothesis of no difference between these two years cannot be rejected. It can be concluded that, with the exception of year 2004 when there was an insignificant change in performance compared to year 2003, there has been an increase in performance in New York City during the period of this study. The same procedure is followed in comparing main effects for race/ethnicity, districts and three two-way interactions.

The code to implement the SCT fit, multiple comparisons and to produce different graphs is written using the statistical software R (R Development Core Team, 2008).

3.3 Results

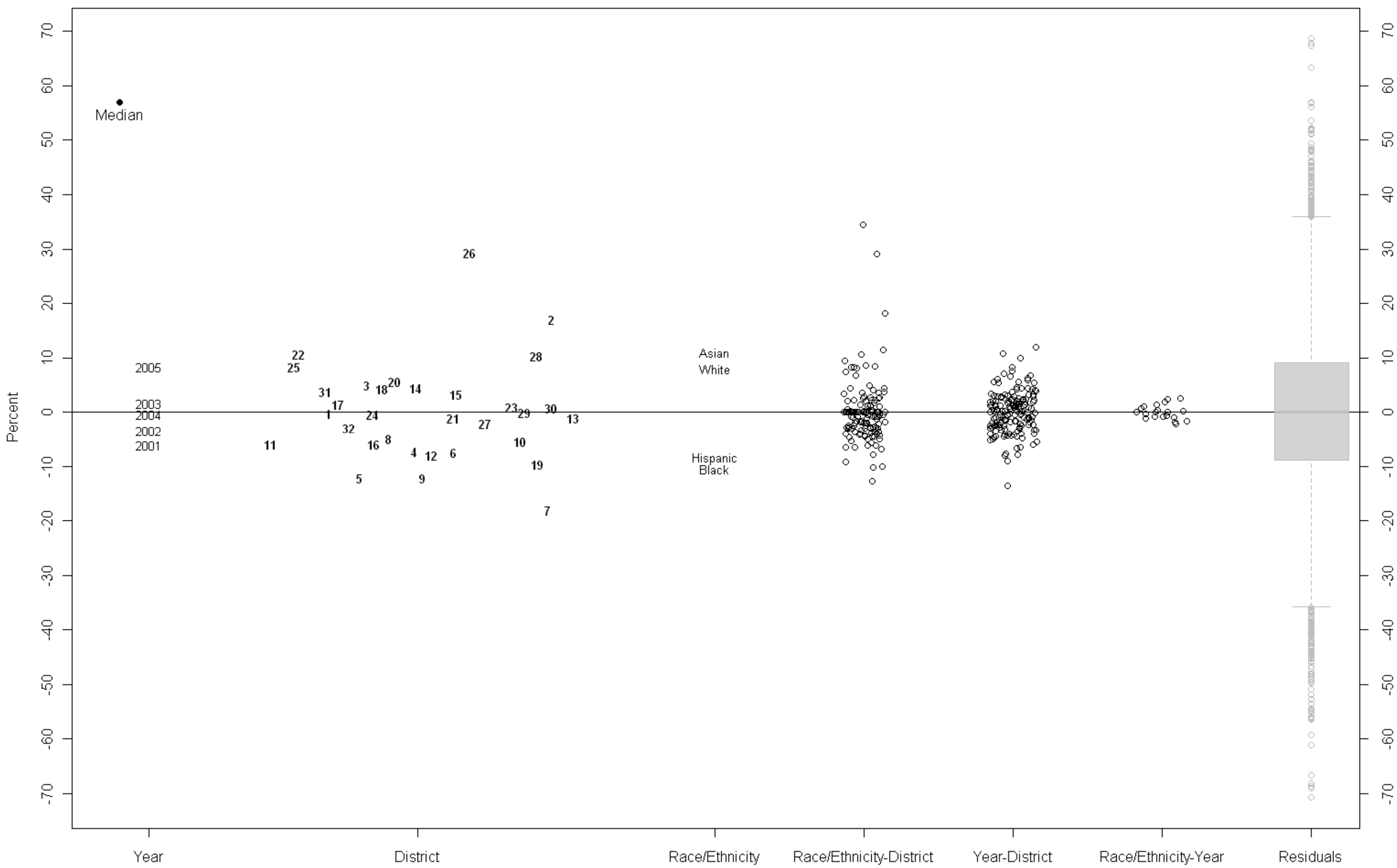
The amount of variability associated with each factor and their two-way interactions are shown in the ANOVA Table (Table 3.4). The effects are displayed in a decreasing order based on mean square values showing that race/ethnicity is the component that affects the most school results in the ELA exam taken by the 4^{th} grade students in public elementary schools in New York City.

Table 3.4 ANOVA Table

Factors	Df	Sum Sq	Mean Sq
Race/Ethnicity	3	630715	210238
Year	4	191891	47973
District	31	1020916	32933
Race/Ethnicity-District	76	91169	1200
Year-District	124	66722	538
Race/Ethnicity-Year	12	4529	377
Residuals	7139	1644233	230

A graphical summary of the results from the SCT fit for ELA are displayed in Figure 3.4 where the median overall performance in ELA exam of 4^{th} grade students in public elementary schools in New York City during the period of study is 56.9 percent.

Figure 3.4. Results from the SCT Fit

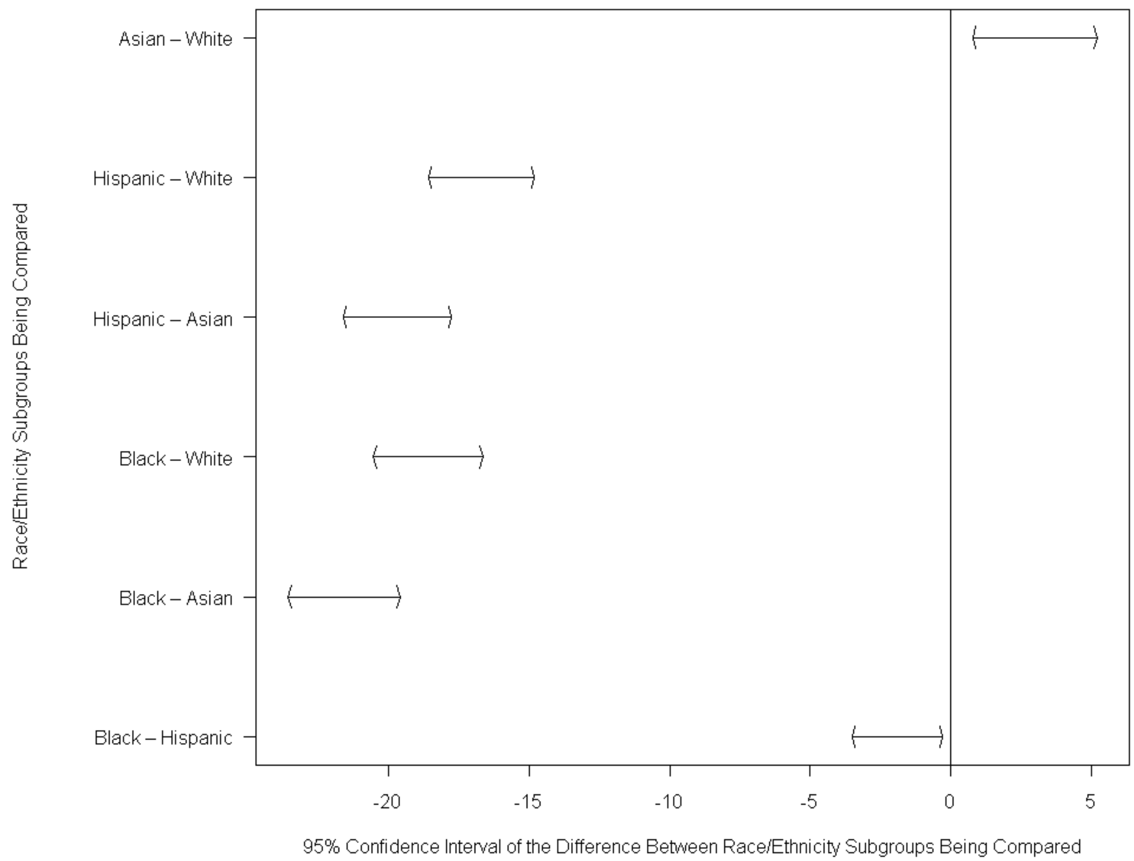


Race/ethnicity subgroup effects, displayed in the center of the graph, show that in New York City there is a big gap between the lowest performing subgroups of Black and Hispanic and the highest performing subgroups of White and Asian. Thus, within the lowest performing subgroups Hispanic students perform better than Blacks (1.9 percent), and within the highest performing subgroups, Asian students perform better than Whites (3.1 percent). The biggest race/ethnicity gap is between the subgroup of Asian and Black (21.2 percent) and the smallest between subgroup of White and Hispanic (17.1 percent). Multiple comparison results between the race/ethnicity subgroups displayed in Figure 3.5 show that none of the 95% confidence intervals contain zero, meaning that all race/ethnicity subgroups effects differ significantly from each other.

Year effects, displayed in the left part of Figure 3.4 show that the performance in ELA exam of 4^{th} grade students during the period of study has been increasing, except between years 2003-2004, when the multiple comparison results show that there was an insignificant change (see Figure 3.3 on section 3.2.4).

Figure 3.6 displays district effects in increasing order of their magnitude from left to right. District numbers are shown at the bottom part of the figure, whereas the respective values of district effects are shown at the top. Results from multiple comparison analysis on district effects show that there are 29 comparisons that differ significantly. The difference between each one of these pairs of districts is bigger than the threshold difference (which is 26 percent) under the null hypothesis of no difference.

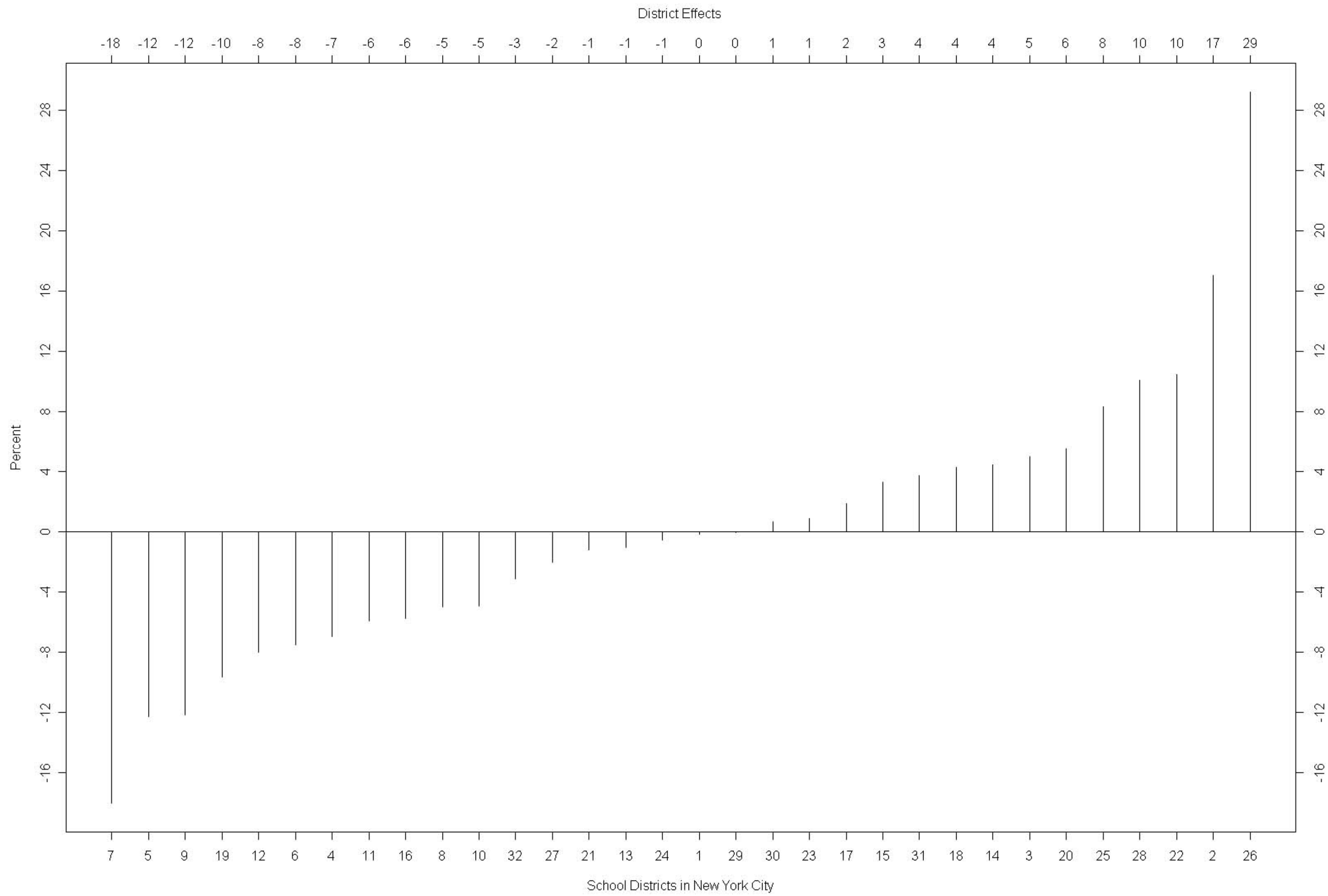
Figure 3.5. Multiple Comparisons Between Race/Ethnicity Subgroups



District 26 is the highest performing district and is statistically different from each district that has an effect smaller than 3 percent (the 22 districts displayed to the left of district 31 in Figure 3.6). District 2, the second highest, is significantly different from the four lowest performing districts. The four remaining significant differences (out of 29) are between the high performing districts 22, 28 and 25 and the lowest performing district 7.

Figure 3.7 displays the spatial distribution of district effects where the worst district 7 (with an effect of -18 percent) and the best districts 2 and 26 (with effects of 17

Figure 3.6. District Effects



percent and 29 percent, respectively) are shown in separate categories. The numbers in parenthesis in legend indicate the total number of districts in each category. Even after taking into account the race/ethnicity effect, schools in Upper Manhattan, Bronx and central Brooklyn with majority of enrolled students from Black and Hispanic subgroups, are located in the worst districts in the city.

Figure 3.7. Map of District Effects

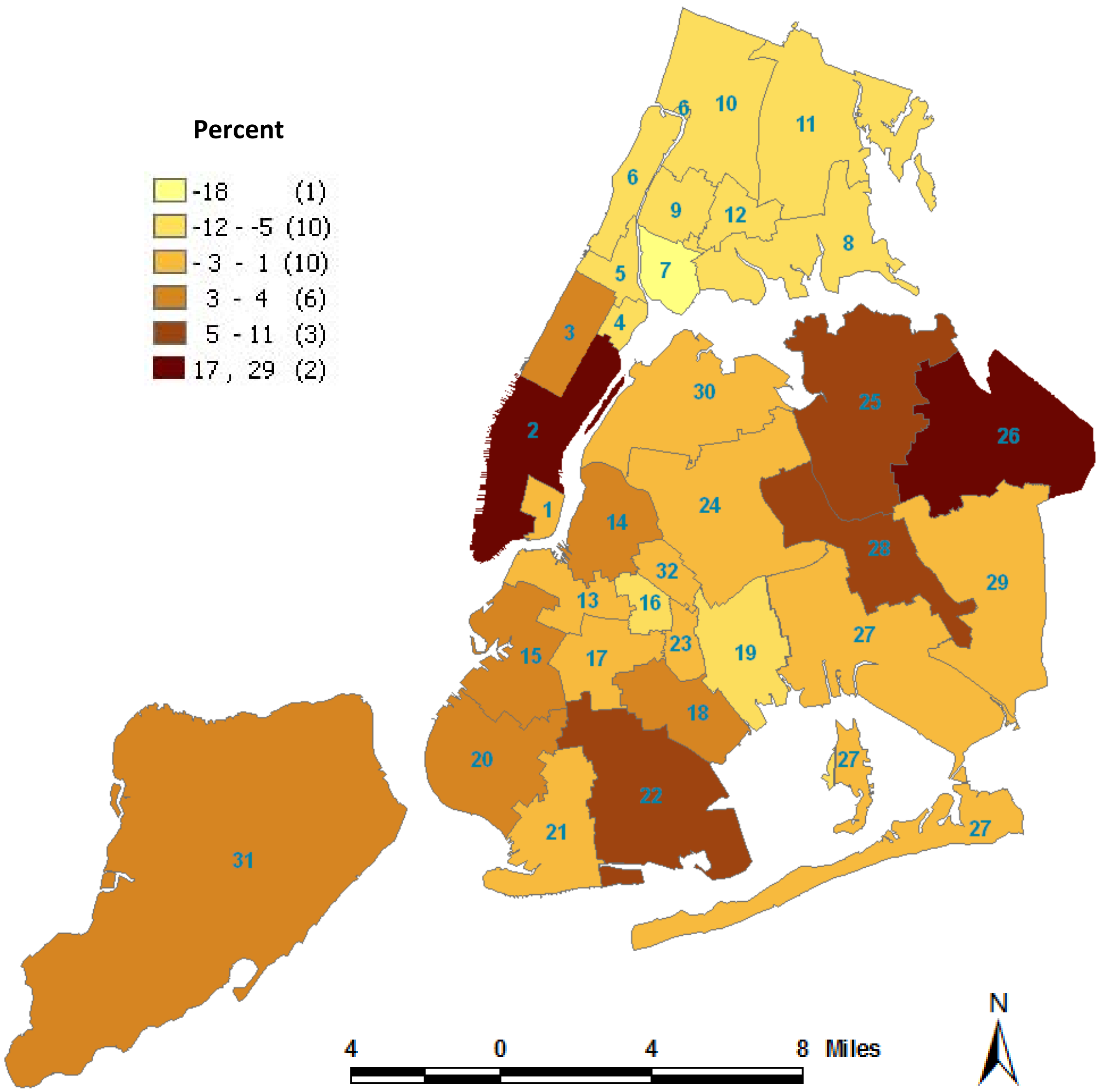


Figure 3.8 is a plot of race/ethnicity and district interaction effects which are ordered by the value of district effects from negative to positive (district numbers are shown on the horizontal axis). Each race/ethnicity subgroup is displayed with the initial letter and different color. Multiple comparison analysis shows that there is no significant difference between these interactions. It indicates that after we have taken into account district and race/ethnicity effects, there is no difference in the performance of race/ethnicity subgroups across districts.

Figure 3.9 is a plot of year and district interaction effects where each year is displayed with the last digit and a different color. Multiple comparison analysis shows that there is no significant difference between these interactions, indicating that once the year and district effect are controlled for, there is no difference in the performance of each district over time.

Race/ethnicity and year interaction effects are shown in Figure 3.10. Multiple comparison analysis shows that there is no significant difference between these interactions, indicating that there is no change in race/ethnicity performance over time and no gap reduction once the year and race/ethnicity effects are taken into account.

Figure 3.8. Race/Ethnicity and District Interaction Effects

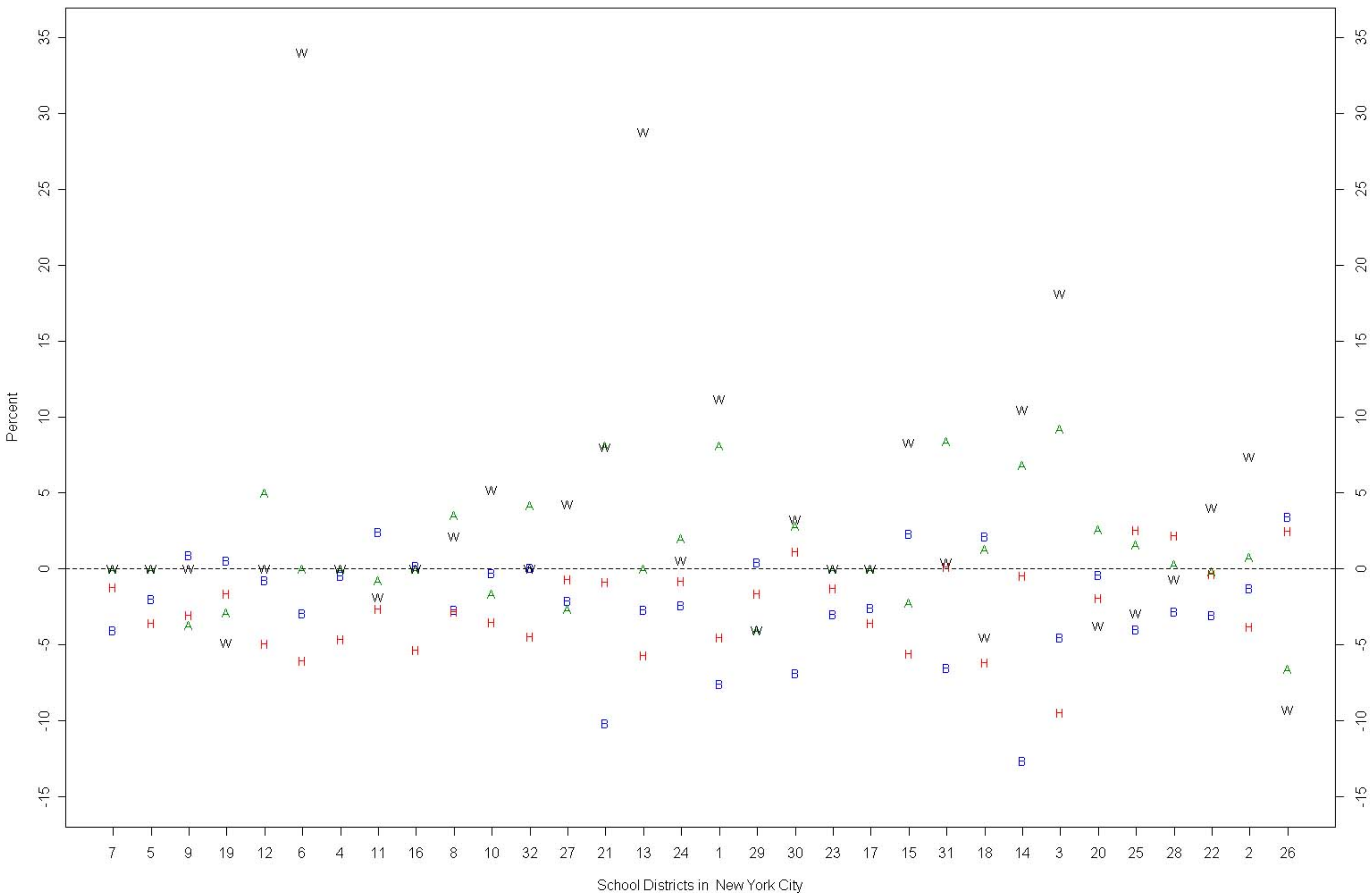


Figure 3.9. Year and District Interaction Effects

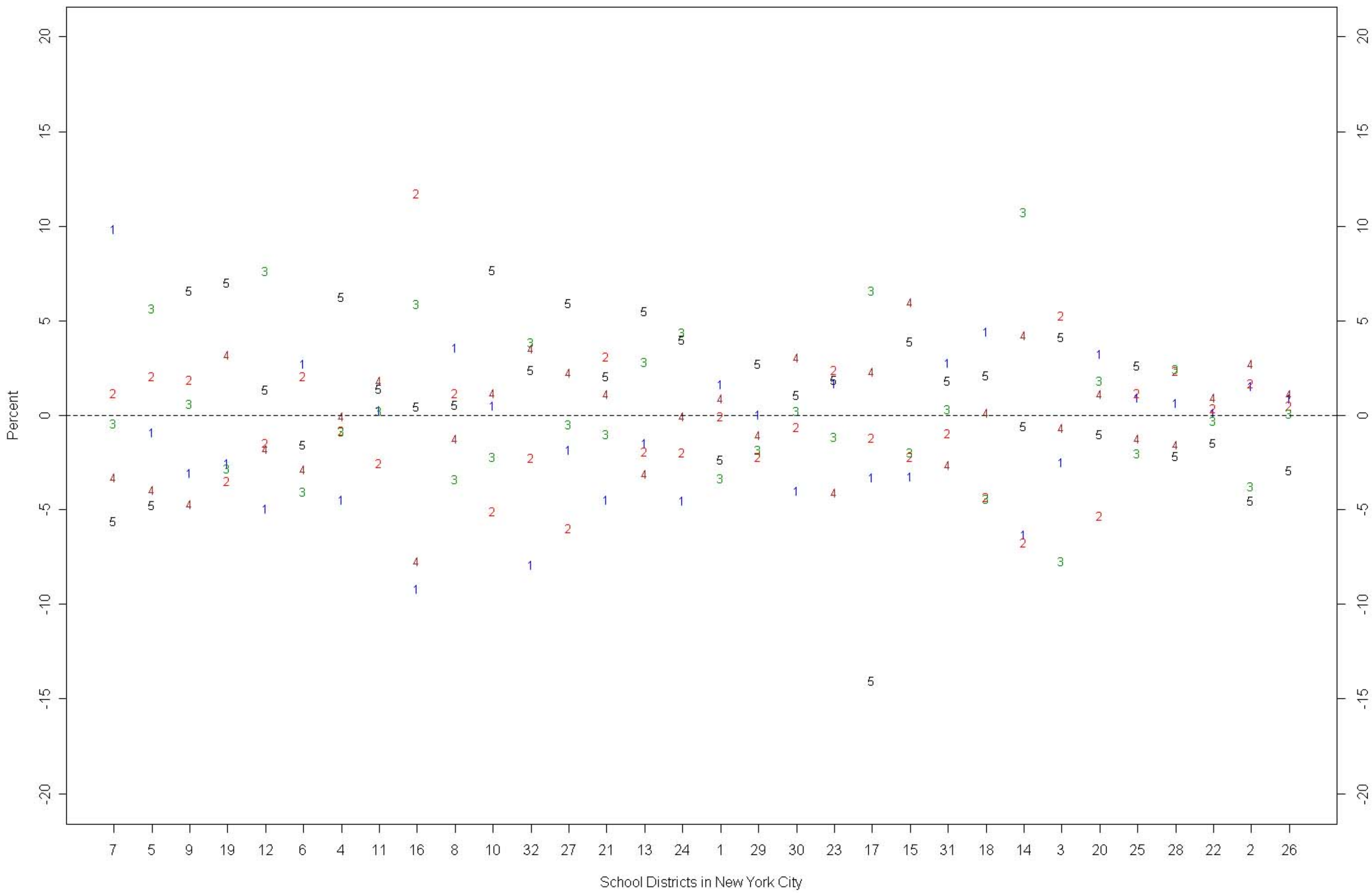
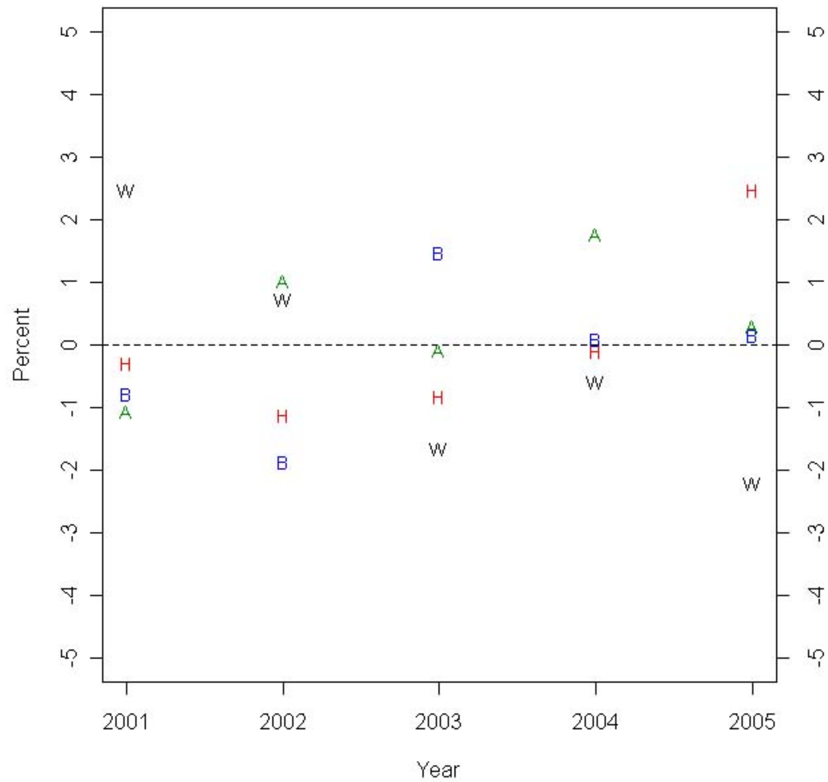


Figure 3.10. Race/Ethnicity and Year Interaction Effects



3.4. Summary

Findings from the analysis in this chapter show that race/ethnicity is the most important component affecting the performance of public elementary schools in New York City during the period of study. The achievement gap between the higher performing racial/ethnic groups (Asian and White) and the lower performing ones (Black and Hispanic) remains substantial. The low performing schools located in Upper Manhattan, Bronx and central Brooklyn enroll high percentages of Black and Hispanic students, resulting in a geographical concentration of low educational achievement. Even after the race/ethnicity effect is taken into account, there is an indication of spatial inequality in educational achievement revealed by the negative effect of school districts in Upper Manhattan and south Bronx on the performance of their schools compared to the positive effect of school districts in north-east Queens and Lower Manhattan.

There was an increase in public elementary school performance throughout the period (before and after the implementation of the reforms) except for the transition academic year 2003-2004. The improvement has been uniform across city indicating that there was no change on the impact of districts over time. Despite this general increase in school performance and even though its reduction has been the main goal of the education reforms, results indicate no decrease in the race/ethnicity gap after the reforms. Similarly, districts had no impact on reducing the race/ethnicity achievement gap.

Improvement on school performance has not always been associated with racial/ethnic gap reduction. This chapter examined results from two academic years after new reforms were implemented in New York City's schools. The first academic year, 2003-2004 is the only one, during the study period, when there was no increase in school performance which may be related with the issues during the implementation of reforms in this transition year. The second academic year 2004-2005 experienced the largest increase in performance during the five year period but such an increase was not associated with a reduction in achievement gap. Future research should consider more years to better assess the impact of the reforms on gap reduction, since the effect of reforms may require some more time to materialize.

Chapter Four

Analysis of Spatial Pattern of School Performance

The increased focus on schools from policy makers and researchers as the main agent of change presents a challenge for geographers to investigate the impact of the geographical context on school performance. School achievement has a geographical component because it is affected by the characteristics of the school catchment area, its surrounding areas, and by the management and policies of the school district to which the school belongs. This chapter uses student and school characteristics to explain the spatial pattern of school performance from a geographical perspective. It addresses and answers the following three questions. What are the student and school characteristics that explain the spatial patterning of school performance and does their relationship change over time? Does the relationship between school performance and student/school characteristics vary in space? Is there a spatial dependence in school performance and what is the spatial process generating it?

To answer the above questions this chapter uses school level data and utilizes spatial regression models, in addition to Ordinary Least Regression, to assess the relationship between school performance and student/school characteristics. Such models are also used to discover the spatial process generating the spatial dependence on school performance. This chapter introduces to the geographical educational literature the Jackknife technique to examine the spatial variation of the relationship between school performance and student/school characteristics.

4.1. Data

As noted earlier, NCLB and New York City education reforms promoted transparency by requiring schools to inform the public about school performance in standardized tests as well as characteristics of the school and its students. These data are made public in Annual Report Cards. Information for each school includes students' characteristics such as the number of students eligible for free lunch, students' attendance, number of students that are recent immigrants, the proportion of students of each race/ethnicity subgroup, special education and English language learners enrollment. School characteristics include information about school size, student-teacher ratio, teacher qualification, their experience and tenure. The availability of these data for each year makes it possible to study the factors that shape school performance over time as well as to assess the effect of reforms. The spatial and non spatial data used in this chapter were provided at request from the New York City's Department of Education.

4.1.1. School Zones

The spatial unit of analysis is important in spatial data analysis because the calculation of different spatial statistics is conditioned on the neighborhood around each spatial unit. The selection of neighbors depends on the way the spatial unit is defined. The spatial units of analysis in this chapter are the polygons representing the school zones (Figure 4.1).

Figure 4.1. Public Elementary Schools and Their Zones

Defining school catchment area has been the focus of several studies (Flowerdew and Pearce, 2001; Martin and Atkinson, 2001; Pearce, 2000). The relationship between the school location and its catchment area is similar to the relationship between a facility and its service area such as a library or hospital. In a walkable city as New York the assumption that people will use the closest facility is a reasonable one and is used, for example, in defining the library service areas by creating Thiessen polygons around the library locations (Japzon and Gong, 2005).

In New York City a child is generally registered at the elementary school located in his/her school zone defined by the New York City's Department of Education. The school zone for each child is determined by the child's home address. In addition to the regulatory reasons, most students for proximity reasons attend the school in their zone. Some parents, however, look for public alternatives outside of their zone, including gifted and talented and special education programs. There are also cases of parents not using their real address but one belonging to a zone with a school they prefer. It is more difficult for students to register in schools outside of their zone because the right to register is given first to those who reside in the school zone. Although the actual flow of students is complex, it is a fair assumption that, based on regulatory and proximity reasons, most students will register in their zone school. Therefore, the school zone is used as a proxy for the residential locations of the students that attend that school.

4.1.2. School Performance

School performance is the dependent variable of the analysis in this chapter. The performance in standardized state tests of 4^{th} grade students in public elementary schools is used to represent school performance. More specifically, the dependent variable is the percentage of 4^{th} grade students performing at level 3 (the proficiency level defined by the state) and level 4 (above the state proficiency level) in the ELA annual state exam. Table 4.1 shows summary statistics of the dependent variable for the 4^{th} grade students in public elementary schools for each academic year starting from 2000-2001 to 2004-2005. The average performance has increased from 46 percent in 2001 to 61 percent in 2005, with a slight decrease from 2003 to 2004. The same pattern is observed as well at the 25th percentile (from 31 percent in 2001 to 48 percent in 2005) and 75th percentile (from 60 percent in 2001 to 75 percent in 2005) suggesting that the improvement in school performance has occurred for the majority of schools. Throughout the period, the school performance ranges from the lowest of 5 percent to the highest at 100 percent. The number of schools per year differs because four schools (out of 616 schools included in this study) had missing data and therefore were excluded.

Table 4.1. Percentage of Students Performing at Levels 3 and 4

	Year				
Percentiles	**2001**	**2002**	**2003**	**2004**	**2005**
25%	31	33	40	36	48
50%	43	46	54	50	61
75%	60	63	68	66	75
Mean	46	49	54	52	61
St. dev	19	19	19	19	18
N of obs.	612	616	612	615	616

4.1.3. Student and School Characteristics

In educational research student and school characteristics are commonly used to explain the school performance (Powers, 2003). Table 4.2 shows the list of selected school and student characteristics used as independent variables in this chapter that are available either at school or grade four levels for each year from 2001 to 2005.

Table 4.2. Independent Variables

Independent Variables	Available for	
	School	Grade 4
Students characteristics		
Eligible for Free Lunch (%)	X	
Attendance (%)	X	
Recent Immigrants (%)	X	
English Language Learners (%)		X
School characteristics		
School size (# of students)	X	
Teachers with master degree (%)	X	
Teachers more than 2 years in the same school (%)	X	
Teachers with greater than 5 years experience anywhere (%)	X	
Teachers fully licensed and permanently assigned to the school (%)	X	
Students/Teachers Ratio	X	
Other		
Special Education (%)		X

Some school variables selected for the analysis in this chapter represent students' socioeconomic background such as: percentage of enrolled students in a given school that are eligible for free lunch, percentage of the recent immigrants measured as the percentage of enrolled students in a given school that immigrated within the last three years, and the percentage of 4th grade students that took the exam and were English language learners. In Annual Report Cards, attendance - calculated as the percentage of days attended by all students - is reported as student characteristic and it is considered the same way in the table above, although it may also represent some measure of school characteristic. Student characteristics provide proxy measures of the socioeconomic status of the schools' population and as such represent factors that are largely beyond the schools' control (Powers 2003).

School characteristics include variables that can be affected by school and district policies such as: school size measured as the number of students enrolled in the school in a given academic year, the number of students per teacher measured as the ratio of the school enrollment to the number of teachers and several teacher characteristics representing teachers' education, experience and tenure. Teachers' characteristics, measured as percentages, include teachers with master degree or higher, teachers with more than two years in the same school, teachers with more than five years of experience anywhere and teachers fully licensed and permanently assigned to the school. Another student characteristic that may affect school results, included as a control variable, is the percentage of 4^{th} grade students in special education that took the exam.

Table 4.3 shows the mean and standard deviation of independent variables for each year over the period of this study. Results indicate that at the mean, students'

Table 4.3. Mean and Standard Deviation of Independent Variables by Year

Independent Variables	Year				
	2001	2002	2003	2004	2005
Students characteristics					
Eligible for Free Lunch (%)	74	73	73	72	70
	(25)	(24)	(24)	(24)	(24)
Attendance (%)	92	93	93	93	93
	(2)	(2)	(2)	(2)	(2)
Recent Immigrants (%)	7	7	7	8	8
	(5)	(5)	(5)	(5)	(5)
English Language Learners (%)	6	5	4	4	4
	(6)	(6)	(5)	(5)	(6)
School characteristics					
School size (# of students)	816	795	776	733	700
	(332)	(324)	(310)	(295)	(286)
Teachers with master degree (%)	76	75	79	82	84
	(12)	(13)	(10)	(9)	(8)
Teachers more than 2 years in the same school (%)	63	65	67	68	69
	(15)	(13)	(11)	(11)	(11)
Teachers with greater than 5 years experience anywhere (%)	54	54	54	55	56
	(13)	(12)	(12)	(12)	(12)
Teachers fully licensed and permanently assigned to the school (%)	85	85	91	99	99
	(13)	(11)	(7)	(2)	(2)
Students/Teachers Ratio	15	16	16	15	15
	(3)	(3)	(3)	(5)	(2)
Other					
Special Education (%)	14	13	13	11	13
	(8)	(7)	(7)	(8)	(8)

Note: Standart deviation in parentheses.

characteristics are stable over time, whereas in general school characteristics have improved, especially after the implementation of reforms. Among school characteristics, average school size has decreased from 816 students in 2001 to 700 students in 2005, a decrease that accelerated after 2003 suggesting that the reforms may have had a role. Similarly, among teachers' characteristics the percentage licensed and the percentage with master degree have increased, particularly after 2003.

Previous research has used census data, either alone or in addition to school level characteristics, to examine the impact of neighborhood on school performance. This chapter does not use census data for several reasons. One reason is that about 20 percent of students in New York City attend hundreds of private and parochial schools (Ravitch and Vitteriti, 2001). Therefore, census data may not reflect the characteristics of families and students attending public schools. For example, in a relatively wealthy area students from higher income families may attend private schools whereas students from lower income families may attend public schools. Thus, census data are not an accurate representation of the families of students attending public schools. Consequently, children who attend private schools are drawn from families that are more affluent than average, and/or those that attach a very high priority to education. Their absence from public schools means that the socioeconomic background of pupils at the school is different than census data would indicate, though to an unknown degree. Therefore, due to the existence of private schools, analysis would suggest that public schools are not performing as well as would be expected from the characteristics of census data. Another reason is that census data are available for year 2000 whereas the period of analysis for this chapter spans from 2001 to 2005, raising concern about their accuracy over time.

Finally, the spatial units used in census data do not correspond to the boundaries of school zones. Moreover, the school zone may overlap with more than one census units. While census units may differ in their characteristics from each other, spatial variation of socioeconomic characteristics may exist even within a census unit itself. Therefore, the interpolation of the census data on the school zone will add another level of inaccuracy.

For all the above reasons this chapter does not use census data to assess the impact of family and parents' characteristics but instead uses only school level data such as the number of students eligible for free lunch as a proxy of family socioeconomic characteristics, students' attendance as a proxy for parents' involvement in students' education and students that are English language learners and recent immigrants as proxies for demographic characteristics.

4.2. Methods and Results

This study explores the spatial pattern of school performance and uses students' socioeconomic background and school characteristics to explain that pattern. It implements a combination of exploratory and regression analysis. The spatial pattern of school performance is investigated using exploratory spatial data analysis (ESDA) methods such as Moran's I and local indicators of spatial autocorrelation (LISA). Ordinary least squares (OLS) regression is used as the first step to identify whether spatial pattern of school performance is explained by the spatial pattern of independent variables. If there is an indication of spatial autocorrelation of residuals then spatial regression analysis are applied. A decision procedure is implemented to determine the type of spatial regression, error or lag model, used to account for the spatial dependence on the residuals. Spatial heterogeneity in the relationship between dependent and each independent variable is investigated using the Jackknife analysis.

4.2.1. Exploratory Spatial Data Analysis

The investigation of the spatial pattern of school performance is implemented through the use of ESDA methods. ESDA is an extension in the spatial context of exploratory data analysis (EDA) which is an approach to analyze data for the purpose of discovering patterns in them and suggesting hypothesis by allowing the data itself to reveal its underlying structure (Tukey, 1977). ESDA techniques intend to discover patterns in the data by imposing as little prior structure as possible, identify atypical locations or spatial outliers, discover patterns of spatial association, clusters or hot spots, and suggest spatial regimes or other forms of spatial heterogeneity (Messner and Anselin, 2003). This perspective in the analysis of spatial data is gaining ground in other areas of research such as crime analysis (Baller et al 2001, Murray et al 2001), child poverty (Voss et al, 2006), health (Dragicevic, Schuurman and Fitzerald, 2004), but has received little attention in the study of the geography of educational achievement.

4.2.1.a. Spatial Weights Matrix

Spatial weights are used to represent the spatial relationship between units of analysis. The creation of the spatial weights matrix is the result of the definition of the neighborhood around each spatial unit and the weight assigned to each neighbor. In this study, the spatial neighbors of each school zone are considered those zones that have simple contiguity (sharing a common border) of a “queen” variety (both borders and common vertices). The weight assigned to each neighbor are row standardized which means that when a school zone is contiguous, for example, with five other school zones, each one of them will get an equal weight of 0.2. The spatial weights matrix created based on the two decisions above is used in all spatial analysis in this chapter.

4.2.1.b. Moran’s I

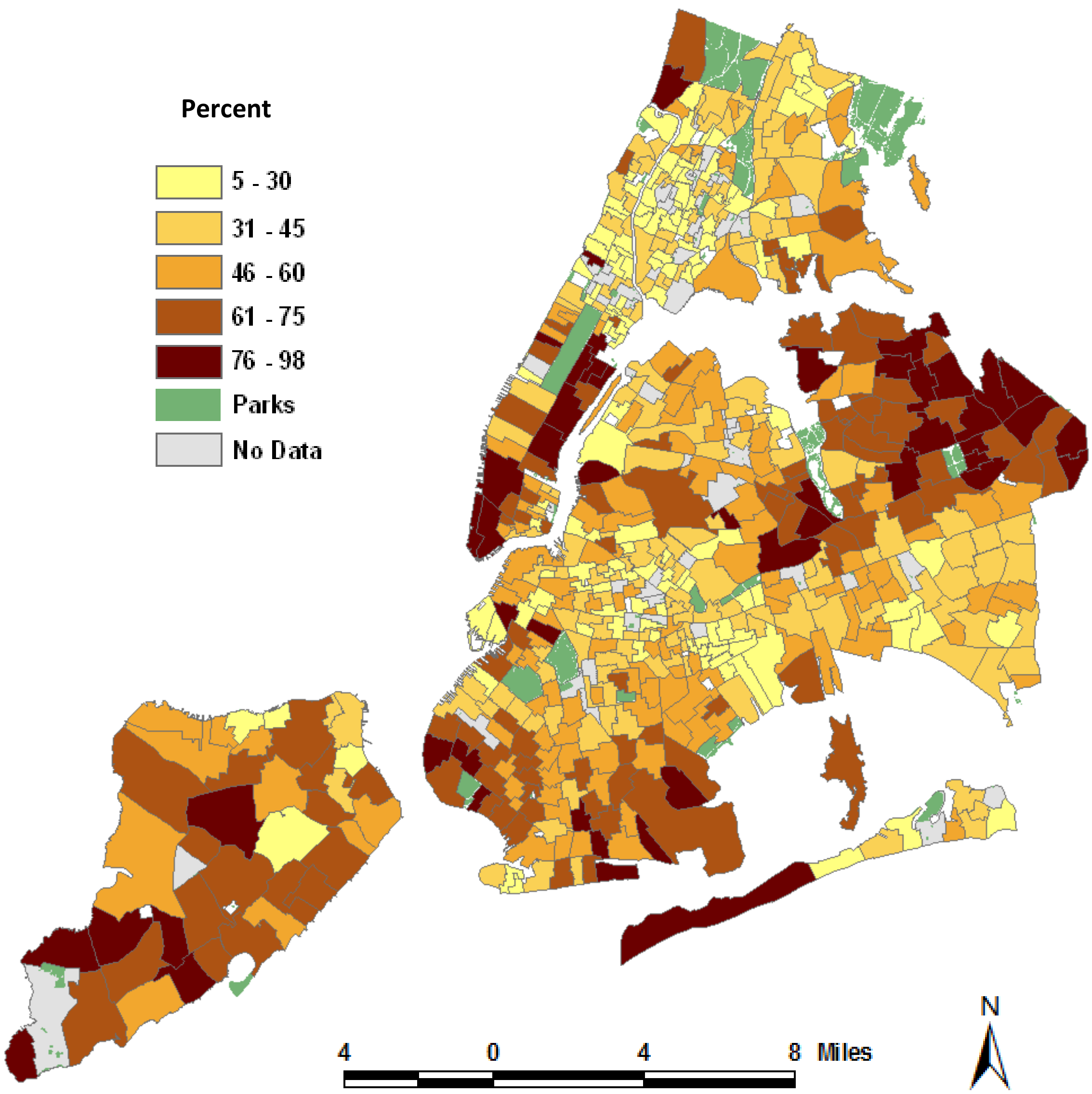
The spatial distribution of school performance in New York City exhibits a clustered pattern (illustrated in Figure 4.2 for year 2001), where schools with similar performance tend to be close in space. The map in Figure 4.2 shows a concentration of high performing schools in lower Manhattan, Upper East Side, north-east Queens and Staten Island. The low performing schools are concentrated in upper Manhattan, south and central Bronx and central Brooklyn.

One of the spatial statistics that quantifies the above concentration of low and high performing schools in space is Moran’s I (O’Sullivan and Unwin, 2003), which is a simple translation of a nonspatial correlation measure to a spatial context. Moran’s I is a single measure describing the general extent of spatial clustering of an attribute across the region of study. It is used to test the null hypothesis of spatial randomness against the alternative hypothesis of the presence of spatial autocorrelation. Moran’s I, as a measure of spatial association, is conditioned on the specific neighborhood structure, represented in a spatial weights matrix. It is calculated as follows:

$$I = \frac{n}{\sum_i \sum_j wij} \frac{\sum_i \sum_j wij(Xi - \overline{X})(Xj - \overline{X})}{\sum_i (Xi - \overline{X})^2}$$

where *n* is the number of spatial units indexed by *i* and *j*, *X* is the variable of interest, $\overline{X}$ is the mean of *X* and *wij* is a matrix of spatial weights.

Figure 4.2. Percentage of Students Meeting the Standard in 2001



The global Moran's I for each year from 2001 to 2005 (Table 4.4) shows a consistent spatial clustering of school performance in New York City over time. The inference for Moran's I is based on a random permutation procedure, which recalculates the statistic many times to generate a reference distribution. The observed statistic is compared to this reference distribution to derive a p-value. Moran's I values in Table 4.4 for each year are significant with p-value = 0.001, derived using 999 permutations to generate the reference distribution.

Table 4.4 shows a decrease in Moran's I from 2002 to 2003 and 2004 to 2005 at the same time when there was a general increase in school performance in the city. This is an indication that in general there has been a decrease in spatial clustering associated with the increase in performance resulting in a slight reduction of spatial disparities. However, this is not the case between 2001 and 2002 when an increase in performance was associated with an increase in Moran's I indicating a slight increase in spatial disparities.

Table 4.4 Global Moran's I

Year	2001	2002	2003	2004	2005
Moran's I	0.54	0.57	0.50	0.52	0.47

The spatial clustering of school performance, as any other geographical phenomena, is a manifestation of the spatial clustering of the factors that impact it such as in this case the clustering of student and school characteristics. Table 4.5 shows the Moran's I of independent variables over time. As expected, there is a distinct spatial clustering of students' characteristics which remains throughout the period. Nevertheless, there are major changes in spatial clustering among teacher characteristics. The Moran's I for the percentage of the teachers fully licensed decreased dramatically from 0.64 in 2001 to 0.05 in 2005, due to efforts from the New York City's Department of Education to increase the number of licensed teachers, resulting in a more uniform spatial distribution. A similar pattern over time, although not as dramatic, is observed with respect to percentage of teachers with master degree and the percentage of teachers with more than five years of experience anywhere.

Table 4.5. Moran's I of Independent Variables Over Time

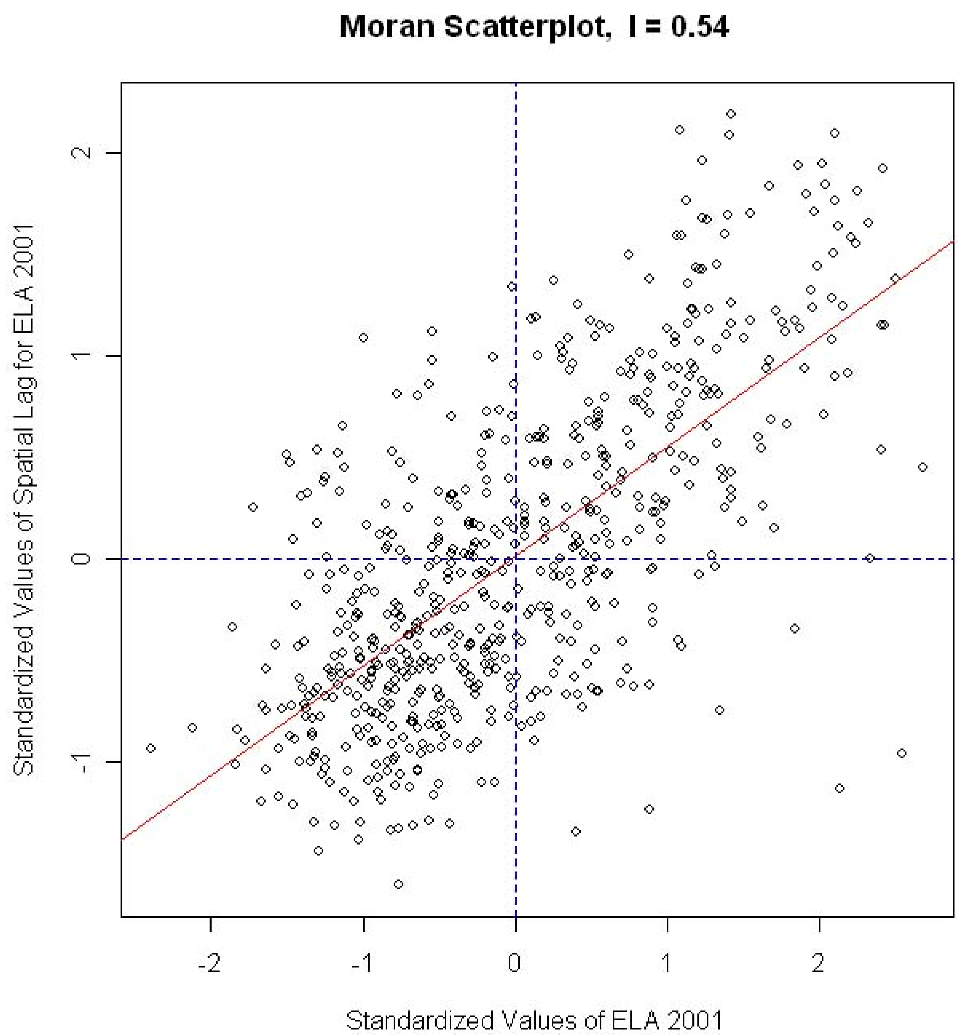
Independent variables	Year				
	2001	**2002**	**2003**	**2004**	**2005**
Students characteristics					
Eligible for Free Lunch (%)	0.66	0.64	0.67	0.67	0.63
Attendance (%)	0.57	0.55	0.53	0.54	0.52
Recent Immigrants (%)	0.58	0.59	0.58	0.57	0.53
English Language Learners (%)	0.44	0.50	0.43	0.33	0.29
School characteristics					
School size (# of students)	0.29	0.28	0.26	0.26	0.29
Teachers with master degree (%)	0.55	0.56	0.46	0.38	0.27
Teachers more than 2 years in the same school (%)	0.22	0.11	0.13	0.14	0.13
Teachers with greater than 5 years experience anywhere (%)	0.28	0.27	0.25	0.20	0.15
Teachers fully licensed and permanently assigned to the school (%)	0.64	0.55	0.44	0.14	0.05
Students/Teachers Ratio	0.29	0.23	0.21	0.14	0.23
Other					
Special Education (%)	0.15	0.14	0.11	0.10	0.09

4.2.1.c. Local Indicators of Spatial Association

Global Moran's I is an indicator of the extent of the overall spatial clustering of school performance, but it cannot identify where within the city the areas of clustering are located. Within the framework of ESDA, a series of new approaches has been developed called local statistics that address this issue (Fotheringham, 1997; Unwin and Unwin, 1996). A local version of Moran's I, one of several local indicators of spatial association (LISA), is particularly useful because it allows for the decomposition of the pattern of spatial association into four categories, corresponding with four quadrants in the Moran's I Scatterplot (Messner and Anselin, 2003).

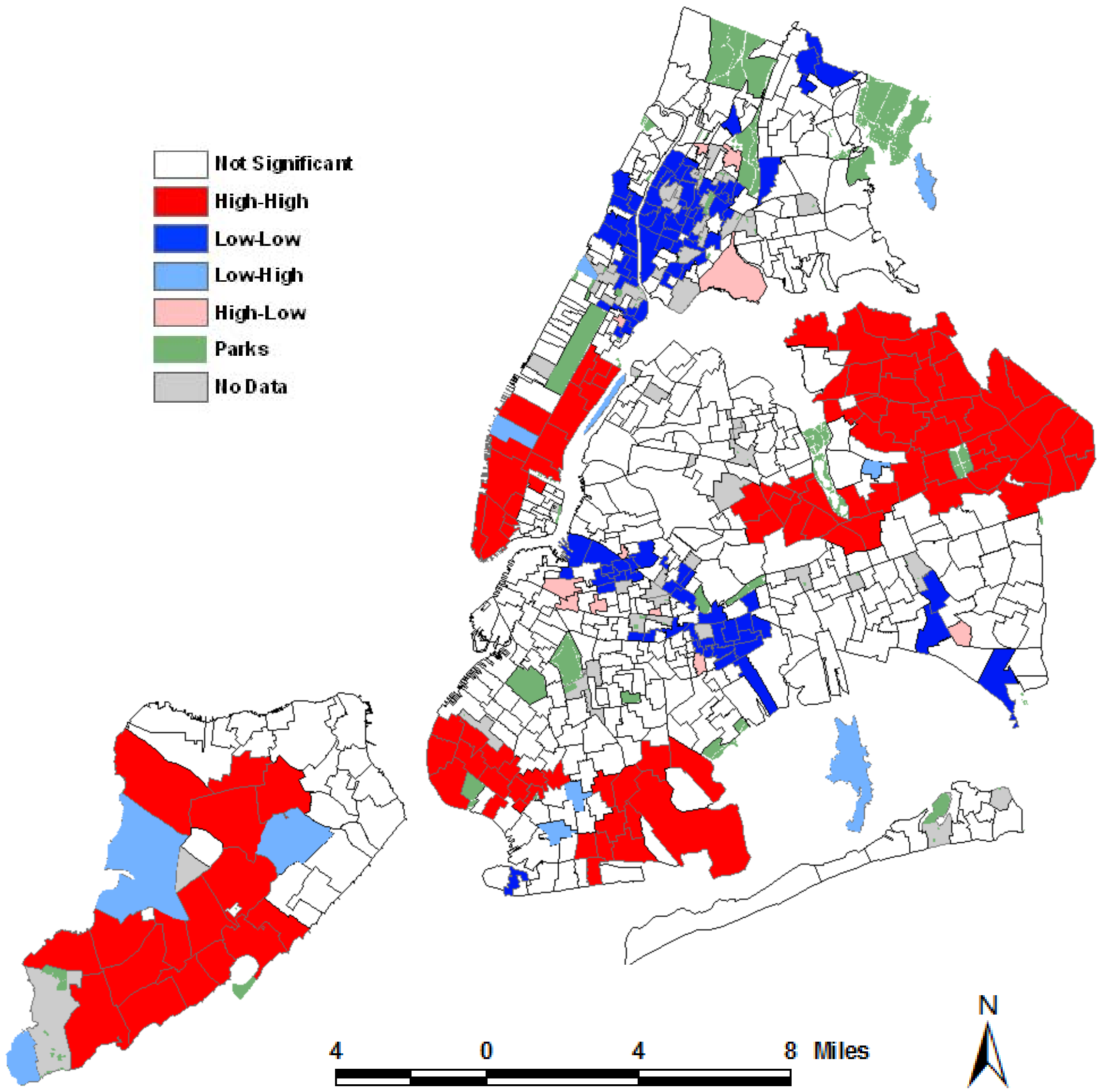
Figure 4.3 illustrates the Moran's I Scatterplot for ELA test in 2001 where the horizontal axis shows the standardized value of the ELA results and the vertical axis shows the standardized value of the average scores for each school's neighbors, as defined by the spatial weights matrix. The upper right quadrant shows those schools with above average performance which share boundaries with neighboring schools that also have above average performance (high-high). The lower left quadrant contains schools with below average performance which share boundaries with neighboring schools that also have below average performance (low-low). The upper right and lower left quadrants include areas of positive spatial autocorrelation and when the corresponding LISA statistics are significant they indicate spatial clustering. The lower right quadrant has schools with above average performance surrounded by schools with below average (high-low) and upper left quadrant has schools with below average performance surrounded by schools with above average (low-high). The lower right and upper left quadrants include areas of negative spatial autocorrelation and when the corresponding LISA statistics are significant they indicate spatial outliers. The Moran's I Scatterplot also displays the relationship between the local and global autocorrelation because the slope of the regression line through the points is the global Moran's I value.

Figure 4.3. Moran's I Scatterplot for Test Results in 2001



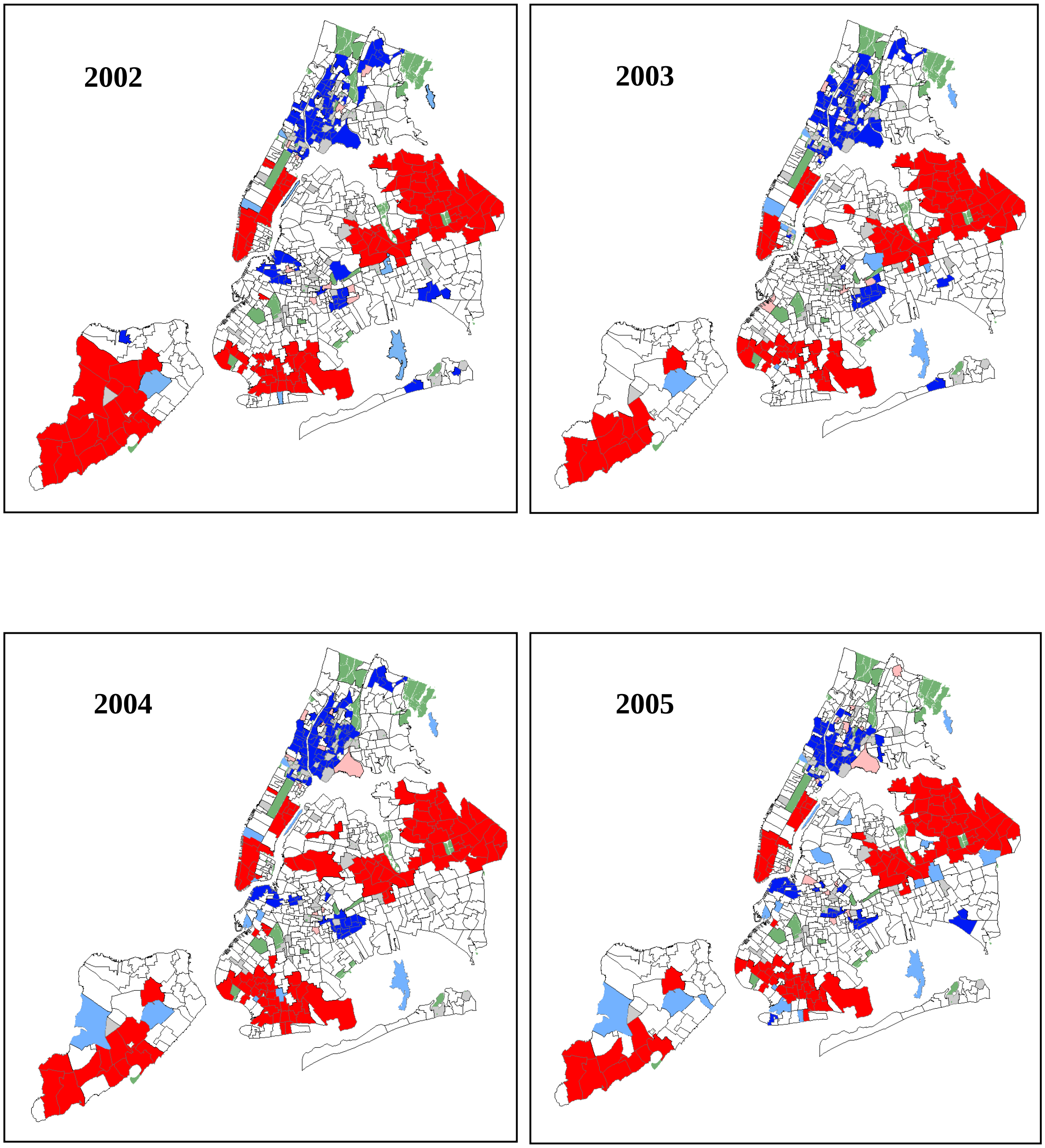
LISA Cluster Map for test results in 2001 (illustrated in Figure 4.4) shows the schools where the LISA statistic rejects the null hypothesis of spatial randomness. It also shows the spatial clusters and spatial outliers as in the four quadrants of the Moran's I Scatterplot described above. The red areas are comprised of high performing schools whereas the blue ones of low performing schools.

Figure 4.4. LISA Cluster Map for Test Results in 2001



There is consistent spatial clustering of school performance over time as shown in Figure 4.5.

Figure 4.5. LISA Cluster Map for Test Results from 2002 to 2005



The results for LISA are obtained using GeoDa 0.9.5 – i (freely downloadable from http://geodacenter.asu.edu/software/downloads), which is a software that implements ESDA and includes functionality ranging from simple mapping to exploratory data analysis, the visualization of global and local spatial autocorrelation and spatial regression (Anselin, Syabri and Kho, 2006).

4.2.2. Ordinary Least Square Regression Analysis

Ordinary least square (OLS) regression analysis is used to establish the extent to which the spatial patterning of school performance can be explained by the spatial patterning of the student and school characteristics. When residuals from OLS fit do not show spatial autocorrelation than the estimates obtained from the OLS regression would be sufficient and therefore there would be no need for spatial regression. Table 4.6 shows the results of weighted OLS regression, using Bisquare Weights as described in section 3.2.3, separately for each year over the period. Independent variables included in the estimation explain the majority of variation in school performance ranging from an adjusted R^2 of 79 percent in 2001 to 74 percent in 2005. Student characteristics are the main contributors explaining 77 percent of the variation in 2001 and 71 percent in 2005. Controlling only for the percentage of student eligible for free lunch, 68 percent is explained in 2001 and 49 percent in 2005. The finding that student characteristics explain majority of school performance, free lunch in particular, is consistent with previous findings in literature. Moreover, the estimation over time reveals that while student characteristics remain important, the contribution of free lunch decreases and that of attendance increases. Controlling for attendance, in addition to free lunch, adds to the adjusted R^2 another 6 percent in 2001 and 18 percent in 2005 suggesting a positive change most likely as a results of education reforms.

Table 4.6. Weighted Ordinary Least Square Regression Estimates

	Year				
Independent variables	**2001**	**2002**	**2003**	**2004**	**2005**
Students characteristics					
Eligible for Free Lunch (%)	-0.41*	-0.37*	-0.32*	.-0.37*	-0.30*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Attendance (%)	2.82*	3.15*	2.76*	3.14*	3.89*
	(0.28)	(0.27)	(0.29)	(0.29)	(0.27)
Recent Immigrants (%)	0.31*	0.34*	0.27*	0.33*	0.25*
	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)
English Language Learners (%)	-0.46*	-0.49*	-0.98*	-0.61*	-0.49*
	(0.06)	(0.06)	(0.09)	(0.09)	(0.07)
School characteristics					
School size (# of students in hundreds)	-0.47*	-0.61*	-0.45*	-0.56*	-0.23
	(0.12)	(0.11)	(0.13)	(0.13)	(0.14)
Teachers with master degree (%)	0.06	0.12*	0.04	-0.03	0.04
	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Teachers more than two years in the same school (%)	0.06*	-0.01	0.24*	0.18*	0.11*
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)
Other					
Special Education (%)	-0.36*	-0.43*	-0.51*	-0.50*	-0.45*
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)
Adjusted-R^2	0.79	0.81	0.76	0.77	0.74

Note: The estimation excludes variables that were not significant. Standard error in parentheses.

* denotes significance at the 5 percent level.

Findings indicate that all variables have the expected sign. The percentage of students eligible for free lunch, a proxy for poverty level in the school zone, has a significant negative effect on school performance. However, the magnitude of its effect decreases over time. School attendance, a proxy for the role of family dedication to education of their children, has a significant positive effect on school performance, with a slight increase in the magnitude of the effect in 2005. Interestingly, school performance increases with the proportion of recent immigrants in the school. As expected, an increase in the proportion of English language learners and of special education students has a negative effect on school performance. The magnitude of the negative effect of English language learners substantially increases in 2003 maybe due to issues related to the transition to a new system of administration, but decreases toward the end of the period to levels close to 2001. The same pattern is observed with special education, but the magnitude in 2005 is higher than in 2001.

Regarding school characteristics, school size as previous research has shown has a significant but small negative effect on school performance, except in 2005 when the effect is not significant. Among teacher variables, the percentage of teachers with more than two years in the same school has a positive significant effect on school performance in all but one year, indicating the important role of teachers' stability in students' performance. Surprisingly, the percentage of teachers with a master degree has no impact on school performance except in 2002 when it has a positive effect. The students/teachers ratio and two teacher variables, the percentage of teachers fully licensed and the percentage of teachers with more than five years anywhere, were not significant and therefore excluded from the regression.

4.2.3. Assessing Spatial Variation in the Relationship Using Jackknife

Regression results in the above section provide global estimates of the relationship between school performance and student/school characteristics. A global regression assumes that such relationship is stable across space and therefore does not provide information about the changes of the relationship in different parts of the study area. However, this assumption should be tested because a spatial variation of the relationship can reveal different social mechanisms of the phenomena under investigation. The existence of the spatial variation of the relationship, known as spatial heterogeneity, should be taken into account before modeling the spatial autocorrelation revealed in the OLS residuals. This is necessary because the residual spatial autocorrelation may be due to spatial heterogeneity (Baller et al, 2001).

The concept of spatial heterogeneity has been discussed more over the last few years in the context of local spatial analysis. A way that this issue has been approached lately has been through the use of Geographically Weighted Regression (GWR) (Fotheringham et al, 2000). However, Wheeler and Tiefelsdorf (2005) raised concerns about several issues related with GWR such as the multicollinearity surrounding the local GWR coefficients at a single location and the fact that GWR ignores that the local models must relate to a global reference model in order to express the local parameters as variation around their global counterparts. This chapter uses the Jackknife technique to assess the spatial variation of the relationship between school performance and student/school characteristics.

The name “Jackknife” is intended to suggest the broad usefulness of a technique "as a substitute for specialized tools that may not be available, just as the Boy Scout’s trusty tool serves so variedly” (Mosteller and Tukey, 1977, pg. 133). The basic idea of Jackknife is to assess the effect of a subset of data (or a particular record) on a dataset not by calculating the statistic on that subset alone but through the effect that results from omitting it. In this chapter, the Jackknife method involves omitting one school from the data and applying the OLS regression on the rest of the data. This procedure is repeated for each school resulting in as many regressions as there are schools. The jackknifed Y_j parameters are defined as:

$$Y_j = k * y_{all} - (k - 1) * y_{(j)} \qquad \text{for } j = 1, \ldots, k$$

where k is the number of schools, y_{all} are the coefficients in the OLS regression fit including all schools and $y_{(j)}$ are the coefficients from the fit that omits the j^{th} school. The jackknifed parameter values are an indication of the influence of the omitted school on the global parameter estimated for regression that includes all schools.

Jackknifed parameter values can also be used to produce unbiased estimates of the parameters for the global OLS model by calculating their respective means and also to derive an estimate of their variance (Mosteller and Tukey, 1977). Table 4.8 shows that the results of weighted OLS global jackknife coefficients, using Bisquare Weights as described in section 3.2.3, are similar to those produced by OLS shown in Table 4.7.

Table 4.7. Weighted Ordinary Least Square Jackknifed Regression Estimates

	Year				
Independent variables	**2001**	**2002**	**2003**	**2004**	**2005**
Students characteristics					
Eligible for Free Lunch (%)	-0.42*	-0.37*	-0.32*	.-0.37*	-0.28*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
Attendance (%)	2.81*	3.18*	2.78*	3.25*	3.69*
	(0.43)	(0.39)	(0.41)	(0.43)	(0.34)
Recent Immigrants (%)	0.33*	0.32*	0.28*	0.31*	0.19
	(0.13)	(0.13)	(0.13)	(0.13)	(0.11)
English Language Learners (%)	-0.46*	-0.49*	-0.95*	-0.62*	-0.49*
	(0.06)	(0.06)	(0.11)	(0.10)	(0.12)
School characteristics					
School size (# of students in hundreds)	-0.48*	-0.62*	-0.48*	-0.56*	-0.20
	(0.14)	(0.12)	(0.16)	(0.17)	(0.16)
Teachers with master degree (%)	0.05	0.12*	0.05	-0.03	0.01
	(0.06)	(0.05)	(0.06)	(0.07)	(0.07)
Teachers more than two years in the same school (%)	0.06	-0.02	0.25*	0.18*	0.11*
	(0.04)	(0.04)	(0.05)	(0.07)	(0.04)
Other					
Special Education (%)	-0.36*	-0.43*	-0.51*	-0.50*	-0.35*
	(0.06)	(0.06)	(0.07)	(0.08)	(0.06)

Note: The estimation excludes variables that were not significant. Standard error in parentheses.

* denotes significance at the 5 percent level.

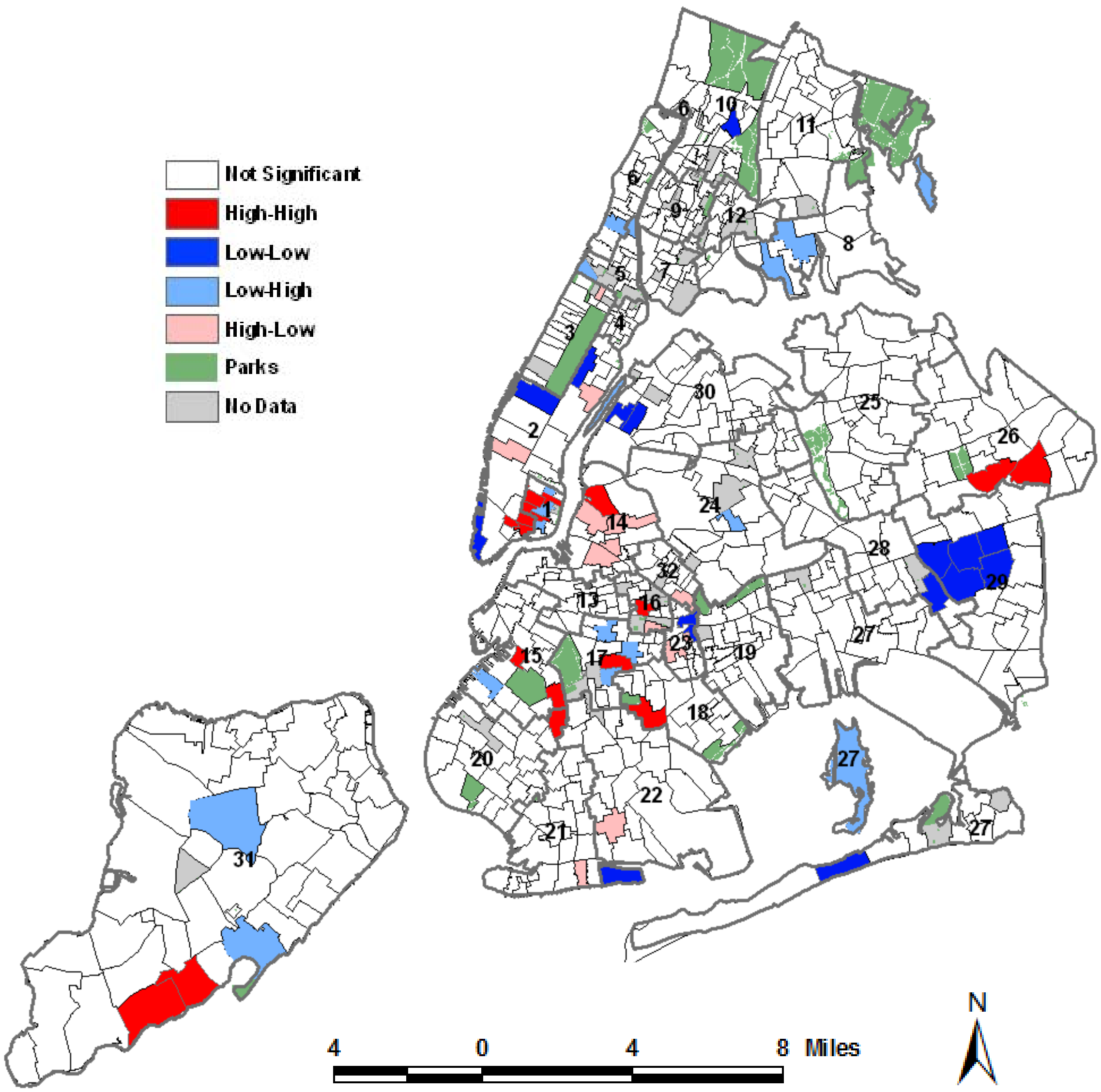
Jackknifed parameters are also powerful indicators of model performance across space (Stetzer, 1982). In this chapter, jackknifed parameters values are used to assess the stability of the relationship over space of each independent variable with school performance. Jackknifed parameter values for each school are treated as an ordinary spatial dataset and they are mapped to explore their spatial variability and to discover outliers. Their spatial clustering, that would suggest spatial heterogeneity, is accessed using Moran's I and LISA.

Figure 4.6 shows a LISA map of jackknifed parameter values for students eligible for free lunch for year 2001, overlaid with district boundaries and labeled with district numbers, which have a global Moran's I of 0.05 and p-value of 0.05, with small clusters of low values and high values. The lack of major clustered areas is an indication that the relationship between free lunch and school performance is the same over the city and that the global coefficient of -0.41 is a good representation of that relationship.

The jackknifed parameters are values around the mean global value and represent the impact that each school has on the overall global value of the coefficient. Schools with very low negative values of the jackknifed parameter are schools with a high percentage of students eligible for free lunch that have, as expected, a low school performance (like schools clustered within District 29 in Queens), or schools that have a low number of students eligible for free lunch and a good performance (the school in Upper East Side in Manhattan). Schools in high value clusters have a higher performance that would be expected from the global relationship based on their high number of students eligible for free lunch (the cluster of schools in downtown

Manhattan) or are schools with a low number of students eligible for free lunch that do not perform as well as would be expected from the global relationship (such as one school in Staten Island).

Figure 4.6. LISA for Jackknifed Coefficients of Eligible for Free Lunch, 2001



The spatial variation of jackknifed parameter values can also be assessed using SCT analysis, described on section 3.2.2, on just one factor, school districts. Jackknifed values have several outliers (because their calculation is sensitive to the large number of schools) which makes SCT a better choice, being a robust method, compare to ANOVA that is affected by outliers. Nevertheless, regarding free lunch jackknifed coefficients neither SCT nor ANOVA show any district to be statistically different from other districts.

Figure 4.7. LISA for Jackknifed Coefficients of Recent Immigrants, 2001

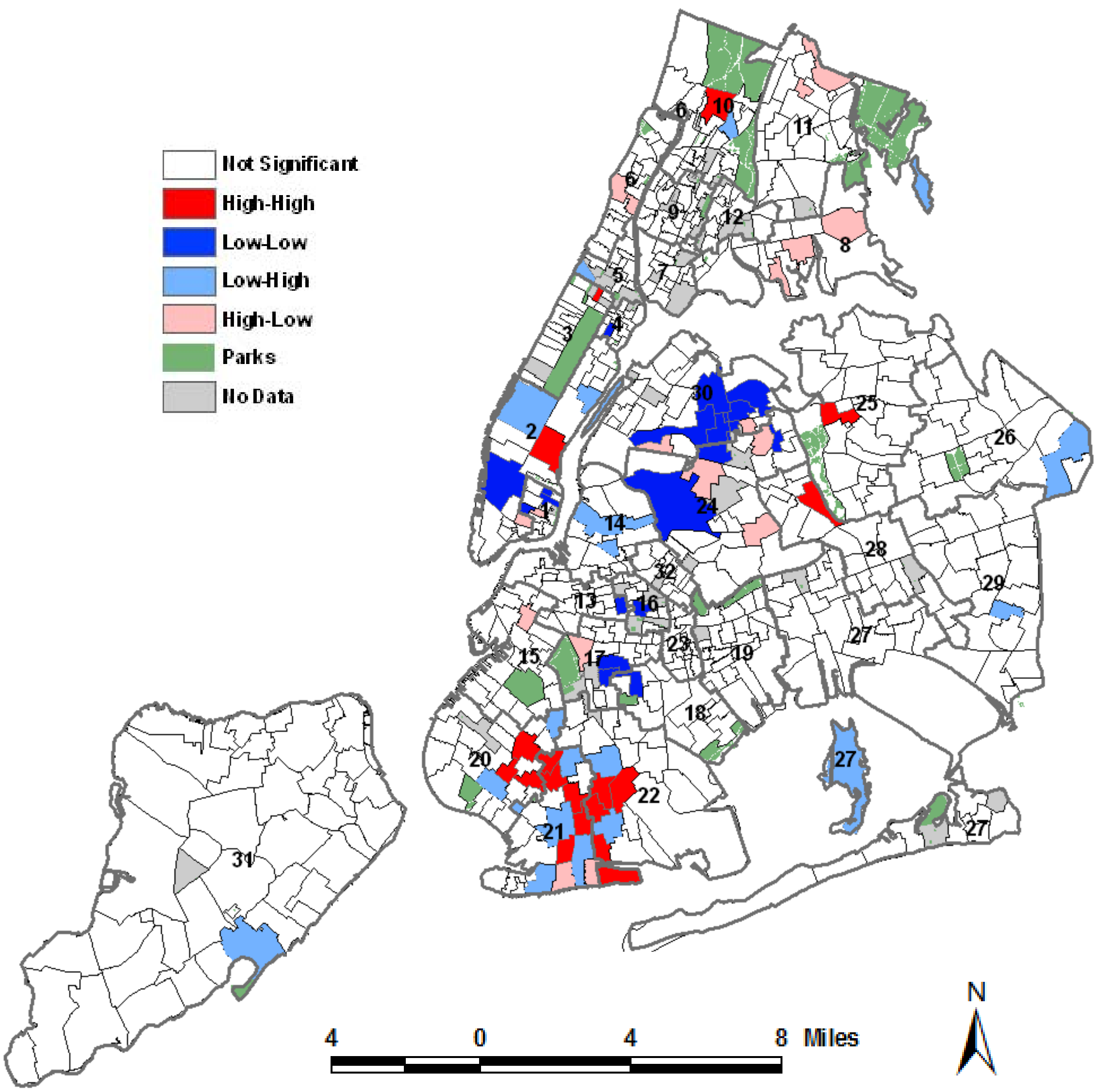
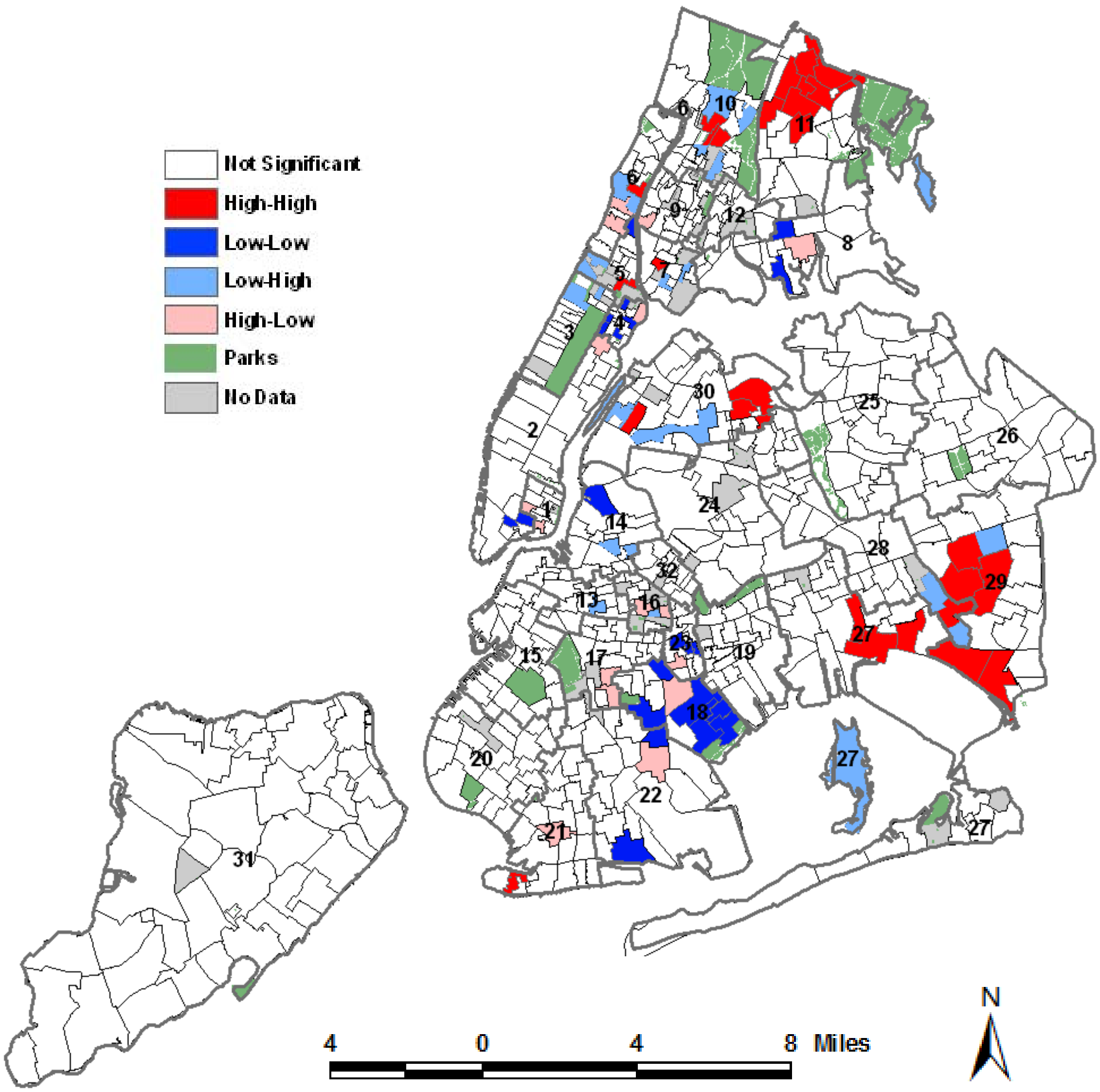


Figure 4.7 shows the LISA map for the jackknifed parameter values of the percentage of students that are recent immigrants which have a positive global relationship with school performance. The schools in the low value clusters within District 24 and 30 have low school performance given what is expected from their number of recent immigrants whereas schools in Districts 20, 21 and 22 in Brooklyn are performing as expected. The schools in the low values cluster have a higher percentage of Hispanic students compare to the schools in the high values cluster where White students are the majority. This suggests that even though the global relationship between recent immigrants and school performance is positive, this is not true for the schools in the low value cluster which are below global average and clustered.

Figure 4.8 shows the LISA map for the jackknifed parameter values of the percentage of students as English language learners (ELL) that have a negative global relationship with school performance. District 11 in Bronx and District 29 in Queens have clusters of schools that have a low school performance even though they have a small percentage of students as English language learners.

Results obtained on the jackknifed parameter values for other variables also indicate that there is no substantial spatial variation suggesting that it is appropriate to fit a global OLS regression (or if there is spatial autocorrelation on the OLS residuals, a global spatial regression). Table 4.8 shows Moran's I values for jackknifed values for each coefficient over time indicating the lack of spatial clustering and therefore of spatial heterogeneity.

Figure 4.8 LISA for Jackknifed Coefficients of ELL, 2001



A SCT analysis was also performed on the jackknifed parameters, using districts and years as factors, to test their variation over time or any interaction with districts. The findings show that, with the exception of English language learners' relationship in 2003, there is no significant difference and all relationships are stable over time and space.

Table 4.8. Moran's I for Jackknifed Coefficients by Year

	Year				
Independent variables	**2001**	**2002**	**2003**	**2004**	**2005**
Students characteristics					
Eligible for Free Lunch (%)	0.05*	0.09*	0.04	0.10*	0.00
Attendance (%)	0.06*	0.14*	0.06*	0.13*	0.07*
Recent Immigrants (%)	0.00	0.02	0.03	0.06*	-0.02
English Language Learners (%)	0.05*	0.10*	0.08*	-0.01	0.03
School characteristics					
School size (# of students in hundreds)	0.03	0.04*	0.01	-0.01	0.04*
Teachers with master degree (%)	0.00	0.02	-0.01	0.04*	-0.01
Teachers more than two years in the same school (%)	0.03	0.06*	-0.01	0.00	-0.06*
Other					
Special Education (%)	-0.02	0.03	0.05*	0.04*	0.04

* denotes significance at the 5 percent level.

The code to calculate the jackknifed parameters using weighted OLS regression is written using the statistical software R (R Development Core Team, 2008).

4.2.4. Spatial Regression

The application of spatial regression in this subsection follows the conceptual and theoretical framework for spatial analysis discussed in Baller et al (2001). Previous research has shown that the performance of children in schools reflects structural factors such as: family background, the quality of the schools and neighborhood effects. If school performance is determined solely by the structural factors included in the regression model, there should be no spatial patterning of school performance beyond that created by covariates' similarities of geographically proximate schools. This situation, referred as "Structural Similarity" by Baller et al (Figure 1, 2001), is assessed by testing for spatial autocorrelation on the OLS regression residuals. For each year there is a significant spatial autocorrelation of the OLS residuals for which the Moran's I values are shown in Table 4.9. The amount of spatial autocorrelation in school performance explained from structural covariates (student and school characteristics), calculated as the difference between the two values of Moran's I for each year, is consistent over time ranging from 0.36 in 2001 to 0.40 in 2005.

Table 4.9. Moran's I of School Performance and OLS Residuals

Year	2001	2002	2003	2004	2005
Moran's I of school performance	0.54	0.57	0.50	0.52	0.47
Moran's I of OLS residuals	0.18	0.18	0.13	0.13	0.07

The presence of spatial autocorrelation in the OLS residuals, an indication of spatial effects, necessitates the use of spatial models. Two spatial models used to account for spatial effects, discussed below, are the spatial error model and the spatial lag model.

4.2.4.a. Spatial Error Model

The spatial error model uses the clustering of the errors to account for the clustering of school performance not explained by measured independent variables. This situation is referred as "Spatial Error Effects" in Baller et al (Figure 1, 2001). In this case, the observed spatial clustering of school performance is accounted for simply by the geographic patterning of measured and unmeasured independent variables. This model would be appropriate in situations when spatial autocorrelation is caused by the "mechanism of grouping forces" when people with common characteristics are found clustered together in space either by choice or as a result of social, economic or political forces (Voss et al, 2006). A spatial error model is specified (in matrix notation) as follows:

$$y = X\beta + u$$
$$u = \rho W u + \varepsilon$$

where y is a (n x 1) vector representing the dependent variable, X is a (n x k) matrix representing the k – 1 independent variables, β is a (k x 1) vector of regression parameters to be estimated, u is a (n x 1) vector of error terms assumed to have a covariance structure as given in the second equation, ρ is a spatial lag parameter to be estimated, W is a (n x n) "weights" matrix defining the "neighborhood" structure in the spatial process such that Wu is a (n x 1) vector of spatial lags of the model disturbance term u and ε is a (n x 1) vector of independently distributed errors.

4.2.4.b. Spatial Lag Model

The spatial lag model incorporates not only the spatial influence of unmeasured independent variables but also the effect of the performance of neighboring schools which is identified in the lagged dependent variable. Spatial lag models are used, for example, in the estimation of spatial dependence of homicide rates (Morenoff et al, 2001).

This situation is referred as “Spatial Lag Effects” in Baller et al Figure 1, (2001). This model would suggest a ”feedback mechanism” when spatial proximity generally increases the interactions and the strength of the feedback as in a diffusion process (Voss et al, 2006). The increase in performance in some schools might have a positive effect on the neighboring schools if the interaction effect between the schools or between neighboring children exists. The improvement in performance may be as a result of the diffusion of educational technologies or as an effect of public policies such as school choice. A spatial lag model is specified (in matrix notation) as follows:

$$y = \lambda Wy + X\beta + \varepsilon$$

where X, β and ε are the same or similar to those presented for the spatial error model. Wy is a (n x 1) vector of spatial lags of the dependent variable y, and λ is a spatial lag parameter to be estimated.

4.2.4.c. Model Selection

The spatial model, an error or lag model, that better represents the data is chosen according to the selection rule described in Anselin (2005). The procedure of the selection of the model is illustrated with the following example of regression analysis with ELA for the year 2001. The first step is to check for “Structural Similarity” by implementing an OLS regression and evaluating the spatial autocorrelation of the residuals. The residuals Moran’s I is 0.18 and its z-value, which is based on a normal approximation and takes into account the fact that these are residuals, is 7.8 and significant. This dependence in regression errors violates the independence assumption for errors in OLS fit, affecting the estimation of regression parameters and their standard errors. A spatial model, either error or lag, would take into account the spatial autocorrelation of the residuals in the estimation of the regression parameters and their significance. While the residuals Moran's I statistic detects misspecifications in the model, it is less helpful in suggesting which model, spatial error or spatial lag, is a better fit. Thus, Lagrange Multiplier (LM) test statistic is used to decide the proper spatial model.

Diagnostics for spatial dependence of OLS results for ELA in year 2001 show that both LM statistics (LM-Error and LM-Lag) are significant. In this case, as described in Anselin (2005), the next step is to compare their robust forms. Only the robust LM-Error is highly significant which means that the spatial error model is the appropriate one to be estimated. The spatial error model is estimated by means of maximum likelihood and adjusted R^2 is not an appropriate measure to compare the results between the spatial error fit and the OLS fit. In this case the comparison between the two fits is based on three measures: log likelihood, the Akaike Information Criterion and the Schwarz Criterion. The best model should have higher value (less negative is better) for the first measure and smaller values for the other two. The comparison of the three measures for year 2001 suggests an improvement of fit for the spatial error specification. The estimated spatial autoregressive coefficient is 0.33 and is highly significant.

The model selection procedure illustrated above is followed for each of the other years. Estimation results indicated that in all years spatial models were a better fit than OLS regression. Table 4.10 shows the estimated coefficients and their respective

Table 4.10. Spatial Regression Estimates

	Year				
Independent variables	**2001**	**2002**	**2003**	**2004**	**2005**
Students characteristics					
Eligible for Free Lunch (%)	-0.39*	-0.34*	-0.32*	.-0.31*	-0.25*
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Attendance (%)	3.19*	3.48*	2.85*	3.05*	3.60*
	(0.29)	(0.27)	(0.30)	(0.34)	(0.32)
Recent Immigrants (%)	0.20*	0.22*	0.13	0.14	0.11
	(0.09)	(0.09)	(0.10)	(0.11)	(0.11)
English Language Learners (%)	-0.53*	-0.56*	-0.91*	-0.56*	-0.44*
	(0.07)	(0.06)	(0.09)	(0.10)	(0.08)
School characteristics					
School size (# of students in hundreds)	-0.33*	-0.47*	-0.32*	-0.43*	-0.18
	(0.12)	(0.11)	(0.13)	(0.15)	(0.16)
Teachers with master degree (%)	0.05	0.11*	0.09	-0.05	0.03
	(0.04)	(0.04)	(0.05)	(0.06)	(0.06)
Teachers more than two years in the same school (%)	0.05	-0.04	0.20*	0.16*	0.09*
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Other					
Special Education (%)	-0.37*	-0.42*	-0.50*	-0.50*	-0.40*
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
Spatial regression model	Error	Error	Error	Lag	Lag
Lambda for Error model and Rho for Lag model	0.33*	0.36*	0.27*	0.17*	0.12*

Note: The estimation excludes variables that were not significant. Standard error in parentheses.

* denotes significance at the 5 percent level.

standard errors obtained from spatial regression model. There are no major changes in the estimated parameters obtained from spatial regression model compare to OLS results. In general, the magnitude of the coefficients from spatial regression is slightly smaller than those from OLS regression and, as expected, the standard errors of the coefficients are slightly bigger. However, the estimates and their significance in the former model are more precise because spatial regression controls for the spatial autocorrelation in the residuals. There are a few differences between OLS and spatial regression results. In 2001, teachers with more than two years experience in the same school turn out to be insignificant when using spatial regression. The relationship between school performance and recent immigrants becomes insignificant from 2003 to 2005 and school size becomes insignificant in 2005.

For years 2001 to 2003 error model is a better fit of the data, whereas for years 2004 and 2005 lag model is a better fit. It is important to note the change between 2003 and 2004, the first year when reforms were fully implemented in New York City, from an error model to a lag model. The error model evaluates the extent to which the clustering of school performance not explained by school and student characteristics can be accounted for by the clustering of error terms. It indicates that there are unmeasured factors at the neighborhood or district level that have an effect in a number of schools that are close in space. These unmeasured effects appear as spatially autocorrelated errors. The lag model, similar to the error model incorporates the effect of unmeasured factors, but in addition it also takes into account another effect, that of neighboring schools. This additional effect in the lag model is an indication of a global change which in this study can be identified with the effect of reforms throughout the city’s public schools. Based on Baller et al (Figure 1, 2001), it can be inferred that in the presence of a lag model school performance is not only affected by its own student and school characteristics but also by those of the neighboring schools, after accounting for the spatial autocorrelation of the errors. The lag model provides an indication of the spatial interaction between the neighboring schools but does not provide the mechanism through which the interaction takes place. Further research is needed to discover the mechanisms behind the process suggested by the lag model.

Figure 4.9 shows the LISA map of residuals from OLS regression in 2001 which have a Moran's I of 0.18. The pattern of low value (over predicted) clusters in District 11 and District 29 is similar to the pattern of jackknifed parameter values for English language learners shown in Figure 4.8. The pattern of negative residuals in District 24 and District 30 and the pattern of positive residuals on District 20, 21 and part of District 22 are similar to the patterns displayed by jackknifed values of the variable of recent immigrant students on Figure 4.7. The spatial autocorrelation of the residuals on the districts mentioned above might be attributed to the local spatial variation of those two variables.

The major areas with spatially autocorrelated residuals are mainly within boundaries of several districts such as, those of negative residual clusters in districts 11, 24, 29 and 30 and those with positive residual clusters in districts 2, 18, 20, 21, 22 and 23. The pattern of the residuals within districts raises the question about the role of these districts in affecting school performance beyond the student and school characteristics.

Figure 4.9. LISA of Residuals from OLS Regression, 2001

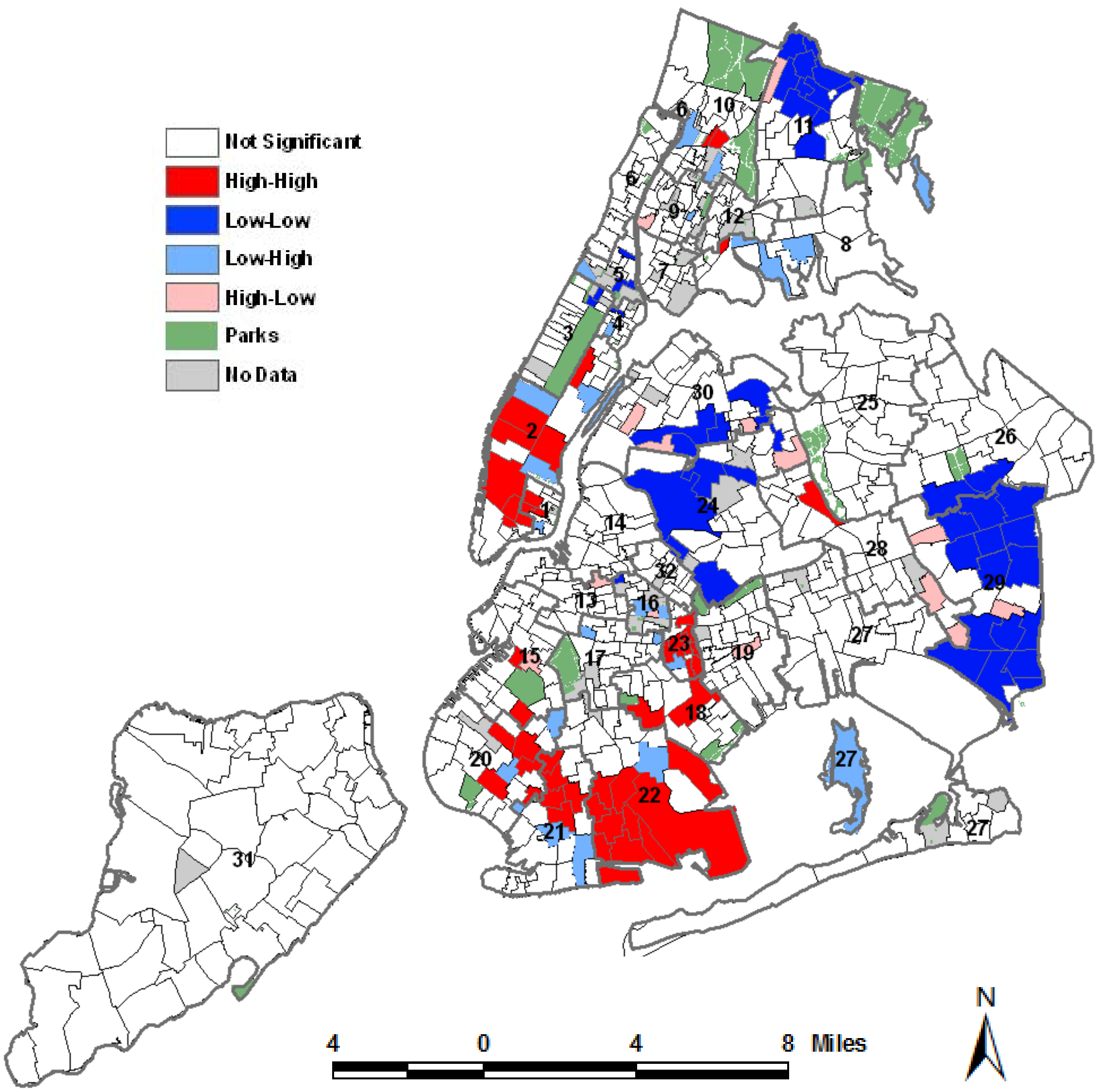


Table 4.10 shows, together with the student and school coefficients, the estimated district effects from a weighted OLS regression. The school performance of District 15 was close to the average value for all districts in 2001 and therefore was used as the reference district to calculate the district effects. District coefficients have the expected sign for the districts that contain spatially autocorrelated residuals shown in Figure 4.11

Table 4.11. Weighted OLS Regression Estiamtes with Districts by Year

	Year				
Independent variables	**2001**	**2002**	**2003**	**2004**	**2005**
Students characteristics					
Eligible for Free Lunch (%)	-0.38*	-0.36*	-0.33*	-0.38*	-0.30*
Attendance (%)	3.10*	3.41*	2.96*	3.10*	3.64*
Recent Immigrants (%)	0.19*	0.27*	0.03	0.12	0.04
English Language Learners (%)	-0.51*	-0.56*	-0.73*	-0.34*	-0.29*
School characteristics					
School size (# of students in hundreds)	-0.29*	-0.32*	-0.19	-0.13	-0.03
Teachers with master degree (%)	0.08	0.11*	0.15*	0.02	0.13*
Teachers more than two years in the same school (%)	0.06*	-0.03	0.16*	0.09*	0.07*
Other					
Special Education (%)	-0.34*	-0.33*	-0.48*	-0.45*	-0.48*
Districts					
1	3	2	4	5	-2
2	8*	8*	8*	6*	2
3	-3	1	-3	-6*	-8*
4	-3	-1	0	0	-1
5	-7*	-11*	0	-9*	-15*
6	-1	1	-6*	-11*	-13*
7	4	-8*	-2	-9*	-14*
8	4	-4	-3	-4	-5*
9	-1	2	2	-5*	-1
10	5*	-3	1	-5*	0
11	-7*	-10*	-6*	-8*	-6*
12	1	3	8*	-3	2
13	0	-2	7*	2	2
14	-1	-4*	10*	1	-3
15 (omitted)					
16	-9*	2	11*	2	0
17	2	-1	14*	3	-1
18	5	-4	1	0	-3

Table 4.11. (cont.)

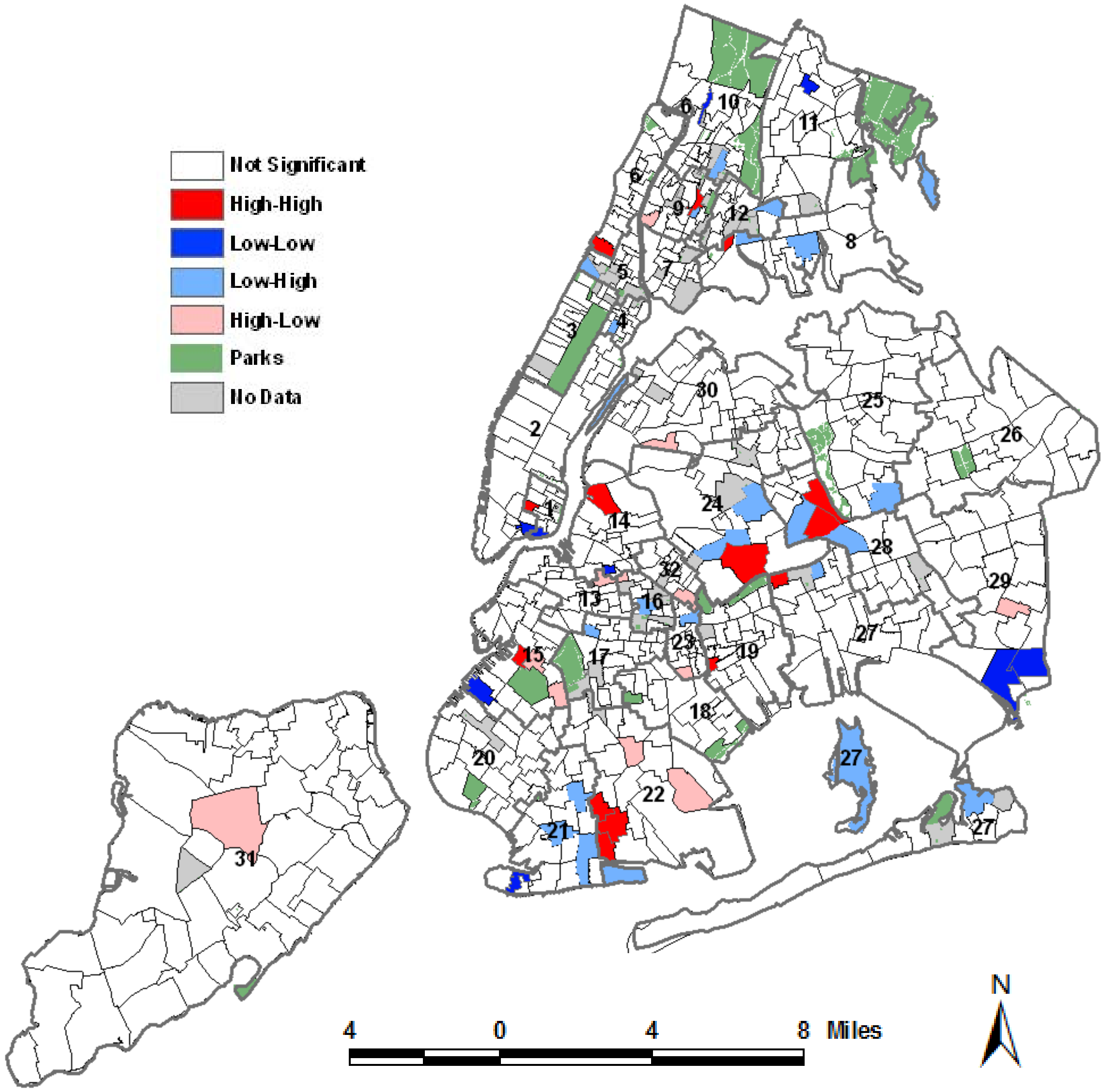
Independent variables	Year				
	2001	**2002**	**2003**	**2004**	**2005**
19	-6*	-5*	-2	-3	-2
20	7*	0	10*	6*	3
21	2	7*	6*	10*	6*
22	8*	3	8*	5*	0
23	13*	6*	8*	2	6*
24	-8*	-7*	2	-3	-4
25	1	-1	0	-2	2
26	-1	-1	2	1	-1
27	-2	-9*	-4	0	2
28	2	2	5*	3	2
29	-11*	-13*	-6*	-11*	-9*
30	-3	-4	1	2	0
31	-2	-4*	-3	-4*	-3
32	-3	2	4	1	-4

* denotes significance at the 5 percent level.

and mentioned above. District 29 has the largest negative effect and together with District 11 are the only two districts that have significant negative effects for each year. District 5 is another district that has a significant negative effect on school performance in for four out of five years. Districts 20, 21, 22 and 23 exhibit significant positive effects in three or four years out of five. District 2, with a reputation of good management, has also a consistent positive impact on school performance.

Once districts are included in the weighted OLS regression, there is no longer spatial autocorrelation in the residuals. Figure 4.10 shows a LISA map of these residuals that exhibit no clustering and have an insignificant Moran's I value of 0.03.

Figure 4.10. LISA of Residuals from OLS with Districts, 2001



4.3. Summary

Findings from the analysis in this chapter show that there is spatial clustering in educational achievement in New York City which persists over time. Measured by global Moran's I, it decreased slightly over time from 0.54 in 2001 to 0.47 in 2005. This spatial variation of school performance reflects the geography of New York City represented by the spatial clustering of student and school characteristics. The part of school performance clustering explained by the student and school characteristics is consistent over time and ranges from 0.36 in 2001 to 0.40 in 2005.

The finding that student characteristics explain majority of school performance is consistent with results in previous research. Among student characteristics, the percentage of students eligible for free lunch, a proxy for poverty level in the school zone, is the most important characteristic. However, the magnitude of its effect decreased over the period of this study. School attendance, a proxy for the role of family dedication to education of their children, has a significant positive effect on school performance, with a slight increase in the magnitude of the effect in 2005. Interestingly, school performance increases with the proportion of recent immigrants in the school, even though this relationship shows local variation maybe due to the composition of the population of immigrants. As expected, an increase in proportion of English language learners and of special education students decreases school performance. The magnitude of the effect of English language learners substantially increases in 2003 maybe due to issues related to the transition to a new system of administration, but decreases toward the end of the period to levels close to 2001. The same pattern is observed with special education, but the magnitude in 2005 is higher than in 2001.

Regarding school characteristics, school size as previous research has shown has a significant but small negative effect on school performance, except in 2005 when the effect is not significant. Among teacher variables, the percentage of teachers with more than two years in the same school has a positive significant effect on school performance in all but one year, indicating the important role of teachers' stability in students' performance. Surprisingly, the percentage of teachers with a master degree has no impact on school performance except in 2002 when it has a positive effect.

Results from Jackknife analysis indicate that there is no spatial variation in the relationship between school performance and student and school characteristics. The relationships are also stable over time, except for English language learners in 2003 that showed a significant increase in the negative relationship. There is spatial dependence in school performance, unaccounted for by student and school characteristics, revealed in the spatial autocorrelation of residuals from the OLS regression. This spatial dependence suggests the use of a spatial regression model; for years 2001 to 2003 error model is a better fit to the data, whereas for years 2004 and 2005 lag model is a better fit. Regression analysis shows that some school districts have an impact on school performance, beyond the effect of student and school characteristics.

.

Chapter Five

Conclusions

The central issues of public education reforms over the last decade have been the increase in school performance, the reduction of the achievement gap between different subgroups and the improvement of school factors affecting performance. These reforms have focused attention on schools as the main agent of change which provides an opportunity for geographers to investigate the impact of geographical context on school performance. In this study, the geographic context ranges from the school catchment area, to its surrounding areas and the school district to which the school belongs.

5.1. Main Research Findings

From a geographical perspective, this study raised and answered several questions related to the issues above. First, is there inequality in educational achievement in New York City across districts and does it change over time? Findings from this study show that there is inequality in educational achievement in New York City across the lowest performing districts in Upper Manhattan and south Bronx compared to the highest performing districts in north-east Queens and Lower Manhattan, after taking into account the race/ethnicity effect. There was an increase in public elementary school performance throughout the period (except for the academic year 2003-2004); furthermore this improvement has been uniform across the City showing that there was no change on the impact of districts over time.

Second, is there variation in the achievement gap between race/ethnicity subgroups across districts? The achievement gap between the higher performing racial/ethnic groups (Asian and White) and the lower performing ones (Black and Hispanic) remains substantial. School districts had no effect in reducing the race/ethnicity achievement gap indicating that this gap does not change over space. In other words, there are no districts that have done a better job in reducing the gap. Even though its reduction has been the main goal of the education reforms, there is no decrease in the race/ethnicity achievement gap after the reforms. Future research should consider more years to better assess the impact of the reforms on gap reduction, since their effect may require some more time to materialize.

Third, what are the student and school characteristics that explain the spatial patterning of school performance and does this relationship change over time? There is a spatial clustering of school performance and the majority of it is explained by the spatial clustering of student and school characteristics. Consistent with previous findings in literature, student characteristics explain a great part of school performance. Among them, the percentage of students eligible for free lunch in particular - a proxy for poverty level in the school zone – is the most important factor and as expected has a negative effect. Another important student characteristic is the school attendance - a proxy for the role of family dedication to education of their children - which has a positive effect. The other student characteristics, percentage of student that are recent immigrants and percentage of student that are English language learners have a smaller effect. School performance increases with the number of recent immigrants even though this relationship shows some local variation maybe due to the composition of the population of immigrants. Regarding school characteristics, school size as previous research has shown has a significant but small negative effect on school performance, except in 2005 when the effect is not significant. Among teacher variables, the percentage of teachers with more than two years in the same school has a positive significant effect on school performance in all but one year, indicating the important role of teachers' stability in students' performance. Surprisingly, the percentage of teachers with a master degree has no impact on school performance except in 2002 when it has a positive effect.

Fourth, does the relationship between school performance and student and school characteristics vary in space? There is no spatial variation in the relationship between school performance and student and school characteristics. This lack of spatial heterogeneity indicates that the application of global regression analysis would be appropriate. The relationships are also stable over time, except for English language learners in 2003 that showed a significant increase in the negative relationship. The lack of major clustering in the relationship does not, however, exclude local spatial variations, common in geographical phenomena.

Fifth, is there a spatial dependence in school performance and what is the spatial process generating it? Findings indicate that there is spatial dependence in school performance, unaccounted for by student and school characteristics, revealed in the spatial autocorrelation of residuals from the OLS regression. This spatial dependence is accounted for by using spatial error and spatial lag models. For years 2001 to 2003 error model is a better fit of the data, whereas for years 2004 and 2005 lag model is a better one. It is important to note the change between 2003 and 2004, the first year when reforms were fully implemented in New York City, from an error model to a lag model. The error model evaluates the extent to which the clustering of school performance not explained by school and student characteristics can be accounted for by the clustering of error terms. It indicates that there are unmeasured factors at the neighborhood or district level that have an effect in a number of schools that are close in space. These unmeasured effects appear as spatially autocorrelated errors. The lag model, similar to the error model incorporates the effect of unmeasured factors, but in addition it also takes into account another effect, that of neighboring schools. This additional effect in the lag model is an indication of a global change which in this study can be identified with the effect of reforms throughout the city's public schools. The lag model provides an indication of the spatial interaction between neighboring schools but does not provide the mechanism through which the interaction takes place. Further research is needed to discover the mechanisms behind the process suggested by the lag model.

5.2. Policy Implications

The findings of this study emphasize the importance of geographic context in educational achievement. Such context is evident in the magnitude and consistency in spatial clustering of important students' characteristics such as the percentage of students eligible for free lunch and their attendance. Results indicate that in general there have been positive changes in the public education system in New York City after the implementation of reforms in the 2003-2004 academic year. These changes are reflected primarily in the decrease in spatial clustering associated with an increase in of school performance, in the continuation of the reduction of the negative effect of poverty, school size and the increase in the positive effect of attendance after the implementation of the reforms. Another positive indication is the increase in the spatial equality of distribution of licensed teachers and teachers with master degree. Among teacher variables, teachers' stability has a small but significant effect on school performance. The impact of the reforms has not been as positive regarding the performance of English language learners and special education students whose negative relationship with school performance has not decreased.

Findings suggest that the spatial inequality of school performance is not only related with the student and school characteristics but in some districts also with the district management and the community within that district. Thus, the education reforms should focus not only on schools but also include school district in their educational policies and accountability systems.

Another important finding of this research is that despite an improvement in school performance in the city, the race/ethnicity gap has not decreased over time even though its reduction was the main goal of the national reform. The spatial clustering of students' socioeconomic characteristics and of low performing race/ethnicity subgroups suggests that policy makers should focus not only in the schools but also the neighborhoods where these schools are located. Furthermore, the spatial concentration of social disadvantage should be taken into consideration in policies regarding allocation of resources that should be spatially focused.

5.3. Methodological Contributions

This study investigates for the first time, to the best of my knowledge, the educational achievement in New York City from a geographical perspective. More specifically, it examines spatial inequality in educational achievement, spatial variation of race/ethnicity achievement gap and explains the spatial pattern and spatial dependency of school performance. In addition to its findings, this research contributes to the geographical literature by introducing a robust framework to explore the impact of the geographic context on the outcome of interest. This framework, with the Square Combining Table method at its core, includes also Bisquare Weights and Multiple Comparison procedures. This framework can also be used in a geographical context of a larger scale such as regional, state or national to investigate the spatial variation of other geographical phenomena. Another important methodological contribution is the introduction to educational literature of the Jackknife technique to examine the spatial heterogeneity of the relationship between school performance and student and school characteristics. The Jackknife technique and the methods applied on jackknifed coefficient values to examine their variation in space and over time can be utilized to study other geographical phenomena in the same way as here. The framework discussed above and the Jackknife technique deserve a place among the tools used by geographers in their geographical analysis.

References

Anselin, L. (2002). Under the Hood: Issues in the Specification and Interpretation of Spatial Regression models. *Agricultural Economics*, 27, 247-267.

Anselin, L. (2005). Exploring Spatial Data with GeoDa™: a Workbook. Available from: https://www.geoda.uiuc.edu/documentation/tutorials/geodaworkbook.pdf

Anselin, L., Syabri, L. and Kho, Y. (2006). GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis*, 38, 5-22.

Aspel, D. (2005, May 04). Teaching Fellows Make their Mark. Gotham Gazette. Available from: http://www.gothamgazette.com/article/education/20050504/6/1405

Baller, R. D., Anselin, L., Messner, S. F., Deane, G. and Hawkins, D. F. (2001). Structural Covariates of U.S. County Homicide Rates: Incorporating Spatial Effects. *Criminalogy*, 39(3): 561-590

Barr, J. (2005). Teacher Location Choice and the Distribution of Quality: Evidence from New York City. *Contemporary Economic Policy*, 23 (4): 585-600

Barrow, L. and Rouse, C. E. (2005). Causality, Causality, Causality: the View of Education Inputs and Outputs from Economics. Working paper No. 05-15, Federal Reserve Bank of Chicago.

Baum, J. (2005, July 21). Setbacks and Successes for Education Councils. Gotham Gazette. Available from: http://www.gothamgazette.com/article/education/20050721/6/1487

Biederman, M. (2004, May 4). Teaching Certificates Are Tougher to Get. Gotham Gazette. Available from: http://www.gothamgazette.com/article/education/20040504/967

Bradford, M. (1991). School-Performance Indicators, the Local Residential Environment, and the Parental Choice. *Environmental and Planning A*, 23: 319-332

Brimicombe, A. (2000). Constructing and Evaluating Contextual Indices Using GIS; a Case of Primary School Performance Tables. *Environmental and Planning A*, 32: 1909-1933.

Brunsdon, C.F., Fotheringam, A.S. and Charlton, M.E. (1998). Geographically Weighted Regression-Modeling Spatial Non-Stationarity. *The Statistician*, 47(3): 431-443

Bryant, M. J., Hammond, K. A., Bocian, K. M., Retiing, M. F., Miller, C. A. and Cardullo R. A. (2008). School Performance Will Fail to Meet Legislated Benchmarks. *Science*, 321

Butler, T. and Hamnett C. (2007). The Geography of Education: Introduction. *Urban Studies*, 44 (7); 1161-1174

Children First History, 2002. The New York City Department of Education. Available from: http://schools.nyc.gov/AboutUs/ChildrenFirstHistory/default.htm

Coleman, J. S. and E. Q. Campbell with C.F. Hobson, J. McPartland, A.M.Mood, F.D. Weinfield and R.L. York. (1966). *Equality of Educational Opportunity*. Washington, DC: U.S. Office of Education

Conduit, E., Brookes, R., Bramley, G. and Fletcher, C.L. (1996). The Value of School Locations. *British Educational Research* Journal, 22: 199-206.

Dragicevic, S., Schuurman, N. and Fitzgerald, J.M. (2004). The Utility of Exploratory Spatial Data Analysis in the Study of Tuberculosis Incidences in an Urban Canadian Population. *Cartographica*, 39(2): 29-39

Driscoll, D., Halcoussis, D. and Svorny, S. (2003). School District Size and Student Performance. *Economics of Education Review*, 22: 193-201

Flowerdew, R. and Pearce, J. (2001). Linking Point and Area Data to Model Primary School Performance Indicators. *Geographical and Environmental Modeling*, 5(1): 23-41

Fotheringham, A.S. (1997). Trends in Quantitative Methods 1: Stressing the Local. *Progress in Human Geography*, 21(1): 88-96

Fotheringham, A.S., Charlton, M. E., Brunsdon, C. (2001). Spatial Variations in School Performance: a Local Analysis Using Geographically Weighted Regression. *Geographical and Environmental Modeling*, 5(1): 43-66

Fotheringham, A.S., Charlton, M. E., Brunsdon, C. (2002). *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. Wiley, New York.

Fuller, B., Wright, J., Gesicki, K. and Kang, E. (2007). Gauging Growth: How to Judge No Child Left Behind? *Educational Researcher*, 36 (5): 268-278

Fusarelli, L. D. (2004). The Potential Impact of the No Child Left Behind Act on Equity and Diversity in American Education. *Educational Policy*, 18 (71)

Garner, C.L. and Raudenbush, S.W. (1991) Neighborhood Effects on Educational Attainment: a Multilevel Analysis. *Sociology of Education*, 64(4): 251-262.

Gibbons, S. (2002). Geography, Resources and Primary School Performance. Centre for the Economics of Education, Available from: http://cee.lse.ac.uk/pubs/default.asp

Gibson, A. and Asthana, S. (1998). Schools, Pupils and Examination Results: Contextualizing School Performance. *British Educational Research Journal*, 24(3):269-282

Godfrey, K. (1985). Fitting by Organized Comparisons: The Square Combining Table Method. In Hoaglin, D. C., Mosteller F. and Tukey , J. W., editors. *Exploring Data, Table, Trends, and Shapes*. Wiley, New York.

Gordon, J. and Monastiriotis, V. (2007). Education, Location, Education: a Spatial Analysis of English Secondary School Public Examination Results *Urban Studies*, 44(7): 1203-1228.

Hamnett, C., Ramsden, M. and Butler, T. (2007). Social Background, Ethnicity, School Composition and Educational Achievement in East London. *Urban Studies*, 44(7): 1255-1280.

Hannaway, J. and Hamilton, L. (2008). Performance – Based Accountability Policies: Implications for School and Classroom Practices.
Available from: http://www.urban.org/publications/411779.html

Hemphill, C., Apsel, D., Man, C. and Wheaton, P. (2005). *New York City's Best Public Elementary Schools: A Parent's Guide*. Teachers College Press.

Herszenhorn, D. M. (2005, April, 18). The Schools Under Bloomberg: Much Tumult, Mixed Progress. New York Times. Available from:
http://www.nytimes.com/2005/04/18/nyregion/metrocampaigns/18schools.html

Herszenhorn, D. M. (2005, May, 19). A Gold Star for Schools: Overview; 4th Graders post Big Reading Gains. New York Times. Available from:
http://www.nytimes.com/2005/05/19/nyregion/19school.html

Hess, A. G., Jr. (1999). Community Participation or Control? From New York to Chicago. *Theory and Practice*, 38 (4): 217-224

Iatarola, P., Stiefel, L. and Schwartz, A.E. (2002). School Performance and Resource Use: The Role of Districts in New York City. In C.Roelike and J.K.Rice (Eds.) *Fiscal Issues in Urban Education* (pp. 93-116), Greenwich, CT: Information Age

Iatarola, P. and Fruchte, N. (2004). District Effectiveness: A Study of Investment Strategies in New York City Public Schools and Districts. *Educational Policy*,18 (3): 491-512

Japzon, A.C. and Gong, H. (2005). A Neighborhood Analysis of Public Use in New York City. *Library Quaterly*, 75(4): 446-463

Jencks, C. and Philips, M. (1998). The Black-White Test Score Gap. Washington, DC: Brookings Institution

Kane, T. J., Rockoff, J. E. and Staiger, D. O. (2008). What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City. *Economics of Education Review*, 27: 615-631

Kim, J. S. and Sunderman, G. L. (2005). Measuring Academic Proficiency Under the No Child Left Behind Act: Implications for Educational Equity. *Educational Researcher*, 34 (8): 3-13

Lee, J. (2006). Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps: An In-depth Look into National and State Reading and Math Outcome Trends. The Civil Rights Project, Harvard University

Linn, R. L., Baker, E. L. And Betebenner, D. W. (2002). Accountability Systems: Implications of Requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, 31 (6): 3-16

Martin, D.and Atkinson, P. (2001). Investigating the Spatial Linkage of Primary School Performance and Catchment Characteristics. *Geographical and Environmental Modeling* 5 (1): 67-83

Messner S.F and Anselin, L. (2003) Spatial Analyses of Homicide with Areal Data. In Goodchild, M. and Janelle, D., editors, *Spatially Integrated Social Science*. Oxford University Press, New York, NY.

Morenoff, J. D., Sampson, R. J. and Raudenbush, S. W. (2001). Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence. Research Report, Population Studies Center

Mosteller, F. and Tukey, J. (1977). *Data Analysis and Regression*. Addison-Wesley

Murray, A.T, McGuffog, I., Western, J.S. and Mullins, P. (2001). Exploratory Spatial Data Analysis Techniques for Examining Urban Crime. *British Journal of Criminalogy*, 41: 309-329

NCLB. 2002. No Child Left Behind. U.S. Department of Education. Available from: http://www.ed.gov/nclb/landing.jhtml

O'Sullivan, D. and Unwin, D.J. (2003). *Geographical Information Analysis*. Wiley, New York.

Pearce, J. (2000). Techniques for Defining School Catchment Areas for Comparison with Census Data. *Computers, Environment and Urban Systems.* 24: 283-303.

Powers, J. (2003). An Analysis of Performance-Based Accountability: Factors Shaping School Performance in Two Urban School Districts. *Educational Policy*, 17 (5): 558-585

R Development Core Team (2008). R: A Language and Environment for Statistical Computing. Available from: http://www.r-project.org/

Ravitch, D. and Viteritti, J. P. (2000). *City Schools: Lessons from New York*. The Johns Hopkins University Press.

Reardon, S.F. (2003). Sources of Educational Inequality: The Growth of Racial/Ethnic and Socioeconomic Test Score Gaps in Kindergarten and First Grade (Working Paper 03-05R). State College: Pennsylvania State University, Population Research Institute. Available from: http://www.pop.psu.edu/general/pubs/working_papers/psu-pri/wp0305R.pdf

Robinson, G. (2004, July 5). School Reforms. Gotham Gazette. Available from: http://www.gothamgazette.com/article/20040705/200/1026

Saulny, S. and Herszenhorn, D. M. (2005, June, 29). Klein Calls School Year a Success, but a Council Report Differs. New York Times. Available from: http://www.nytimes.com/2005/06/29/nyregion/29school.html

Schwartz, A. E. and Stiefel, L. (2005). Public Education in the Dynamic City: Lessons from New York City. Paper Prepared for the Conference on Urban Dynamics in New York City, Federal Reserve Bank of New York.

Singer, J. (1991). Types of Factors and Their Structural Layouts. In Hoaglin, D. C., Mosteller F. and Tukey , J. W., editors. *Fundamentals of Exploratory Analysis of Variance*. Wiley, New York, pp. 50-69

Stetzer, F. (1982). The Analysis of Spatial Parameter Variation with Jackknifed Parameters. *Journal of Regional Science.* 22 (2): 177-189

Stiefel, L., Bel Hadj Amor, H. and Schwartz, A. E. (2005). Best Schools, Worst Schools and School Efficiency: a Reconciliation and Assessment of Alternative Classification Systems. In Fowler, W. J., Jr. 9Ed.) Developpments in School Finance, 2004: Fiscal Proceedings From the Annual State Data Conference of July 2004 (81-101). Washington, DC: Government Printing Office (NCES 2005-865)

Stiefel, L., Schwartz A. M. and Ellen, I. G. (2006). Disentangling the Racial Test Score Gap: Probing the Evidence in a Large Urban School District. *Journal of Policy Management and Analysis.* 26 (1): 7-30

Stiefel, L., Schwartz, A. E. and Chellman, C. C. (2007). So many Children Left Behind: Segregation and the Impact of Subgroup Reporting in No Child Left Behind on the Racial Test Score Gap. *Educational Policy*, 21 (3): 527-550

Sunderman, G. L., Orfield, G. and Kim, J. S. (2006). The Principals Denied by NCLB Are Central to Visionary School Reform. *The Education Digest*, 72 (2): 19-24

Tukey, J. (1977). *Exploratory Data Analysis*. Addison-Wesley

Tukey, J. and Hoaglin, D. C. (1991). Qualitative and Quantitative Confidence. In Hoaglin, D. C., Mosteller F. and Tukey , J. W., editors. *Fundamentals of Exploratory Analysis of Variance*. Wiley, New York, pp. 336-364

Unwin, A. and Unwin, D. (1996). Exploratory Spatial Data Analysis with Local Statistics. *The Statistician*, 47(3): 415-421

Voss, P. R., Long, D. D., Hammer, R. B. and Friedman, S. (2006). County Child Poverty Rates in the US: a Spatial Regression Approach. *Popul Res Policy Rev*, 25: 369-391

Wheeler, D. and Tiefelsdorf, M. (2005). Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression. *Journal of Geographical Systems,* 7: 161-187

Wolff, J. (2002, May). Community School Boards' Uncertain Future. Gotham Gazette. Available from: http://www.gothamgazette.com/article/education/20020501/6/245