EXPLAINING ACHIEVEMENT DISPARITIES BETWEEN THE UNITED STATES AND SOUTH KOREA

by

Eun Jung Park
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Public Policy

Committee:

______________________________	David J. Armor, Chair
______________________________	A. Lee Fritschler
______________________________	Penelope Earley
	Soo-Yong Byun, External Reader
______________________________	James P. Pfiffner, Program Director
______________________________	Edward Rhodes, Dean
Date: __________________________	Spring 2013 George Mason University Fairfax, VA

Explaining Achievement Disparities between the United States and South Korea

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Eun Jung Park
Master of Public Policy
Georgetown University, 2009
Master of International Studies
Yonsei University, 2007
Bachelor of Arts
The George Washington University, 2003

Dissertation Director: David J. Armor, Professor Emeritus
School of Public Policy

Spring Semester 2013
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of the members of my Dissertation Committee. Professor David Armor, as teacher, advisor, mentor, supporter, and the Chair, guided me through every step of the long, dark, and seemingly endless tunnel of the Ph.D. program. This dissertation is the product of his patience as he went through drafts after drafts reviewing each and every sentence of the work to make sure that I do not go off the tangent. I am truly indebted to his support and trust.

I thoroughly enjoyed Professor Fritschler's insights and perspectives. He never missed the points that needed clarifications and explanations, and at the same time, he has been the keenest supporter of my work. He is an admirable teacher, and he is no doubt my role model and the mentor.

My sincere gratitude goes to Professor Penelope Earley for her valuable advice and insights throughout the dissertation stage. Also, I thank Professor Soo-yong Byun for willingly accepting my invitation to serve on the committee as an external reader. His expertise on the subject matter and his comments on my drafts have been very valuable. I have been lucky with so much support and encouragement. I thank Professor Catherine Rudder for providing me encouragement and believing in me. I also thank Professor James Pfiffner for the support and guidance. I am particularly grateful to Beth Eck and Shannon Williams for providing me with the necessary support to finish on time. I will always remember numerous coffee breaks and happy hours with my colleagues, Nazia, Hezekiah, Callie, Dae-In, Yujin, and Kay, and I thank them for suffering together and cheering me on when I needed encouragement. Also, I thank Sara Allen, Matthew Dickert, and Ian Stanford for editing the draft.

Last but not the least, I dedicate this dissertation to my mother, Chil-Soon, and my father, Heung-Sik. They have been the most ardent supporters of my educational pursuit and they never stopped believing that I can make my own living at some point in my life. I am greatly indebted to their financial and emotional support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ABSTRACT

EXPLAINING ACHIEVEMENT DISPARITIES BETWEEN THE UNITED STATES AND SOUTH KOREA

Eun Jung Park, Ph.D.

George Mason University, 2013

Dissertation Director: Dr. David J. Armor

The educational systems and environments of the United States and South Korea drastically differ, and yet little research has focused on explaining which factors lead to differences in student achievement outcomes. Using the data on mathematics and reading literacy of 15-year-olds from the Program for International Student Assessment, this dissertation aims to reveal how student- and school-level factors are associated with student achievement outcomes within and between these two countries. I find that (1) schools are differentiated in both countries but the extent to which schools are segregated along the line of family SES is greater in the United States than in Korea; (2) Within-country examination revealed that school factors and their relationships with student achievement differ considerably between Korea and the United States. For instance, school autonomy measures have strong and positive relationship with school performance

in Korea, whereas they have no statistical significant relationship with school performance in the United States. Also, school accountability measures, including positing achievement data is positively associated with student achievement in the U.S., but is negatively related with achievement in Korea; (3) Korean educational success is largely attained by the role played by parents, and to a lesser extent by the school factors. Parental involvement in education beyond the family's socioeconomic status is found to have strong relationship with student achievement; (4) Shadow education is the major factor that explains the U.S.-Korea achievement gap. The shadow education participation is positively related with student achievement in Korea, whereas it is negatively related with student achievement in the United States. Whether the finding indicates the causal relationship needs to be further examined. In other words, the question remains whether shadow education causes Korean students to achieve higher and U.S. students to achieve lower, or whether there exists the issue of self-selection, where in Korea, high achievers participate in shadow education to excel higher, and in the U.S., low-achievers participate in shadow education for remedial purposes.

CHAPTER ONE: INTRODUCTION

In recent decades, the availability of international comparative data that ranks countries based on student assessment scores has grown considerably, gaining widespread attention. Particularly for the United States, where improving the quality of education has been a major goal in recent decades, the rankings reported by these international assessment tests are quite disappointing and have caused considerable concern among policy makers. For instance, in the 2009 Program for International Student Assessment (PISA) data, the U.S. students scored significantly below the average of Organization for Economic Cooperation and Development (OECD) countries and below many other industrialized countries in Europe and Asia. South Korea, on the other hand, has had among the highest scores of any country (See Figure 1).

The results of these international assessments have intensified debates on the quality of the U.S. public education. Already in 1983, a major report called *A National at Risk*, in conjunction with other education reports of the early 1980s, criticized the quality of American education as “mediocre”, and argued that reversing the declines in education is essential (Vinovskis 2008, 16). The authors of *A Nation at Risk* warned that the nation is in economic and social danger unless the quality of education is dramatically improved. The report begins:

> Our nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technology innovation is being overtaken by competitors throughout the world…We report to the American people that

> while we can take justifiable pride in what our schools and colleges have historically accomplished and contributed to the United States and well-being of its people, the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people (National Commission on Excellence in Education 1983, 5).

Despite severe criticisms that the report "was too pessimistic and misrepresented the data on student achievement" and "created a false and menacing sense of impending doom", the report was generally well received by the policy makers and the public (Vinovskis 2008, 16–17). Likewise, policy makers referenced international data to make arguments for educational reforms as they believed that using international surveys of student achievement, as in the case of *A Nation at Risk*, is more authoritative than using within-country research (Porter and Gamoran 2002).

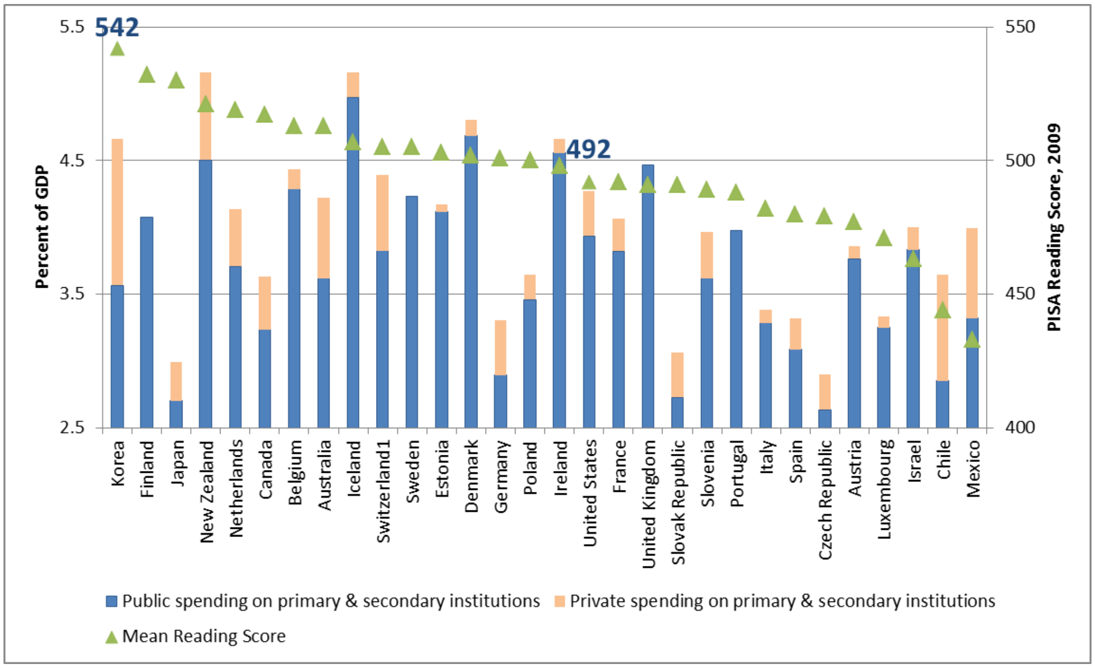

Figure 1 National mean 2009 PISA reading score and expenditures on primary and secondary educational institutions

Source: Education at a Glance (2009), OECD <http://dx.doi.org/10.1787/888932662599>; PISA 2009

Yet, comparing countries like the U.S. and S. Korea based on the relative ranking of their student assessment scores is problematic unless the research takes into account the contextual differences between the countries. For instance, the relative importance of family- and school-related factors is likely to depend on the country's level of economic development (Heyneman and Loxley 1983). In addition, countries have adopted different educational structures (e.g. centralized versus decentralized systems) and may have very different types of educational resources and programs. Figure 1 clearly shows that the public expenditure on primary secondary educational institutions is higher for the U.S.

than for Korea, but when the public and private expenditures are combined, Korea spends more on educational institutions relative to GDP than the United States.

Another example of differing educational resources concerns the use of shadow education, a term coined to refer to outside-of-school educational services (See Stevenson and Baker 1992; Buchmann, Condron, and Roscigno 2010). While some researchers limit the use of 'shadow education' to refer to private, supplementary educational services (e.g. Bray 1999), use of the term has been expanded to include various forms in different cultures. Buchmann et al. (2010) defined shadow education as "educational activities, such as tutoring and extra classes, occurring outside of the formal channels of an educational system that are 'designed to improve a student's chance of successfully moving through the allocation process" (Buchmann, Condron, and Roscigno 2010, 436). And yet, researchers have noted that shadow education plays a different role in different countries. Baker et al. (2001) argued that in some countries (i.e. U.S.) students participate in shadow education for remedial purposes while students in other countries (i.e. Korea) participate for enrichment purposes.

This dissertation examines academic achievement differences between U.S. and Korean students, taking into account differences in their educational contexts. Among other findings, this study will demonstrate that the U.S. and South Korea make very different use of shadow education resources, with dramatically differing outcomes.

U.S. versus S. Korean Challenges

In the U.S., educational policy debates focus on both equity and excellence. Ever since the *Brown v. Board of Education* (1954) Supreme Court ruling, equity and equal educational opportunity have been the major themes of the American public education system. Until recently, however, achievement gaps between white and minority students have not been narrowed despite a series of educational reforms that expanded the role of federal government. According to recent statistics, national high school dropout rate was 8.1 percent in 2009, much improved from 10.9 percent in 2000, but the dropout rates are still significantly higher for minority students and students from disadvantaged families. The dropout rates for Hispanic and Black students were 17.6 and 9.3 percent respectively, compared to 5.2 percent for white students. The dropout rate for students from families with low income was 15.8 percent (Stillwell and Sable 2013).

On the excellence front, the debate grows about the effectiveness of school programs in raising student achievement. A study published almost a half century ago by James Coleman and his colleagues is still dominant in the field of education. The influential Coleman Report concluded that family-related variables are dominant factors influencing academic achievement (Coleman et al. 1966), although Coleman's findings are being challenged again by new research (Konstantopoulos and Borman 2011). On the other hand, research findings on the effect of school-related factors are less consistent. Some researchers have found insignificant and little substantive school effects on student achievement (e.g. Hanushek 1986; Hanushek 1989; Hanushek, Rivkin, and Taylor 1996), while others have found positive and significant school effects (e.g. Greenwald, Hedges, and Laine 1996; Darling-Hammond 2000).

Similarly, the question of whether a series of educational reforms that expanded the role of federal government in education was successful in raising school effectiveness remains to be answered. A successor of Elementary and Secondary Education Act (ESEA), the No Child Left Behind (NCLB) Act was designed to raise achievement outcomes of low-achieving students by setting academic standards and holding schools accountable for learning outcomes. Increased accountability, transparency, and higher standards were expected to raise school efficiency, but the policy encountered criticism for failing to amend major shortfalls of the previous ESEA, including weak implementation and monitoring of the Title I funds (targeted for disadvantaged and at-risk students) and inequities in resource distributions (Manna 2011). Also, supplemental programs, such as the Supplementary Educational Services (SES), which requires school districts to offer after-school remedial instructions to at-risk students free of charge, is also under criticism for poor implementation and monitoring, low participation, and inadequate data for effective program evaluations (Munoz, Potter, and Ross 2008; Ross et al. 2008; Bergeron 2010). Also, the effectiveness of such programs on raising achievement is being questioned (e.g. Munoz, Potter, and Ross 2008; Ross et al. 2008; Farkas and Durham 2007).

The South Korean public education system also deals with the issue of educational equity and excellence (Seth 2002; H. Park 2005; Jaekyung Lee 2007; S. Kim and Lee 2010). South Korea's centralized, rigid system that controls the national curriculum barely offers any exit-based alternatives[1] to students and their parents. High School Equalization Policy, adopted in 1970s, was initially designed to eliminate competition among middle school students entering high schools by eliminating the

[1] Hirschman (1970) *Exit, Voice, Loyalty*

ranking of high schools and “equalizing” the quality of all Korean high schools, regardless of whether they are public or private. In doing so, however, competition among high school students who seek to enter highly-ranked colleges was intensified (S. Kim and Lee 2010). In elementary and middle schools, some forty students are lumped into a classroom based on their age regardless of their academic standing, and they are taught the same materials designed by the central government regardless of their academic aspiration.[2] Consequently, the policy is blamed as “downward equalization [standardization] (S. Kim and Lee 2010), which fails to adequately prepare students for the annual college entrance exam.[3] Park, Hyunjoon (2005) showed, using the 2000 PISA study, that despite their high national mean performance, Korean students in the top 10th percentile scored relatively low on the 2000 reading literacy portion of the PISA – 608 compared to the OECD average of 625 and Americans’ score of 636 at the same percentile.

Notwithstanding the strong criticisms against the policy, the South Korean government could not reach consensus on repealing the Equalization Policy. Parents, unsatisfied with the poor quality of public schools, rely heavily on shadow education to supplement their children’s educational needs. Research shows that the majority of Korean students utilize shadow education, and an individual student spends on average

---

[2] Korea has a single track system following a 6-3-3-4 pattern. Tracking begins at the upper secondary-level where schools are categorized into “academic” and “vocational/technical” schools. Academic schools offer two differ curriculum track, “liberal arts (humanities and social sciences)” and “natural sciences” (see Byun et al. 2012 for more on tracking in the Korean educational system).

[3] About 590,000 Korean students took the nine-hour, multiple-choice college entrance exam, administered by the nation’s Education Ministry in 2008. This exam is administered once every year, and students apply to universities and colleges based on the results. Refer to *The Wall Street Journal* article, “On College-Entrance Exam Day, All of South Korea is Put to the Test” (S. Park 2008) and to *The Economist* article, “The One-Shot Economy” (The Economist 2011).

about USD 242 per month on shadow education (Statistics Korea 2010). Critics argue that household income disparity is likely to lead to a disparity in students' consumption, in terms of both quantity and quality of shadow education, and consequently will lead to a disparity in academic achievement (H. Park, Byun, and Kim 2011). In spite of the criticism that the Equalization Policy itself is creating inequality in educational opportunities, the policy still remains a core educational policy in Korea. Its underlying idea of equality and equal access to a quality public education system is appealing and difficult to overturn.

South Korea faces educational issues that are vastly different from those of the United States. While Koreans' fundamental striving for educational equity and excellence are similar to those of Americans, their educational environment and social contexts are different from those of Americans. Therefore, researchers and policy makers need an in-depth analysis of achievement factors and their relationship with achievement outcomes at the student- and school-levels, rather than a simple comparison of achievement scores.

Cross-National Achievement Comparisons & Large-Scale Data

International surveys of student achievement are useful in understanding achievement-related factors and their effectiveness in improving students' achievement outcomes. Educational researchers agree that "education in one country can be better understood in comparison to education in other countries" (Porter and Gamoran 2002, 4). Among other possibilities, cross-national comparisons will reveal whether the factors that influence achievement outcomes do so differently across different countries, raising the possibility of interactions between the usual achievement predictors and a student's country of residence.

Also, large variations in cross-national surveys allow researchers to examine certain educational phenomena that do not necessarily have sufficient variation for a within-country assessment (Porter and Gamoran 2002, 15). For instance, some countries have centrally controlled public education systems while others have more decentralized systems. Some public education systems are entirely financed publicly while others charge student fees. Some education systems emphasize lecture-based education styles while others lean toward student-centered learning. Some have teachers with a higher social status than others. In this sense, studying other countries' education systems can help us discover alternative approaches to teaching and learning (Porter and Gamoran 2002, 5).

Increased sophistication in terms of data collection, sampling, measurement, and survey methods allows researchers to engage in more rigorous analysis of achievement outcomes across nations. However, due to differences in culture and context, it remains more difficult to conduct international comparative studies. Many methodological advances have been made over the past several decades in international comparative work (Porter and Gamoran 2002). In addition, Raudenbush and Kim (2002) noted that statistical methods have greatly enhanced the capacity of researchers to summarize evidence from large-scale, multilevel surveys such as the Trends in International Mathematics and Science Study (TIMSS) and the International Adult Literacy Survey (IALS).

The fact that these cross-sectional international surveys lack a pre-test makes causal inferences problematic. However, statistical development such as item response models, estimation procedures for multilevel data, and new approaches for handling missing data have greatly enhanced the quality with which international comparative data can be analyzed (Raudenbush and Kim 2002). Similarly, as Porter and Gamoran (2002) noted, these international comparative data have more value in "generating hypotheses" than "testing hypothesis". The comparative data can be used to generate hypotheses, which then can be tested using domestic, panel data. In sum, the increased availability of international survey data and more advanced, sophisticated research methods allow researchers to examine the relationships between critical factors that determine student achievement.

Research questions

The main purpose of this dissertation is to investigate factors that determine student achievement and that might also explain differences between U.S. and S. Korean educational outcomes. Also, by using multi-level analysis, this study examines the relative importance of student- and school-level factors in explaining achievement outcomes in these two countries. While there have been a few studies that examined international comparative research using multi-level modeling (e.g. Buchman 2002; H. Park 2005; Shin et al. 2009), the range of achievement factors examined in these studies were somewhat limited. In this dissertation, I examine whether the wide range of achievement factors are consistently effective (or not effective) in raising student achievement within two vastly different educational contexts of the U.S. and South Korea. More specific research questions include:

1. What role do the family's socioeconomic status (SES) and social/cultural capital factors play in achievement outcomes for Korean and U.S. students? Is there a difference in the relative importance of family background factors on within-country achievement outcomes between Korea and the United States?
2. What role do differences in school characteristics play in achievement outcomes of Korean and U.S. students? What is the relative importance of school-level factors in explaining within-country achievement outcomes in Korea and the United States?
3. What role do shadow education and government policies (e.g. accountability, class size, and teacher performance evaluation) play in explaining between-country achievement outcomes? Do shadow education and government policies contribute to the country differences in achievement outcomes?

CHAPTER TWO: LITERATURE REVIEW

The main purpose of this dissertation is to examine whether the determinants of academic achievement for students is consistent across different educational contexts and environments. To do so, this dissertation examines the relationship of achievement factors in two different educational contexts, namely in S. Korea and the United States. These two countries have many contrasting features of their educational systems such as centralized versus decentralized systems, standardized versus flexible curricula, and a highly regulated system versus one allowing greater autonomy.

A. Educational Structures and Policies: the United States vs. S. Korea

The U.S. decentralized system

A majority of American students attend free, universal K-12 formal education, but the notion of formal education in the United States came into being during the mid-19th century as a local initiative (Timpane 1974). Local communities were responsible for organizing and financing their public schools, and the federal government had minimal intervention in education as the U.S. Constitution does not endow the federal government with any formal role in this area. The dispersed control of educational institutions promotes "wider representation of legitimate interests in education" (Lauglo and McLean 1985, 5) and encourages flexibility in formulating their educational initiatives.

Yet, a decentralized system can generate difficulties in reaching consensus among various stakeholders on educational goals and the means to attain them. The U.S. education system has “functional” decentralization with complex relationships and roles of each governing entity (Hill 2003). Hill argued that education codes are “compendia of laws and regulations” and the pessimists about governance doubt whether those “rules can be so perfectly aligned”, particularly in the complex American system (Hill 2003, 61–62). The U.S. education system not only involves the multi-layered (federal, state, and local) governing bodies but it also includes various stakeholders (e.g. parents, teachers unions, private educational enterprises). In other words, schools have to balance accountability to parents, students, teachers, school boards, and the broader community (Hill 2003).

Variations in the level of educational attainment and achievement among students and local public school districts have been noted as a problem, and the role of federal, state, and local governments in addressing the problem has been debated. The federal government remained on the sidelines of formal education until it became evident that the government has a role in ensuring equality in educational opportunity. The 1954 Supreme Court ruling in the *Brown v. Board of Education* established that the racially segregated education system deprived black children of equality in opportunity because “separate educational facilities are inherently unequal”. Subsequently, the Congress passed the 1964 Civil Rights Act to enforce desegregation of schools. The report entitled “Equality of Educational Opportunity” (EEO also known as the “Coleman Report”), commissioned by the Office of Education in 1964 in response to the Civil Rights Act, identified a large achievement gap between white students and African-American students (Coleman et al. 1966).

Governance and Financing

Researchers have identified the decentralized nature of educational financing as the major issue in the inequality of educational opportunity (Timpane 1974; Moser and Rubenstein 2002). About 87 percent of U.S. students attend public schools and about 16 percent of those students attending public schools live in poverty (see Education Law Center 2012). According to the National Report, the extent to which poverty is concentrated in school districts within states is a striking feature of the U.S. education system. The report shows that 10 percent of school districts have the U.S. Census standard poverty concentrations over 30 percent. Moreover, 17 states serve more than a tenth of their students in high-poverty schools and in five states, over a fifth of the state's students are in such districts (Education Law Center 2012).

Such disparity exists in educational financing across school districts and states because the control and financing of education lie in the authority of local and state governments. According to recent data, about 48.3 percent of total financing comes from states, 43.5 percent comes from local districts, and the remaining 8.2 percent comes from the federal government (National Center for Education Statistics 2009). Because most local districts rely on property taxes to raise school funding, there is a large variation in the amount of local funding (Guthrie 1997; K. Wong 1991; Augenblick, Myers, and

Anderson 1997). Consequently, the inequities in the quality of public education resulting from inequities in the local tax base have persisted (Guthrie 1997).

The role of state governments in education financing is to smooth out the inequities resulting from the unequal tax bases. The California state court decision, *Serrano v. Priest* (1971), argued the resources available to educate children were a function of school-district wealth and not the wealth of the state as a whole. Thus, the inequities in educational financing created by this system violated the equal protection clause of the California constitution. The court decision, along with similar subsequent decisions in other states, played a considerable role in improving the equality of the distribution of educational resources throughout the nation; however, the role of state governments in addressing social inequities is also limited as a large percentage of state funding is used to address territorial inequities (K. Wong 1991).

In this decentralized system, the role of state and federal governments is complicated by other political issues. For instance, researchers pointed out that even when state and/or federal governments have the means to address the equity issue, equitable resource distribution can ignore the classic moral hazard problem (Wood and Theobald 2003). In other words, if local districts recognize that state revenues are forthcoming to the point of equalization, then they have no incentives to raise their own resources.

Furthermore, in a decentralized system, it is difficult to determine the equitable and adequate levels of educational resources. In a centralized education system, educational resources are centrally controlled and generally equally distributed to schools and students. Therefore, students would receive adequate level of funding when the total pool of resources is large. But in a decentralized system, it is difficult to reach consensus among stakeholders on the "right" level of educational financing. This is particularly so because there seems to be no direct link between educational resources and educational outcomes (Hanushek, Rivkin, and Taylor 1996; Armor 2007). The debate on the 'right' measure of educational outcomes further complicates the issue as the social effects of education are indirect and far in the future.

The Increasing Federal Role

The federal government has sought to expand its role of providing financial assistance to students of disadvantaged families during the past several decades. During the 1960s, the election of President Lyndon B. Johnson and a Democratic Congress set the stage for a series of federal programs designed to help disadvantaged students. The passage of the Elementary and Secondary Education Act (ESEA) in 1965, particularly the Title I program, together with Head Start program for early childhood education, was the major educational initiative that aimed to provide equal educational opportunity (Vinovskis 2008, 11). However, because of limited funding (which remained less than 10 percent of total educational financing), poor implementation, and an ineffective funding mechanism, the ESEA remained "a laudable but overly ambitious promise" to its goal toward eliminating poverty and providing equal opportunity to all children (Vinovskis 2008, 11).

During the past several decades, the Congress and presidents with education agendas have sought to enhance the federal role in education by improving the efficacy of ESEA. In 2002, the ESEA was reauthorized as the No Child Left Behind (NCLB) Act. The key provisions of the policy with some subsequent additions follow (Vinovskis 2008, 169–170; Department of Education 2012):

- States determine educational curriculum and set academic proficiency standards;
- School districts must conduct annual testing and report results by race & economic status;
- Progress requirements must be established so that 100 percent proficiency is attained by 2014; and
- Core subjects be taught by highly qualified teachers (HQT) by the end of the 2005-06.

The major educational reforms mainly sought to improve the quality of schools, particularly high poverty schools, through high standards, strict accountability, and improved teacher quality.

The passage of NCLB reflects the government's belief that resources alone are not sufficient to ensure equity in education and programs that strengthen standards and accountability are necessary (Cross 2004). Hanushek and Raymond (2004) hypothesized that the lack of consistent relationships between school inputs and outputs is likely a result of inefficient use of resources due to lack of strong incentives. They argued that states that adopted some form of accountability produced on average higher achievement gains on the NAEP tests than states that did not have such programs (Hanushek and Raymond 2004). They posed two possible explanations: (1) accountability systems may work by virtue of their closure of information about schools; and (2) accountability systems can be effective since they impose both positive and negative consequences to the results of teachers and administrator behavior.

While researchers, including Hanushek and Raymond, believe accountability systems, regardless of any design flaws, are effective in boosting academic performance of schools (Hanushek and Raymond 2004), they remain skeptical of the accountability systems in attaining the goal of educational equity. Hanushek and Raymond (2004) found that accountability significantly increases the state's achievement gain, but both blacks and Hispanics show lower gains on state tests. Similarly, Gordon and Armor (2004) found no significant gains for whites, blacks or Hispanics even in states with stronger accountability systems (c.f. Armor 2007). In sum, the question of whether accountability policies that require schools to test and report student progress encourage more effective and efficient use of resources as to produce better student outputs remains to be answered.

Similarly, NCLB focused on increased attention to teacher education and teacher preparation as a way to improve the quality of public education. A set of policy initiatives focused on designing professional standards, strengthening teacher education and certification requirements, and increasing investments in professional development (see, Darling-Hammond, Wei, and Johnson 2009). Yet critics argued that teacher education has never been the main part of federal education initiatives but only can be characterized as an adjunct to other programs (Earley 2000). Earley (2000) argued that even in the multi-billion-dollar ESEA, “teacher education is embedded in a larger authority as an enabling rather than transforming element,” and the consequence of this adjunct status of teacher education within a large number of separate federal categorical programs “is a patchwork of segmented efforts that individually may have only modest impact on a limited population of teachers” (Earley 2000, 27). Even with the NCLB’s requirement for “highly qualified” teachers, as the definition of “highly qualified” teachers is left up to the states to decide, the states and local districts can “evade having truly ‘highly qualified’ teachers” (Vinovskis 2008, 226).

Public School Choice and Supplementary Educational Services

Despite concerns about inequity and unequal educational opportunities embedded in the system, the role of the federal government in implementing national policy is limited. In fact, handling the differences in 50 states has been noted as a great challenge in fabricating national education policy (Cross 2004). Also, the pro-choice movement that favors flexibility over centralization criticizes the expanded federal role in education. Chubb and Moe criticized the American system as being “too heavily bureaucratic – too hierarchical, too rue-bound, too formalistic – to allow for the kind of autonomy and professionalism schools need if they are to perform well” (Chubb and Moe 1990, 26). Pro-choice supporters contend that under the common public school model, public schools provide a standardized curriculum and treat everyone equally irrespective of social class, culture, race, or religion, and hence the model fails to adequately meet the increasingly diverse needs (Boyd and Miretzky 2003, 5). Some supporters of private-provision of educational services continuously fight for school choice provisions, and in 2002, under the NCLB Act, school choice provision was expanded. Public schools that have failed to meet adequate yearly progress for two consecutive years were required to provide students with options to be relocated to other schools in the district. The public school choice option, however, was criticized on several grounds: (1) potential information gaps in school choice setting; (2) potential social stratification, in terms of racial, religion, and income segregation; and (3) large variability in terms of school characteristics that deter systematic analysis of charter school effects (for review, Vergari 2009).

The ability of charter schools to create more effective and innovative school systems is also in question. As a more innovative form of operating public schools, charter schools received attention from educational policy makers and the public. Charter schools receive approval from a state-authorized agency and are regarded as public schools, but they are freed from other "higher-level authority" outside the school and are regarded as more autonomous and innovative. Yet lottery based enrollment and self-selection issue deter educationalists from closely monitoring the general success of charter schools (Ravitch 2010).

The Supplementary Educational Services provision, which requires school districts to pay for after-school lessons to low-income students attending Title I schools, offers similar concerns. Public schools that fails to meet adequate progress for three consecutive years are required to provide students with privately-operated supplementary educational services for free. Even in the past, the government has financed outside-of-school remedial education for low-achieving, low-income students. Studies of the effectiveness of this approach, however, have been inconclusive. Dynarski et al. (2004) examined the effect of the 21st Century Community Learning Centers Program, one of the federally funded after-school programs that provided funding to 2,250 school districts to support school-based programs in 7,000 public schools, and found that the program had little positive impact for both elementary and middle school students. Furthermore, while research suggests that one-to-one tutoring is the most effective form of remediation (Wasik and Slavin 1993; Farkas and Durham 2007), critiques are less optimistic about the efficiency of the program in closing the achievement gap (Farkas and Hall 2000).

Summary

After almost a decade of government efforts to improve public education through an increased federal role, expanded supplementary educational services, and school choice, there does not seem to be a significant rise in student achievement or closure of the racial achievement gap (Jaekyung Lee 2006). Student achievement in reading, based on their results in the National Assessment of Educational Progress (NAEP), remained flat after NCLB was implemented and achievement in math is growing at the same rate after NCLB as before (Jaekyung Lee 2006). PISA results provide some optimistic picture on the effectiveness of the education reforms. Compared to the U.S. students' reading scores in 2000, which had a mean value of 504 with a standard deviation of 105, the scores increased to a mean value of 513 with a standard deviation of 87 in 2009. In 2009, the score for the 10th percentile was 398 points compared to 363 points in 2000. The score of U.S. students in the 90th percentile, however, decreased from 636 in 2000 to 625 in 2009. Yet even with the increase in reading scores of low-achieving students (e.g. students at the 10th percentile), the mean score of U.S. students is lower than those of other OECD countries, including South Korea (m=552), Finland (m=546), and Canada (m=525). Researchers tend to agree that the major educational initiatives undertaken by the U.S. government have failed to substantially improve learning for K-12 students (for review, Vinovskis 2008) and that corrective measures to improve the nation's educational system are much needed.

The centralized S. Korean system

After the destructive Korean War, the young Korean government was limited in its capacity to establish governing agencies, including the educational system. In 1945, at the time of liberation from Japanese colonial rule, only 65 percent of children at the age of primary school and 20 percent of children at the age of secondary school were enrolled in schools (S. Kim and Lee 2010). The Republic of Korea (S. Korea) was established in 1948, and the newly established democratic nation sought to expand and establish a foundation for democratic education by provide universal education to its citizens. But, the nation was extremely impoverished as a consequence of the prolonged Japanese occupation and the Korean War (1950-53) because the government had a low taxing capacity. Its effective tax rate was just 9.9 percent of its nation's GDP, and because of extreme corruption (Seth 2002), the government lacked adequate capacity to finance the expansion. Consequently, the Korean government experienced difficulties in keeping up with the rapidly expanding student population and experienced shortages in educational capacity. The government contributed only about 10 percent of its total cost of education, mostly for building schools and educational facilities, and the rest was financed through student fees and parent contributions (Seth 2002; S. Kim and Lee 2010). In the 1960s and 70s, due to quantitative expansion, Korean education suffered from over-crowded classrooms, oversized schools, a shortage of fully qualified teachers and educational facilities, as well as intense competition in the college entrance system (Ministry of Education, Science, and Technology 2012).

Increasing income and demand for education, together with the limited educational facilities, resulted in excessive competition for better schools. Schools were allowed to select their students by administering their own admissions procedures and elite high schools were able to attract students in the top tier, leading to serious inequalities between these elite schools and the rest in terms of student and teacher quality (Byun and Kim 2010). Consequently, parents with economic means sought to secure various ways to ensure entrance to high quality secondary education, including hiring private tutors for entrance exams. The phenomenon called *ipsi-jiok* (entrance examination hell) plagued the society, and the nation called for a nation-wide campaign to "rescue some 1.8 million middle school students from the *ipsi-jiok*" (Korea Policy Portal 2007). The Ministry of Education, Science, and Technology (MEST) notes that such shortfalls necessitated the reform of the entrance examination system to normalize education at all school levels (MEST 2013a).

Equalization Policy and Central Control of Education

In the early 1970s, equalization policy was adopted by the Park administration as a way to alleviate the competition. The policy exemplifies the South Korean government's rigid and uniform educational policy that reflects its determination to centrally control its children's educational outcomes. The middle school (grades 7-9) equalization policy was first adopted in 1969 and spread throughout the country by 1971, and the high school (grades 10-12) equalization policy was first implemented in 1974 and was expanded to other major cities in the 1980s. The policy forfeited individual school's authority, regardless of whether it was privately or publicly managed, to select its students through individually administered entrance exams, and instead established a system of random allocation of students within separate school districts (S. Kim and Lee 2010). By lottery, students were randomly assigned to schools in the district, and under the system, even private schools were forced to take all students assigned by the Ministry of Education (S. Kim and Lee 2010).

Yet the policy remains controversial. First, the government was able to reduce competition among middle school students but not among high school students who seek to enter higher education (S. Kim and Lee 2010). Second, the policy is blamed as "downward equalization policy" (Byun and Kim 2010; S. Kim and Lee 2010). In a large, heterogeneous classroom setting where students are assigned based on age and not on ability, teachers face a challenge in standardizing the level of instruction. Furthermore, while the purpose of the policy was to reduce the demand for private tutoring, the demand drastically increased (S. Kim and Lee 2010; Seth 2002). Because parents believe that public schools that provide the same quality of educational services to everyone do not adequately prepare students to survive the severe competition for university admission and job search, they resort to supplementary educational services (Seth 2002).

During the 1980s, the government focused on educational aspects related to the normalization and improvement of educational quality (MEST 2012). The government established a broadcasting system dedicated to programs on education. It sought to diversify schools by establishing "special purpose" schools for religion, athletics, arts, agriculture, foreign language, and science specialty. Also, it adopted the educational tax system to finance educational reforms. Furthermore, the government prohibited participation in private tutoring. Until the prohibition of the use of private tutoring was ruled unconstitutional in 2000, the educational practice remained illegal and prospered only in the black market (Bray 2009).

The central control of education has its advantages in attaining educational goals. The hierarchical relationship of educational decision-making authorities makes it easier to establish clear objectives and goals in education, and it also makes it easier for the government unit to formulate long-term planning that can produce desirable outcomes since the governing environment is likely to produce predictable funding and a stable regulatory environment. Korean education is highly standardized in line with the government's goal of egalitarian education. Both public and private schools were highly regulated by the central government in terms of tuition fees, teacher salaries, and the school curricula (S. Kim and Lee 2010; Byun and Kim 2010).

And yet, the central control of education has resulted in highly standardized and uniform educational outcomes, and the government policies have been criticized for undermining educational excellence (Byun and Kim 2010). Critics noted the monopolistic power of the concentrated authority on education. In such systems where one publicly-operated and financed education system dominates, there often “lacked meaningful accountability and tangible incentives to improve, that it exhibited the characteristic flaws of a command-and-control enterprise, that it enjoyed a virtual monopoly” (Peterson 2003).

In spite of the relatively higher achievement scores of Korean students on international assessment surveys compared to their counterparts in other developed countries, the public has demanded that the government improve the quality of public education. Despite the rapid economic growth the country has undergone during the past several decades, public expenditure on education did not experience similar growth. Public expenditure on education remains at about 3 to 4 percent of the country’s GDP while the private sector spends about 7 percent of the GDP on education (Jisoon Lee 2001). The public share of total education spending at the primary and secondary levels has actually declined from 81 percent in 2000 to 76 percent in 2009 (OECD 2012a). This figure is well below the OECD average of 92 percent. Consequently, parents provide their children with supplementary educational materials, textbooks, school uniforms, and they spend close to 25 percent of their household income on education (Jisoon Lee 2001). To complement the inadequacy in public education, a large portion of school-aged children rely on shadow education. More discomforting than the statistics on the proportion of students that use private tutoring is the inequality in educational spending; recent statistics indicate that families in the top 20th percentile of the income distribution spend more than 8 times the amount of families in the bottom 20th percentile in shadow education (Statistics 2011). The nation acknowledges the need for major education reforms that aim to improve the quality of public education and to subsequently lower the private burden on education.

Current Educational Policies

Current educational policies involve increasing the diversity in educational services to promote excellence and to reduce private expenditures on supplementary educational services. Below is a summary of key educational policies listed on the website of the Ministry of Education, Science, and Technology:

- Excellent schools and diverse curricula
- Creativity and character education
- Teacher expertise
- Reduction of private education expenditure

Diverse Curricula:

The Ministry of Education and Human Resources Development (MOE)and the Ministry of Science and Technology (MOST) was merged in to the Ministry of Education, Science, and Technology (MEST) in 2008, and it oversees the national education system and the national curriculum to ensure equal educational opportunity for all and maintain the quality of education (J.-H. Lee 2012). The government has undergone seven curriculum revisions “to meet national and social needs” (MEST 2013a). The Seventh Curriculum was introduced in 1997 for primary-level students in 2000 and was expanded through upper secondary level by 2004. The Ministry of Education, Science, and Technology states that the Seventh Curriculum “attempts to break away from the spoon-fed and short-sighted approach to education of the past towards a new approach in the classroom to produce human resources capable of facing new challenges” (MEST 2013a). Likewise, the goal of the curriculum revision was to acknowledge individual differences and their different needs and to emphasize “individual talent, aptitude, and creativity” (MEST 2013a). During the 11th and 12th grades in the upper secondary school, students are allowed the opportunity to choose their curriculum and courses to fit their needs to best “[facilitate] their future path” (MEST 2013a).

In tandem with curriculum differentiation, grouping practices have increased in Korea (Byun and Kim 2010). Also, some private schools were allowed greater autonomy in terms of curricular and financial management.

Education for Creativity:

Korean education has been noted as instruction-based, memorization-based learning. Increasingly, the need for education that fosters creativity and innovative thinking has been emphasized. Shifting from rote-based and teacher-centered instruction, the government is expanding the “departmentalized classroom system”, in which students move to different classrooms for different subjects (MEST 2013b). The MEST states that this style of instruction “boosts students’ interest in learning as it provides a tailored environment for each course” (MEST 2013b). In addition, the government seeks to expand extracurricular activities that involve hands-on activities “in order to develop good character and a sense of responsibility towards the community” (MEST 2013b).

Teacher Expertise:

Even during the earlier years of institutional formation, the government put a heavy emphasis on teacher education. During the 1960s and 70s, the government established the Graduate School of Education to provide in-service training and education for teachers. The central government highly regulates teacher qualification, and in order to be fully qualified to teach, teachers are required to have a credential or license in addition to the education diploma (see OECD 2012a). Despite the highly selective process, Korea experiences over-supply of teachers (E. Kim, Kim, and Han 2009) in part due to the social status and stability teachers enjoy in the Korean society. As teachers are highly respected and competitively compensated, competition for entrance to universities of education tended to be fierce and only those at the top tier in their class could secure admission to these universities. For instance, teachers in Korea are paid starting salaries of about 30 percent higher than other full-time, full-year workers with similar level of education (OECD 2012a). Likewise, teacher quality tends to be generally high in Korean schools compared to schools in other countries. According to the OECD report, Korea is one of a few countries that require applicants to pass competitive exams to enter pre-service training programs for all levels of education (Finland, Greece, Ireland, Israel and Turkey are the others) (OECD 2012a, 489).

Even so, the government has blamed much of the public's dissatisfaction towards public education on teachers (E. Kim, Kim, and Han 2009) and called for improved teacher quality. In 2010, the government has adopted a teacher evaluation policy "as a way to monitor and improve teacher quality" (MEST 2013b). It aims to "build up teachers' professional quality" by requiring teachers with poor evaluation scores to receive supplementary training (MEST 2013b). It gives teachers with high performances the opportunity for personal research or education at institutions at home and abroad for them "to upgrade their expertise" (MEST 2013b). They are designated as "Master Teachers" who give consultations to peers and develop and distribute effective teaching methods (MEST 2013b).

Reduction of Private Education Expenditure:

Various policy measures have been implemented to curb private spending on private education. The government has constantly argued that cutting expenses on private education is a challenge to be overcome to secure fair educational opportunities for all students (J.-H. Lee 2012). Ranging from complete banning of shadow education to establishing a free broadcasting educational channel, the government has actively sought to reduce the demand for shadow education. The Ministry states that as a result of various government policies, including establishing vocational educational systems, diversifying curricula, promoting creativity and character education, and enhancing teacher expertise, the growth in the private spending on private education has decreased (J.-H. Lee 2012). It argues, "spending on private education, which has increased 10 percent every year, went down to 3.5 percent for the first time in 2010" (MEST 2013b).

Whether these policy measures are effective at reducing private education expenditures remains inconclusive. For instance, the special purpose high schools, which were created to meet individual student's educational needs and to educate talented students in foreign languages and science, are recognized as "a short-cut to prestigious universities among parents" (J. Kim and Shin 2012, 30). Subsequently, competition to enter these high schools intensified, resulting in higher demand for shadow education among middle school students and parents (J. Kim and Shin 2012). Also, the reduction in shadow education expenditures could be a result of the 2007-08 economic recession and not necessarily a result of successful policy measures.

Summary

The government has set forth a direct policy measure to reduce the financial burden of private households. After-school programs were offered at affordable rates and broadcasting services were implemented to provide free supplementary educational services. However, critics argued that these policy measures have not adequately contributed to the ease of the *ipsi-jiok* (admission hell) and shadow education frenzy (Seth 2002). The government, however, seeks to expand after-school programs based on the reasoning outlined as follows: (a) supplementing the educational function of schools by provision of diverse programs that may not be offered through the regular curriculum,

(b) alleviating private tutoring expenses, (c) actualizing the educational welfare through narrowing the educational gaps between social classes and regions, and (d) bridging a partnership between schools and community through after-school programs (c.f. H.-J. Park, Byun, and Jo 2012, 5). Consequently, it is of interest to researchers whether the current educational policies have been effective in achieving the educational goals of the Korean government. It is particularly important to examine the demands and effects of shadow education on student achievement as critics remain skeptical about the effect of publicly-funded after-school programs and on their ability to reduce public demand for shadow education (H.-J. Park, Byun, and Jo 2012).

Summary

The U.S. and S. Korean governments share similar concerns on the quality of their educational systems. They seek to promote educational equity and at the same time, improve the quality of education. Because these countries' public education systems have different characteristics, they have implemented contrasting policies to attain their educational goals. The U.S. government focuses on establishing standards and a national curriculum, while the Korean government emphasizes diversity in curriculum. These countries similarly emphasize on strengthening teacher qualifications, increasing school choice, and expanding the financing of supplementary educational services.

B. Factors Affecting Academic Achievement

School effectiveness can be measured in a variety of ways, but this dissertation focuses on the effectiveness of schools in raising students' academic achievement.

Because achievement is a function of family background factors, school-level factors, and other peer group characteristics, it is necessary to properly control for other determinants in order to closely estimate school effects.

Much of the empirical literature on educational achievement is based on the educational production function. A standard production function, or input-output model, focused on relationships of various input variables and educational output or outcome. This approach assumes that output of the educational process, or achievement of individual students in this case, is related directly to a series of inputs (Hanushek 1989). These inputs typically include a student's genetic endowments or learning capacities and family characteristics, which are not controlled by policy makers, and various school characteristics such as teachers, curricula, and resources that are controlled by policy makers. Also, educational researchers consider peer inputs, measured by aggregate summaries of the socio-demographic characteristics of other students in the school, when estimating the production function (Summers and Wolfe 1997; Hanushek 1989; Willms 2010). Hence,

Equation 1 Educational Production Function

A = F (Genetic &family factors, school characteristics and practices, peer group effects)

This equation is a reduced form, and the accuracy and reliability of the results depend on a variety of measurement, sampling, and technical estimation issues (see Hanushek 1986; Hanushek 1989). Also, this equation fails to capture other major unobservable variables, such as attitudinal variables that affect achievement but are themselves determined by other inputs (Summers and Wolfe 1997). More importantly, educational researchers began paying more attention to the process of these inputs, or the relationships among the input variables. Summers and Wolfe (1977) pointed out that studies have either relied on an additive form of the production function or have segmented and estimated the sample based on the family factors, but they have generally failed to examine the interaction terms to allow the effect of policy variables to vary with the family background factors. Furthermore, Hanushek (1986) pointed to “differential effects” of school inputs on achievement outcomes. He argued that existing measures of the characteristics of teachers and schools are seriously flawed and thus are poor indicators of the true effect of schools (Hanushek 1989), and because of the nature of social science, in which natural experiments are not feasible, it is virtually impossible to control for all of the unobservable factors and make causal claims based on the production function.

Notwithstanding such limitations, past studies have examined and revealed relationships of various input variables and achievement. Yet, despite the large quantity of studies on the topic, no consensual agreement seems to have been reached among educational researchers. This section reviews past literature on the list of family background factors and school factors that have been shown to affect academic achievement. In addition, past research on the effectiveness of government policies and practices intended to raise student achievement outcomes is summarized.

Family-level Factors and Academic Achievement

Ample research has focused on proper and adequate ways of measuring family background factors. A good measure of family socioeconomic status (SES) and other background characteristics is crucial because family background factors have been noted as the strongest correlates of student achievement (Coleman et al. 1966; White 1982; Hauser and Sewell 1986; Baker, Goesling, and LeTendre 2002; Buchmann 2002; Rothstein 2004; Sirin 2005; Armor 2007). This line of research argued that the impact of family's socioeconomic status on children's educational outcomes is deterministic and it plays a role in intergenerational transmission of status, through the transmission of financial capital, cultural resources, and social capital from parents to children (Buchmann 2002).

Also, the importance of properly measuring family background factors is to allow adequate controlling for family background factors that could have confounding effects with other variables of interests on academic achievement. In one of the major educational studies conducted in the United States, Coleman, in his attempts to measure the impact of schools on students' learning outcomes, noted the importance of controlling variations in family background factors and the subsequent variations in student inputs (c.f. Buchmann 2002). Coleman et al. stated:

> In the attempt to discover effects of school factors on achievement, perhaps the principal villain is the fact that student populations in different schools differ at the outset, and because of this difference, it is not possible merely to judge the quality of a school by the achievements of the students leaving it. It is necessary to control in some way for the variations

> in student input with which the teachers and staff of the school are confronted (Coleman et al. 1966, 395).

Researchers have sought ways to effectively measure the impact of family background factors on educational success. Some noted that SES is a complex and multidimensional concept and thus a single composite measure of SES can create ambiguity in interpreting research findings (White 1982; Sirin 2005). Also, because many of the family-background factors are highly correlated and heavily confounded, Sirin (2005) argued that the strength of these variables on achievement depend on which set of family-related variables are included. However, researchers who are mostly interested in the mediating role of family background factors on the educational outcomes, usually combine parental education, occupation, and income variables in to a single composite measure of socioeconomic status (see Buchmann 2002, 153).

Family's socioeconomic status is measured by parental education, parental occupation, and parental income, and the tripartite measures of family status are highly interconnected. Parental education and parental occupation are highly correlated and mostly determined by each other. Jencks and others (1981) pointed out that those who have higher level of educational attainment are more likely to secure jobs with higher prestige. Furthermore, parental occupation and parental income are also closely related. Those who have occupations with higher prestige are more likely to earn higher income.

Similarly, in meta-analyses conducted by White (1982) and Sirin (2005), family SES was confirmed to be one of the strongest correlates of children's academic success. Among the tripartite SES factors, Sirin (2005) found that parental education is the most commonly used SES component and they concluded that parental education is the most important predictor of academic achievement. Blau and Duncan (1967) further argued that a father's education and occupational status explain the son's educational attainment and that the father's education, occupation, and income explain the son's occupational status.

Also, parental income and childhood poverty are highly correlated and they are closely associated with children's educational success. Poor children living in high-poverty communities, as compared with their counterparts residing in communities with lower rates of poverty, have limited access to jobs and high-quality public and private services, including child care, schools, and community centers (McLoyd 1998). Poverty, measured by eligibility for free lunch, is one of the variables with the largest effect size on achievement (Sirin 2005).

The impact of family structure on children's educational success is better understood in relation to family wealth. Children growing up in a single-parent household are more likely to experience poverty than children in a two-parent household. In 1992, about 45 percent of children from a single-mother household were living below the poverty line (McLanahan and Sandefur 1994). The risk of poverty in single-mother families is particularly high for many reasons, including inadequate child support from fathers, low wages for women, and low community support (McLanahan and Sandefur 1994; McLoyd 1998). McLanahan and Sandefur (1997) argued that loss of income due to parent separation costs schooling achievement, by lowering the quality of schools children attend and by leading to lack of parent involvement in schools.

When Coleman reported that the inequalities in educational outcomes resulted from the inequalities in socioeconomic status cannot easily be overcome, an intense debate ensued about whether what families "are" (i.e. status) matter more than what families "do" (i.e. process) for educational success. Arguments that genetic heredity and resource transfer are keys to educational success were on one line of thought. Herrnstein and Murray (1996) argued that parents exert genetic and heredity influence on their children's cognitive ability and that the ability cannot easily be altered. Similarly, researchers argued that children from disadvantaged families possess a lower chance of educational success because the family's financial resources are directly related with the parents' decisions on resource allocation (Becker 1991; Jencks et al. 1972; Sirin 2005). In fact, more disadvantaged individuals and groups tend to have less access to educational services, particularly preschools and higher educational services, where the cost of attending is high (Jencks et al. 1972).

Researchers examined how much of what parents "do" matters on educational success. Sirin (2005) argued that parents' location in the socioeconomic structure has a strong impact on students' academic achievement, not only directly by providing resources at home but also indirectly by providing social capital that is necessary to succeed in school (p. 438). Furthermore, the indirect influence of poverty is as important as the direct influence of material hardship. Gershoff and others focus on "the *processes* by which family income affects children" (Gershoff et al. 2007, 70). In particular, material hardship affects parenting because material hardship influences parents' mental health and relationships with partners. They argued that the stress of raising a family on a low income is posited to negatively affect parents' mental health and behavior and in turn, to negatively affect children (Gershoff et al. 2007, 70).

Also, the effect of status variables is oftentimes mediated by the process variables. Armor (2003) argued that while there is a close correlation between parents' IQ and children's cognitive ability, the relationship between parents' IQ and children's IQ becomes weaker when the relationship is adjusted for environmental factors. Mothers with higher IQ and higher levels of educational attainment are likely to create more favorable home environment by providing better parenting, emotional support, and nurturing (Armor 2003).

Similarly, whether or not a child lives in a single-mother household is an important predictor of educational success, not only because of the loss of financial resources but also because of parenting practices. Children in single-mother families spend less time with their mothers than children in two-parent families, and single mothers exercise less control over their children than mothers in two-parent families (McLanahan and Sandefur 1994). Also, single-motherhood is problematic because it is often associated with teenage, out-of-wedlock pregnancy. Armor (2003) argued that teenage mothers have a higher risk of having low birth weight children, pay inadequate attention to children's nutrition, and are less likely to be skilled in favorable parenting behavior for children's intellectual development. Furthermore, teenage mothers are more likely to drop out of high school and less likely to attend higher education (Armor 2003, 83). In sum, along with the loss of economic/financial resources, family structure and a single-mother household affect children's cognitive development and educational success through practices and processes.

Studies were conducted in a range of countries to examine the role of social origins in determining educational and occupational status and mobility (Buchmann 2002). Buchmann (2002) reviewed a representative sample of international research on the relationship between socioeconomic status and educational attainment and achievement since 1970. She found that most studies in the sample conceptualize socioeconomic status as either the father's education and occupation or a composite measure of these and other family background factors (Buchmann 2002, 155).

She also noted the systematic approach to the measurement of family background factors in the sample of international literature. Occupational status, for instance, is measured via scales that have been developed to generalize the prestige associated with occupations across a wide range of societies. The Socioeconomic Index (SEI) scale, created by Duncan (1961) was one of the earliest to be formulated and she noted that much of the international literature used a modified SEI scale for father's occupational status. Also, in an effort to create a cross-culturally reliable scale, the Standard International Occupational Prestige (SIOP) and the International Socioeconomic Index (ISEI) of occupational status were formulated based on the Duncan SEI.

Similarly with the case of occupational status, scales have been created to measure educational attainment with the goal of ensuring cross-national comparability (Buchmann 2002, 164). The International Standard Classification of Education (ISCED) and the Comparative Analysis of Social Mobility in Industrial Nations (CASMIN) categories were used extensively to measure and facilitate comparative research on social stratification and mobility. The importance of including mother's education has received attention, particularly in cases where males are absent from the household, and in many cases, mother's education is used as a measure for parental education. This approach is generally accepted by the scholarship, because in many cases, maternal and paternal education levels are highly correlated (see Buchmann 2002, 164). Similarly, researchers use the higher of the two parents' education levels as a measure of parental education. In other cases, researchers use the sum of both parents' schooling.

Buchmann (2002) noted that developing a cross-nationally comparable measure of family wealth or parental income has been a challenge for international researchers. For one thing, it is difficult to get high response rates on income survey questions (p. 165). Also, income and wealth categories are seldom compatible cross-nationally (p. 165). Indices of home possessions were used instead as a proxy measure for family wealth, and some researchers noted that these indices better capture long-term effect of family wealth, rather than simple flow of wealth (see R. Wong 1998; Buchmann 2002).

School-level Factors and Academic Achievement

The effects of school-related factors have also been studied extensively in the research literature. After controlling for family background factors, some researchers found that school-related factors exert little substantive influence on academic achievement. The Coleman Report noted that "the inequalities imposed on children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the end of school" (Coleman et al.

1966, 325), and the finding was confirmed by a number of scholars (Hanushek, Rivkin, and Taylor 1996; Hanushek 1999; Armor 2007). Because cognitive skills and learning behaviors are largely established before children enter school, school factors are not likely to override the large effect of non-school factors (Armor 2007).

Contradicting the findings conducted in the U.S., more substantive impact of school-related factors on academic achievement were found in studies of developing countries. In the earlier years, Heyneman (1976) published an article, referred to as the "Coleman Report for a developing country", in which he replicated the design of the Coleman report and found significant effects of school facilities and weak effects of family background on academic achievements of students from 67 primary schools in Uganda. In subsequent research, Heyneman and Loxley (1982; 1983), using data from the 1970s, concluded a stronger impact of school-related factors on academic achievement relative to family SES factor in developing countries. Heyneman and Loxley argued the strong impact of school factors in developing countries is possibly resulted from the weak correlation between family background and school factors, resulting in less ambiguity in the efficacy of school physical facilities and teachers (Heyneman and Loxley 1983, 1180). Recently, the Heyneman and Loxley's findings have been challenged (Baker, Goesling, and LeTendre 2002).

Similarly, Fuller and Clarke (1994) reviewed school effects studies in developing countries and found some reports that show positive school effects on achievement. They found that many studies reported simple school resources variables, concluding the availability of textbooks and supplementary reading materials and the presence of a school library are significantly related with achievement at the primary level. They noted, however, that not only is the amount of studies conducted at the secondary level lower than those at the primary level, but the school effects research is less consistent. Also, the effect of teacher attributes, including teachers' subject-matter knowledge, experience, and salary, was found to be rather mixed both at the primary and secondary level (for review, Fuller and Clarke 1994).

Likewise, while scholars tend to reach a consensual agreement on the importance of family background factor, research on school effects revealed mixed results, across both the level of development of the country and the level of education. Below is the summary of school inputs generally found in the literature; these include school sector (private vs. public), school resources, teachers, and school policies and practices. A large number of existing studies have sought to reveal their effects, but there exists large variations in data collection, research methods, and study findings.

1. School Sector

One of the school effects debate is centered on the school sector question. Since the Coleman, et al. (1982) study on the effects of Catholic and other private schools on students' educational outcomes, the private versus public school debate has been on-going. Coleman et al. (1982) concluded that students in Catholic and other private schools score better on achievement tests than do students from public schools. Also, average scores of Catholic and private schools are higher than those of other sectors. Using the first wave of "High School and Beyond" data, they concluded that the public-private school differences at the sophomore and senior level are one-third to one-half of a standard deviation (Coleman, Hoffer, and Kilgore 1982, 68). Even after a range of family background factors are accounted for, the authors argued "from a low of one-fifth of the initial difference to a high of about four-fifths of the initial difference" remain (Coleman, Hoffer, and Kilgore 1982, 71).

The findings gave weight to the voices of proponents of privatization of public schools who claimed that the governance structures of public schools are prone to inefficiencies. For instance, Chubb and Moe (1990) argued that the American educational system is "bureaucratic and political" and that "the bureaucracy problem is the more immediate explanation" for the poor performance of U.S. schools and that "the political problem is the more fundamental" which explains the bureaucracy problem (Chubb and Moe 1990, 26–27). They argued for a market-based approach to educational governance –including innovative approaches to teaching and learning and competition among schools. They saw that public schools (or "old institutions") were "owned" by vested interests, including teachers' unions, administrators and school boards, and argued for a replacement system where almost all "higher-level authority" outside the school was eliminated (Chubb and Moe 1990, 39; Hess 2003).

In more recent studies, however, a group of researchers found contradictory results. In a cross-national study conducted using the 2006 and 2009 PISA surveys, the authors found a positive and statistically significant effect of public schools on math and science scores of students in 10 large developed OECD countries (Sousa, Park, and Armor 2012). Similarly, S. Lubienski and Lubienski (2006) argued that the private school-public school differences are largely due to differences in the student populations these sectors serve. Using a nationally representative sample of elementary and middle school students, the authors examined whether achievement differences persist across students attending different sectors, even after a range of demographic characteristics are controlled for. They found that demographic differences between public and private schools accounted for the relatively high raw NAEP mathematics scores of private schools (S.Lubienski and Lubienski 2006, 679). In subsequent research, S. Lubienski, Lubienski, and Crane (2008) argued that students' math scores in private schools tend to be lower than their counterparts in public schools, because private schools tend to ignore existing national expertise regarding mathematics curriculum, instruction, and teacher education (S. Lubienski, Lubienski, and Crane 2008, 133). The authors noted that schools are not a "black box," but the actual internal processes of schools matter (S. Lubienski, Lubienski, and Crane 2008).

2. School Resources

The impact of school resources on educational outcomes is another controversial variable in educational research. Some researchers argued that the longitudinal data on school spending and student achievement reflects that school resources have little or no impact on achievement outcomes. Despite a substantial growth in school expenditures over the past several decades, Hanushek, Rivkin, and Taylor (1996) argued that American students educational outcomes remained relatively flat. They noted that despite the rise in real educational expenditures at an annual rate of 3.5 percent, the school attainment of Americans remained relatively flat since 1970s. Also, various test measures, including the SAT, indicate a rather flat level of academic performance (Hanushek, Rivkin, and Taylor 1996). Others criticized the use of aggregate measures to examine the relationship between school expenditures and educational outcomes. Hedges and Greenwald (1996) pointed out that for the several past decades, poverty rates and divorce rates have increased and have negatively affected student achievement. They contended that the fact that achievement has not declined substantially during the 1972-90 period when the country experienced a decline in favorable home environment or social capital indicates that increased school expenditures had a positive effect on student educational outcomes.

Specific school inputs were examined separately and together with other school resource variables to determine their effects on achievement. The effect of school size (enrollment) and class size were frequently examined. Smaller schools were often found to be closely related with higher school performance (Darling-Hammond 2000). However, whether the size of the school is independently related with performance is not clear as there are other issues (e.g. urbanity, school sector) involved (for review, S. Lubienski, Lubienski, and Crane 2008). In terms of class size, a doctoral dissertation by Black (1954) reviewed the effect of class size on achievement, and while thirty-five of the surveys found smaller classes to be more effective, there were another thirty-two which found the effect to be inconclusive (c.f. Summers and Wolfe 1977). In an experimental study conducted by Finn and Achilles (1999) in Tennessee, class size reduction had positive effect on achievement, particularly for minority students and low-achievers. But the validity of their research design was challenged by other researchers (e.g. Hanushek 1999).

Hanushek (1989; 1999) argued that teacher-pupil ratio is not systematically related with achievement and the effect of class size is too small to be cost-effective. Even in studies conducted in developing countries, the effect of class size is not significant. Only 2 of 21 studies reviewed by Fuller and Clarke (1994) have revealed a significant effect on achievement.

The effect of quality of school facilities on students' academic performance seems to depend on the level of the country's development. Simple physical facilities, e.g. whether student has access to more or less playground space, a new or old building, or better building conditions, did not affect students' achievement (Summers and Wolfe 1997). Hanushek (1989) found that of 74 estimated coefficients of the studies that examined facilities in the U.S., only 12 of them had statistically significant effect on achievement, and only 7 of these had positive effect. He concluded that there is no systematic relationship between school facilities, ranging from spending to individual characteristics, and achievement (Hanushek 1989, 47). Studies conducted in developing countries, however, found a positive effect of the quality of school facilities on achievement. Of eight studies reviewed by Fuller and Clarke, six of them found positive effects on achievement of students at the primary level. Only one study examined the effect of the quality of school facilities on achievement at the secondary level, and the study found a positive effect (see Fuller and Clarke 1994). These findings suggests that, as Summers and Wolfe (1997) noted, "although school inputs do not seem to relate to achievement growth, it does not mean that reducing these expenditures to zero is the logical policy recommendation" (Summers and Wolfe 1997, 646).

Teacher salary is another variable that is determined by the level of school resources; however, only a few studies have found any direct influence of teacher salaries on student achievement. In most countries, teacher salaries are linearly related with the teacher's seniority, and Fuller and Clarke (1994) noted that in rapidly expanding education systems, younger teachers, who are paid less than more experienced teachers, often are better trained. Researchers found that teacher training levels are related to achievement, but the salary levels are not independently associated with achievement. In the U.S., Hanushek reviewed 69 estimated coefficients that examined the effect of teacher salaries, and found that only 11 of them had a positive and statistically significant effect on achievement (Hanushek 1989).

In sum, the debate on the importance of school resources, or school expenditures per pupil, on educational outcomes is inconclusive. Hanushek (1989) argued that of 65 equations he examined, only 16 estimated coefficients demonstrated statistically significant relationship between expenditures per pupil and achievement. While this finding is difficult to interpret, due to variations in measurement, price levels, and family input measures, Hanushek argued that the finding points to "no strong or systematic relationship between school expenditures and student performance" (Hanushek 1989, 47). Yet Hedges and Greenwald (1996) found that per-pupil expenditure has a strong and consistent relationship with student academic achievement, and that positive relationship is large enough to be educationally important.

3. Teacher Qualities

Research findings on the relationship between teacher effects and student achievement are rather mixed. Teacher education and content knowledge, experience, and instruction are major features used as proxies for teacher and teaching quality.

Generally, researchers agree on the positive effect of teacher experience on achievement. In a review of studies conducted in the U.S., Hanushek (1989) found that of 140 estimated coefficients, 40 of them reported positive and statistically significant effects. Of 90 equations that found a statistically insignificant relationship – 44 estimated coefficients had positive effects. Hanushek concluded that "a clear majority of estimated coefficients point in the expected direction, and almost 30 percent of the estimated coefficients are statistically significant by conventional standards" (Hanushek 1989, 47). In a more recent study, Fetler (1999) found teaching experience to be associated with achievement gains in high school mathematics. Rowan, Correnti, and Miller (2002) found similar results for elementary mathematics. However, Hanushek pointed to possible selection effects associated with teacher experience and achievement. He argued that the positive relationship between experience and achievement may result from more experienced teachers teaching better (or high achieving) students (Hanushek 1989) and inexperienced teachers assigned to low-achieving and low-income students (Desimone and Long 2010).

In addition, teacher experience seems to be conditionally related with achievement. Researchers argued that teacher experience matters only up to a certain point. For instance, Murnane found that experience of one, two, or three years over having none mattered greatly on achievement, but the extra benefit between three and five years was less. He found no effect of experience on achievement beyond five years (Murnane 1975). Also, Summers and Wolfe (1977) pointed out that the effect of teacher experience varies among student's grade level. They found that students whose third-grade score was above grade level benefitted from more experience, but those who were very much below grade level were negatively affected. They concluded that younger students "did best with newer teachers who perhaps have an undampened enthusiasm for teaching those who find it hard to learn" (Summers and Wolfe 1997, 644).

The relationship between teacher education and achievement is less consistent. Summers and Wolfe (1977) found that teachers who received B.A.'s from higher rated colleges were associated with students whose learning rate was greater, and it was students from lower income families who benefitted the most. Greenwald, Hedges, and Laine (1996), in their meta-analysis of studies conducted in the U.S., found that that the teachers who attend better colleges and/or score higher on standardized tests produce greater gains in student achievement, but they are less likely to teach low-SES, Black, or Hispanic students. Hanushek (1989), however, did not conclude that teacher education is systematically related with student achievement. Of 113 estimated coefficients that examined the relationship in the U.S., 100 of them showed no statistical relationship between teacher education and student achievement, and of 13 estimated coefficients that had statistical significance at the conventional confidence level, only 8 had a positive effect.

Teacher's subject matter knowledge is also less consistently related with achievement (for review, Darling-Hammond 2000; Desimone and Long 2010). Most studies that examined teachers' scores on the subject matter tests of the National Teacher Examinations (NTE) and the teachers' performance, measuring by student outcomes, did not find a consistent relationship (Darling-Hammond 2000).

Some studies, however, have found positive teacher effects. Armor (2007), using the 1996 NAEP data, concluded that while teacher experience and teacher certification are insignificantly related with a student's math score, whether the teacher had a major/minor in math is statistically and positively related with a student's math score. The result was consistent even after major family background factors were considered. In their review of studies conducted in developing countries, Fuller and Clarke (1994) found that all 4 of 4 analyses found a significant effect of a teacher's knowledge on students' achievement.

Other proxy measures for teacher quality also show mixed results. For instance, some researchers have found the positive effect of teacher certification on student math achievement (e.g. Darling-Hammond 2000), while others have found insignificant effect on math (e.g. Armor 2007). Also, researchers have argued that while in-service teacher training is not independently associated with achievement, when it interacts with prior levels of teacher education or training, in-service teacher training can be effective (for review, Fuller and Clarke 1994).

Similarly, studies of teaching practices and instruction revealed complex findings. Previous literature has found a student-centered, didactic approach to learning to be more effective in raising achievement (e.g. Anderson 2004). But in a more recent cross-national study conducted using the 2006 PISA study, authors found that even in the 10 largest OECD countries, the student-centered pedagogy was negatively related with achievement (Sousa, Park, and Armor 2012). Studies conducted in developing countries also confirmed the finding (for review, Fuller and Clarke 1994, 132). Armor (2007) argued that there are very few large-scale, systematic studies that have examined various curricular and teaching techniques, and even those studies have revealed methodological problems, resulting in validity issues. In sum, there is no solid evidence to conclude which pedagogical approach is associated with higher achievement (Armor 2007).

In addition, as with other school effects variables, data and evidence used in the policy process are often corrupted with "bias […] embedded in the data or evidence itself, bias associated with analysis, and the biases of those in the policy world who use the information" (Earley 2000, 35). In general, the field of teacher education suffers from inadequate data to address many of the emerging research questions. For instance, the decentralized education system and the teacher preparation enterprise make it difficult to write an accurate description of highly variable teacher education, in terms of size and characteristics (Floden 2010). More systematic data are emerging (e.g. Teacher Education Study in Mathematics and US Department of Education's Institute of Education Science's study on early reading instruction), and the results are to be examined (Floden 2010).

While there seems to exist some complexities in revealing consistent positive teacher effects, researchers generally agree on the important role that teachers play. For instance, there exists a cumulative effect of teacher quality on student achievement. Studies have found that students who are assigned to several ineffective teachers in a row have significantly lower achievement (Darling-Hammond 2000). Also, researchers found a strong effect of classroom climate, including the student-teacher relationship, on achievement (for review, Anderson 2004).

4. Instructional Time

Studies conducted in the U.S. found a substantive effect of the time spent on instruction and student achievement (Carroll 1963; Brown and Saks 1986; Corey et al. 2012). Brown and Saks (1986) found that time has positive effects in both subjects, math and reading, in both grades, grade 2 and grade 5, but the effects vary significantly across teachers in mathematics instruction (Brown and Saks 1986, 498). Also, time of engagement on tasks is a stronger indicator than the amount of time spent on school subjects (Berliner 1990). Some researchers also noted that while there exists positive effects of staying longer in school, these effects exist with diminishing returns on educational outcomes (for review, Corey et al. 2012, 147).

5. Shadow education

The growing trend in the use of informal, privately-operated education services to supplement formal primary and secondary schooling is gaining popularity throughout the world. Every nation in the world has formal outside-of-school classes and private tutors that are used to help students navigate a successful passage through the formal education system and into adulthood (Baker and LeTendre 2005). These institutionalized forms of fee-based academic instructions and private tutoring offered to students to supplement formal primary and secondary schooling is so pervasive in many countries that the form of private tutoring is referred to by researchers as the "shadow education system" (Stevenson and Baker 1992; Bray 2006). The term "shadow" refers to an out-of-school system that "mimics the mainstream, growing as the mainstream grew and changing as the mainstream changed shape" (c.f. Bray 2010) and mostly at primary and secondary levels. While some scholars such as Bray emphasize the "privateness" in terms of financing when categorizing after-school instructions as "shadow education", most research focuses on "privateness" in terms of operation and management of the educational services. Likewise, outside-school learning includes tutoring, review sessions, proprietary cram schools, and other related practices, and all of these instructional services that are geared toward improving students' academic achievement outcomes outside of formal school hours are all referred to as "private tutoring" for the purpose of this study, and the term is used interchangeably with "shadow education".

The issue of shadow education has recently received attention of the United Nations Educational, Scientific, and Cultural Organization (UNESCO) and other educational researchers. With the belief that a market-driven education affects mass schooling at the national level in terms of both equity and quality of education, most debate was centered on how the growth of shadow education will exacerbate the issue of inequities. In fact, households with higher income and higher levels of parental education were found to utilize more private education, compared to households with lesser income and socioeconomic status (Bray 1999). Consequently, most research tended to focus on the adverse effects of shadow education and consequent policy responses to curb the demand for tutoring. Some nations, including South Korea, have tried to ban private tutoring although without much success.

Yet more emphasis should be imposed on the effect of shadow education on raising academic achievement of target students and the possibility of using shadow education as a policy tool. To date, the impact of shadow education has not yet been widely researched. Baker, one of the most actively engaged researchers of private tutoring throughout the world, criticizes the inadequate level and depth of systematic research on the issue (Baker et al. 2001). However, because of the informal nature of shadow education, researchers claim that it is not easy to collect data and make conclusions based on oftentimes incomplete survey data. Scarcity of research in the topic is particularly surprising in South Korea, where the majority of students' reliance on shadow education has become almost a norm. Chung (2002) argues that in Korea, public resentment to private tutoring is so strong that it requires "a gut for any educational policy makers to try an objective judgment on the merits and demerits of [private tutoring]" (Chung 2002, 1). Likewise, many stories reported in the news media emphasize the harms resulted from heavy reliance of private tutoring, but arguably these "demerits" have not been tested with scientific rigor as to conclude whether or not the hypothesis is true. Being a politically sensitive issue, data is not widely published to allow rigorous scientific study, and even when the data is available, most are not of the nationally representative sample (Dang 2007).

In both the U.S. and South Korea, the role of shadow education plays, or is growing to play, a significant role in their public education systems. Both are experiencing a growth of the shadow education industry. It is worth examining the effect of these informal educational instructions on academic achievement, since Korean parents and students' reliance on the informal educational sector is significant and rapidly growing. Data shows that 87.9 percent of Korean elementary students, 74.3 percent of middle school students, and 62.8 percent of general high school students were involved in private tutoring in 2009.[4] The Majority of Korean students rely on private tutoring to satisfy their educational needs, and Korean parents complain about the increasing burden of educational costs on their budget. According to recent data, families in Seoul spend about 16 percent of their income on shadow education.[5] Similarly, the private tutoring industry is arguably one of the most rapid industries in the United States. *The New York Times* and *Washington Post* have published articles that point out private tutoring industry as one industry that is immune to recession.[6] Under the NCLB, the government has required the Title I schools to provide after-school supplemental educational services to low-income, low-achieving students at free or charge. Baker et al., using the 1995 Third International Mathematics and Science Study (TIMSS) data, showed that more than 30 percent of 8th graders in the U.S. participate in shadow education. The same data indicates around 50 percent of Korean 8th graders participate in private tutoring.

[4] Statistics Korea, http://kostat.go.kr/portal/korea/index.action; referenced in Lee, Soojeong and Roger Shouse, (2011).
[5] Refer to *The Economist* article, "The one-shot society" (Dec 17, 2011).
[6] Sullivan, Paul. "As Private Tutoring Booms, Parents Look at the Returns" *The NY Times* (Aug, 2010); Sarah Maslin Nir. "Like a Monitor More Than a Tutor" *The NY Times* (Nov 2010).

The growing but still limited existing studies reveal rather mixed results on the impacts of shadow education. Kang (2007), using an instrumental variable (IV) that is strongly correlated with the expenditures on private tutoring but exogenous to the determinants of academic achievement, revealed negligible causal effect of private tutoring on student achievement. Kang and Ryu (2007), using a different method, also found similar results. Another study that examined the effect of pre-school tutoring on Korean students' grade-point-averages also found no statistically significant effect (J. T. Lee, Kim, and Yoon 2004).

Research on the effect of private tutoring on achievement is more difficult in the U.S. due to the decentralized nature of the American education system, there are a number of studies that examine the effect of after-school lessons on academic performance. Dynarski and others (2004) examined the effect of the 21st Century Community Learning Centers Program, one of the federally funded after-school programs that provided funding to 2,250 school districts to support school-based programs in 7,000 public schools, and found that while the program had little positive impact for both elementary and middle school students. In addition, evaluation studies of the Supplemental Educational Services under the NCLB Act of 2001 provided little insight on the effect of supplementary educational services as a handful of these studies found mixed results. While some researchers concluded the effects are statistically significant, most of the researchers found that the effect is not statistically or substantively significant on raising academic achievement of at-risk students and closing the achievement gap with the majority students.[7] Furthermore, while research suggests that one-on-one tutoring is the most effective form of remediation (Farkas and Durham 2007; Wasik and Slavin, 1993), critics are less optimistic about the efficiency of the program in closing the achievement gap (Farkas and Hall 2000).

These studies, however, pose reliability questions. Because of the decentralized system that allows local school districts and state authorities to establish academic standards, the system suffers from lack of consistencies in achievement measures and subsequently, comparison of students' academic achievement across states lack validity (Grissmer et al. 2000). With the publication of the National Assessment of Educational Progress (NAEP) test, testing achievement outcomes of general American students became possible; however, even the data is limited in examining the effect of private tutoring on academic achievement due to the lack of a direct measure of students' use of private tutoring. Furthermore, the findings from these program evaluation studies might not provide policy makers with credible results, as these studies have employed quasi-experimental studies with potential selection bias issues (Zimmer at al. 2007).

Recent publications of international assessment datasets, including the Trends in International Mathematics and Science Study (TIMSS) and the Program for International Student Assessment (PISA), allow researchers to engage in more complete comparative educational research. Using the 1995 TIMSS study, Baker and others (2001) conducted

[7] Zimmer, Ron et al. (2007) found, of the 9 large districts examined, statistically significant effect in 5 districts (2 districts were dropped from the sample due to small number of participants & 2 districts had no statistically significant effect), but the effect size was only about 0.09 with the confidence interval of [0.03-0.14]. Also, Chicago Public School Districts (2007) found students who participated in the SES program tended to score 0.8 adjusted scale scores higher than other eligible students who did not participate. But Munoz et al (2008) and Rickles and White (2006) did not find statistically significant effect of SES on achievement.

an extensive research to reveal whether there exists cross-national variations in the use of shadow education. Because the TIMSS survey does not directly ask why students participate in shadow education activities, Baker and others made inferences based on a series of bivariate and multivariate analyses of other information provided in the study (Baker et al. 2001, 6). Using the logistic regression of the use of shadow education (use vs. no use) as a function of student math scores, they made a series of conclusions. First, in some countries, the use of shadow education is positively related with the math scores indicating an enrichment strategy as the dominant role for shadow education, whereas in other countries, the relationship is clearly negative suggesting a remedial strategy.[8] South Korea, Romania and Thailand are among nation states that use shadow education as an enrichment strategy, and the U.S. and Japan are among nation states that use it as a remedial strategy. In addition, the authors further concluded that the presence of high-stakes tests are not related at all to cross-national variation in the use of shadow education and that the national use of shadow education is not associated with national achievement levels.

A handful of subsequent studies confirmed the conclusions suggested by Baker et al. (2001). Jae Kyung Lee (2007) used the same dataset to describe the cross-national variations in the prevalence and causes of private tutoring around the world, particularly between the U.S. and Korea. Lee hypothesized, based on Baker and others' conclusions

[8] Baker and others (2001) contended that enrichment strategy as the modal use indicates that students with high performance in mathematics generally use shadow education for strategic advantages in future educational contests; remedial strategy as the modal use, on the other hand, indicates that students with low performance in mathematics tend to use shadow education to maintain minimal or otherwise acceptable achievement levels in school.

on the cross-national variations in the modal use of shadow education, that private tutoring is more a function of enrichment needs in Korea but more of a function of remediation needs in the U.S. (Jaekyung Lee 2007, 1211). Using the 1995 TIMSS study and private tutoring in math as a dependent variable, Lee concluded that private tutoring in Korea is generally a function of college aspirations and academic enrichment needs, whereas in the U.S., it is more closely related to academic remediation needs. Similar conclusions were drawn by researchers who used different datasets. Using the 2003 PISA study, Southgate (2011) tested the hypothesis posed by Baker et al. (2001) that a national modal strategy affects the decision making process of whether or not a family will purchase shadow education. Southgate also used the use of shadow education as a dependent variable and the logistic regression analysis to replicate the 2001 Baker et al. study, and concluded that South Korea, the Czech Republic and Thailand use enrichment strategy, whereas the U.S. and Canada are among other nation states that use remedial strategy.

Summary

Ample efforts have been paid to reveal school effects, but educational researchers have failed to reach a consensus on these school effects on student achievement. The school effects literature is criticized on various fronts, including the adequacy of the measurement of family background factors used as control variables (see Buchmann 2002, 167). Fuller and Clarke (1994) argued that using proxy measures that are not compatible across nations and cultures can lead to misspecification of student background factors, which can lead to bias in the school effects.

Similarly, as the debate between Greenwald, Hedges, and Laine (1996) and Hanushek (1989; 1996) suggests, there exists limitations in data collection and research methodologies. Because of the feasibility issues related with social science subjects, it is costly to conduct randomized experiments, and studies that utilize quasi-experiments are challenged by the unobservable variables that cannot be easily controlled (e.g. student innate ability, motivation and other attitudinal variables). Also, as some other researchers have pointed out, due to the lack of data that examines students at various stages of school life (e.g. longitudinal data) at diverse educational settings (e.g. comparative international surveys), it is difficult to analyze and find the most effective school inputs on student achievement (e.g. Bryk and Raudenbush 1988).

Peer Group Effects

Previous literature has confirmed that school composition is closely related with student achievement. Student demographics, the aggregated socioeconomic status of the student body, and school climate are frequently considered contextual variables.

Student demographics, such as the proportion of ethnic, racial groups, the proportion of female students, and the proportion of immigrant students, and their effects on school and student academic performance were considered. Studies have found that demographic issues accounted for a large portion of between-school achievement variances (S. Lubienski, Lubienski, and Crane 2008). Summers and Wolfe (1977) found that black and non-black students benefitted when they were in schools with a 40 to 60 percent black student body, rather than in schools that were more racially segregated.

The aggregated SES of the student body has been shown to be independently associated with student outcomes. Willms (2010) argued that all school systems have some degree of segregation, and he used the term ‘horizontal segregation’ to notate the segregation associated with SES, and ‘vertical segregation’ to refer to the extent that students with differing levels of academic performance are segregated among schools. He found that school systems that are horizontally segregated tend to be also vertically segregated (Willms 2010). In other words, schools with a high mean SES tend to have higher student academic performance.

When students are segregated based on their academic performance, as are in the case of ability grouping and ‘vertical segregation’ of schools, the academic outcome is often exacerbated. Summers and Wolfe (1977) argued that being in a student body with more low achievers has a negative effect on learning for all students, and that ability grouping adversely affect low-achievers. Similarly, when lower SES students are grouped in a lower SES school, their lower educational outcomes can be exacerbated (Perry and McConney 2010). Perry and McConney (2010) found that increases in the mean SES of the school are consistently related with increases in student academic achievement.

Sense of community and parental involvement in school affairs are important determinants of school climate. In fact, Bryk, Lee, and Holland (1993) argued for the importance of sense of community within Catholic schools (for review, S. Lubienski et al. 2008). In addition, parent involvement and their expectations on school standards have been found to positively affect students and school performance (S. Lubienski, Lubienski, and Crane 2008).

Purpose of the Study

The existing school effects studies are inconclusive. Buchmann (2002) summarized some of the issues related with the school effect studies. One issue involves the compatibility of school-level measures across multiple contexts. While it is important to gauge measures specific to local contexts, the use of widely divergent measures or concepts arguably leads to results that are less comparable (Buchmann 2002, 168). Also on the methodological front, the reliance of school effects studies on OLS regression analysis of education is likely to create problems as the standard regression models are likely to overestimate the standard error of population parameters. This is particularly true, since most of the educational data have nested data structures. A new generation of "effective school" research revisited the school and family effects questions using the multi-level modeling, which takes into account the hierarchical structure of educational data (cf. Buchmann 2002, 168).

The 'new generation' of school-effects studies uses the multi-level modeling and generates findings that contrast to findings from earlier research. For instance, using the International Association for the Evaluation of Educational Achievement (IEA) data, Lockheed and Longford (1991) found stronger family-effects than school-effects across countries, regardless of the level of economic development. Such findings contradict earlier findings by Heyneman and Loxley (1982), which stated that in developing countries, the school effects are greater than the family effects.

Using the HLM method, Borman and Dowling (2010) and Konstantopoulos and Borman (2011) revisited Coleman's Equality of Educational Opportunity study and re-examined the school-effects. They found that school factors exert significant influence on student achievement. Even those studies that concluded stronger family-effects than school-effects on achievement outcomes, they also found that school-level factors explain a large amount of the explained variance in student achievement when international comparative data were used. For instance, Lockheed and Longford (1991) found that school-level differences attributed to 32 percent of the explained variance in student achievement in Thailand.

Asides from whether school factors matter on achievement outcomes, the use of the multi-level modeling technique allows researchers to properly examine the importance of each factor and/or cross-level interaction of variables across countries. A few studies have conducted comparative studies between two vastly different countries in terms of educational institutions and outcomes. For example, the United States and South Korea are often forced into a comparison in terms of students' achievement scores on internationally conducted standardized examinations because of the differences in their educational systems and their relative rankings on student assessments. Hyunjoon Park (2005) examined the institutional differences on the effect of family factors on student achievement. Using 2000 PISA data, he found that institutional features of educational systems influence the way in which family SES is associated with student achievement. Similarly, Shin, Lee, and Kim (2009), using the 2003 PISA data, examined the relative importance of student- and school-level factors that affect student achievement in Korea, Japan, and the United States, and found that there exist different patterns of the relations between student- and school- level predictors and student achievement. These findings suggest that the relative rankings of U.S. and Korean students are rendered meaningless unless researchers can compare and contrast the respective differences in individual and family resources and in the institutional and organizational characteristics of schools.

This dissertation seeks to supplement the prior literature by expanding the scope of analysis and using proper measures and research methodologies to examine the relationships between student- and school-level factors and achievement outcomes. While most domestic research has failed to reveal any strong correlations between school resources and practice factors, international comparisons might reveal some interesting relationships. By taking advantage of the PISA survey dataset that includes a wide range of information on students' family and school information, I attempted to include a more complete set of control variables. The study includes many school practice variables and policy measures, and consequently, it can provide interested policy makers and educational researchers with useful information on the effects (or associations) of such school practices/ policies on student achievement. Particularly as the demand for shadow education increases, it is in the interest of educational researchers and policy makers to reveal the effect of shadow education on achievement and to examine the effectiveness of using supplementary educational services as a policy tool.

CHAPTER THREE: CONCEPTUAL FRAMEWORK & HYPOTHESES

Conceptual Framework

An input-output model for the relationships between academic achievement and a variety of potential causal factors shows how achievement is influenced by different educational contexts. Student achievement, the output measure, was determined as a function of observable input variables, individual-, family-, and school-level factors (See Figure 1). Most individual- and family-level factors are outside the scope of policy influence. These variables, including student gender, family SES, and a family's social and cultural capital measured by home educational resources and parental involvement, directly influence student achievement outcomes. Some family-level factors also indirectly influence achievement via school factors as the family's social status or social class often determines the quality of school the children attend.

The country's structural and governance factors, as well as more specific policy measures, largely shape school-level characteristics. For instance, school regulatory policies shape the degree of autonomy each school has in determining its own curriculum. Also, financing mechanisms determined by the higher-level authority can influence a school's educational resources. The way schools are financed can determine their educational resources. Furthermore, teacher-related policies, including teacher certification and pre-service training, influence the quality of teachers at schools. These school-level factors exert influence on student achievement.

The student's participation in shadow education can be regarded as another proxy measure of a family's social/cultural capital. Previous literature established that parental involvement and family SES determine the use of shadow education (Byun and Park 2012). To some extent, student participation in shadow education can be influenced by policy measures. As was discussed in the previous chapters, governments have utilized shadow education as a policy tool. For instance, the U.S. government has publicly subsidized student participation in shadow education, and the Korean government has banned student participation in shadow education for over a decade. It is also important to examine within- and between-school achievement disparities after major achievement factors are considered.

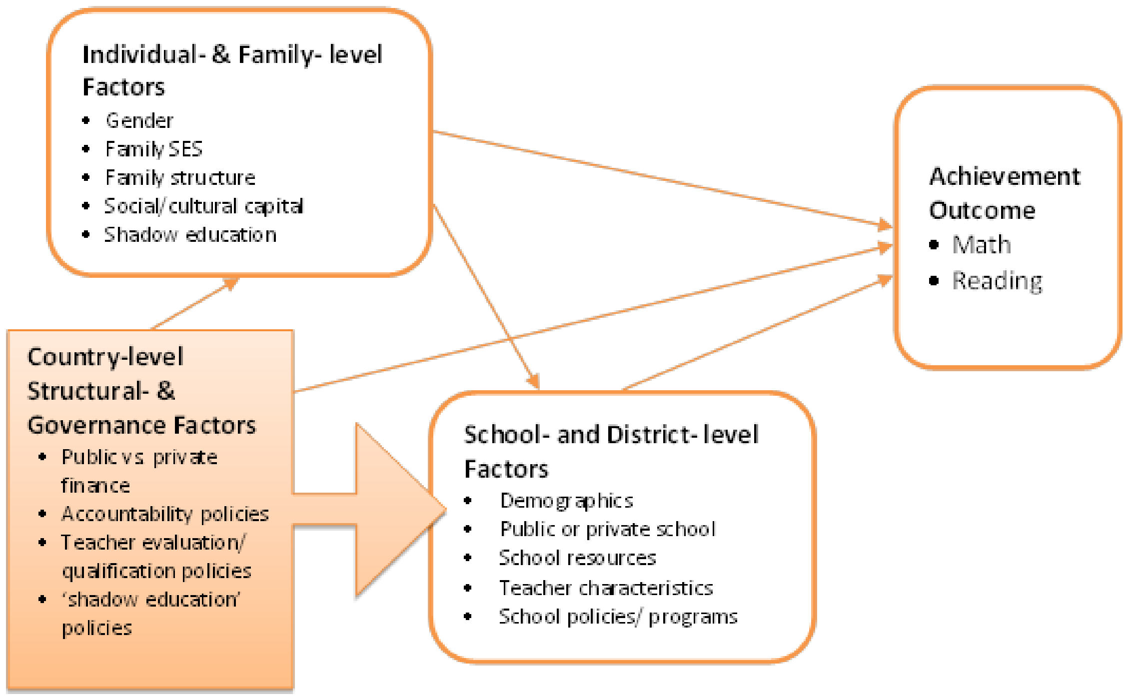

Figure 2 Conceptual Model

Research Questions/ Hypotheses

An increasing number of studies use the international assessment datasets to examine achievement disparities and correlates of those disparities among participating countries throughout the world. Despite this research, important questions remain to be answered. This study aims to answer three research questions:

1. What role do family SES and 'social/cultural capital' factors play in achievement outcomes for Korean and U.S. students? Is there a difference in the relative

importance of family background factors on within-country achievement outcomes between Korea and the United States?

2. What role do school characteristic differences play in achievement outcomes of Korean and U.S. students? What is the relative importance of the school-level factors in explaining within-country achievement outcomes of Korea and the United States?
3. What role do 'shadow education' and other government policies (e.g. accountability, class size, teacher performance evaluation) play in explaining differences between KOR-US achievement outcomes? Do shadow education and government policies contribute to or lessen the country differences?

Below is a list of hypotheses:

Firstly, this dissertation investigates to what extent the cross-national differences are attributed to family, socioeconomic, and cultural factors. As established in the literature review section, family background and socioeconomic status factors significantly influence student achievement across countries. I examine whether this common finding is consistent for both Korean and U.S. students. I hypothesize that family, socioeconomic, and cultural factors have statistically and a substantively significant relationship with Korean and U.S. student achievement, but the extent to which these factors influence student achievement may differ between the two countries.

> (H1) Although socioeconomic, family and cultural factors play an important role in explaining the within-country achievement outcomes of

> students in Korea and the U.S., the extent to which these factors affect achievement differ between these two countries.

Next, I will examine whether differences in school factors influence achievement outcomes in these two countries. Based on the existing literature that school-related factors attribute little influence to the achievement of students in developed countries with fully functioning educational systems, I hypothesize that despite significant differences in the organization and structure of the education systems in these two countries, these differences do not play an important role in achievement outcomes in the U.S. or S. Korea.

> (H2) After controlling for individual and family-level factors, the extent to which the school resource and practice factors explain within-country achievement differs between Korea and the United States.

Finally, this study examines how educational policies interact with the existing educational contexts and systems in each country. Without considering the overarching educational environment, previous studies have reported mixed findings on the effect of policy measures, including: the effect of outside-of-school lessons; accountability measures; and, the central curriculum. My hypothesis is that these policies interact differently with other school-level variables in South Korea and in the U.S., as these two countries have adopted vastly different educational systems.

> (H3) After controlling for other student- and school-level variables, the measures of government policies and shadow education show statistically and substantively significant relationship with student achievement. But

the relationship has different magnitudes and directions in Korea and the United States.

CHAPTER FOUR: DATA & METHODS

Data

This Study uses the Program for International Student Assessment (PISA), sponsored by the Organization for Economic Cooperation and Development (OECD). PISA is an international study that was initially conducted in the year 2000 and has been administered every three years since then. In 2000, a total of 43 countries/economies participated in the study; the participation increased to more than 70 countries/economies for the 2009 survey.[9] Because PISA includes samples from countries/economies of various developmental stages, PISA provides comprehensive information on education systems worldwide. Availability of such large-scale international studies allows researchers to engage in more rigorous hypothesis testing and benchmarking, but researchers need to be well aware of the importance of adjusting for between-country differences and background conditions, such as cultural factors (Porter and Gamoran 2002).

The purpose of PISA is to evaluate "to what extent students at the end of compulsory education, can apply their knowledge to real-life situations and be equipped for full participation in society" (OECD 2013), and thus, PISA does not restrict its testing to curriculum-based materials. For that purpose, PISA targets 15-year-old students in

[9]Refer to OECD Programme for International Student Assessment homepage for the list of participating countries/economies.

participating countries/economies. The age-based sampling, while not perfect in examining educational effects (related to curriculum-based effects), allows researchers to examine “the cumulative yield of education systems at an age when schooling is still largely universal” (OECD 2010, 10). In addition, the 15-year-old students are often in the final level of compulsory education. In most countries, including the more developed ones, education is universally provided and oftentimes compulsory through middle school, but upper secondary levels are not compulsory. For example, In South Korea, middle school education became compulsory in 2005 but upper secondary level education is neither universal nor compulsory. Therefore, assessing the knowledge and skills of students nearing the compulsory level of education has meaningful implications for educational researchers and policy makers.

Using the 15-year-old students as the target population corresponds with the purpose of the PISA study; PISA intends to measure the knowledge and skills of a group of individuals who were born within a comparable reference period, but who may have undergone different educational experiences in and outside schools (OECD 2010). In each testing year, PISA focuses on one specific domain of literacy while testing students on all three domains. For example, the focus was in reading literacy for the year 2000 testing, and the focus shifted to math in 2003 and to science in 2006. As in 2000, reading literacy was the focus of the 2009 PISA study. The reading framework was updated and the 2009 study also included the assessment of electronic texts (OECD 2010). The 2009 study also includes measures of achievement in mathematics and science; but less information is collected about specific school instructional programs in math and science.

Rather than focusing on specific knowledge acquisition, PISA assesses students' acquisition of broader concepts and skills that are essential in adult life. In reading, "the capacity to develop interpretations of written material and to reflect on the concept and qualities of text" are considered as central skills (OECD 2010, 12). In mathematics, "being able to reason quantitatively and to represent relationships or dependencies" rather than the ability to answer familiar textbook questions is the central framework (OECD 2010, 12). Similarly in science, "having specific knowledge, such as the names of plants and animals, is of less value than understanding broad topics such as energy consumption, biodiversity and human health in thinking about the issues under debate in the adult community" (OECD 2010, 12). The 2009 PISA study includes the frameworks for assessing mathematics and science that were fully developed in the 2003 and 2006 testing (OECD 2010).

PISA mainly uses paper-and-pencil tests, and in total, 390 minutes of test items are covered (OECD 2010). Rather than having to test each sampled student with the whole item battery, PISA uses a rotated booklet design and assigns a subset of the item pool to each student (OECD 2009a, 80). This design allows for "overcoming the conflicting demands of limited student-level testing time and the broad coverage of the assessment domain" (OECD 2009a, 91). Student assessment scores are then calculated using the Rasch Model. The Rasch Model, one of the Item Response Theory models, was designed to build a systematic continuum on which both item difficulty and student ability are located. Because the item difficulty and the student ability are linked by a logistic function, it is possible to compute the probability that a student succeeds on an item. The Rasch Model is then used to create a scale on which every item and every student will be located (OECD 2009a, 94). PISA, based on this design, created assessments that last a total of two hours for each student. In a range of countries and economies, an additional 40 minutes are devoted to the assessment of reading and understanding electronic texts (OECD 2010).

In addition to the paper-and-pencil tests, students answer a background questionnaire, which takes about 30 minutes to complete. In these questionnaires, they are asked various questions regarding themselves and their homes and schools. Principals from participating schools also answer a 20-minute questionnaire about their schools and community. In some countries/ economies, optional short questionnaires are given to parents. The United States does not participate in parent questionnaire, so researchers who focus on the U.S. use the student questionnaire to obtain information about students' home and family backgrounds. PISA allows researchers to conduct rigorous analysis on various family- and school-effects on academic achievement. Although causality cannot be drawn based on the cross-sectional data, the wide range of information PISA provides on a student's family, school, and community allows researchers to take into account most of the critical factors that influence students' academic achievement. Based on the findings, it is possible to offer plausible causal inferences (Porter and Gamoran 2002).

Sampling

PISA conducts two-stage, stratified sampling. Because two-stage sampling is cost-effective and practical, it is widely utilized in education surveys. It is particularly beneficial for researchers who are interested in understanding the variability in student performance in a particular educational context, such as the classroom and the school (OECD 2009, 51; see A. Bryk and Raudenbush 1992).

Yet compared to a simple random sampling, the two-stage, stratified sampling design increases uncertainty associated with any population estimates (OECD 2009a). For example, in the two-stage sampling, selected students attending the same school cannot be considered as independent observations, as is in the simple random sampling. Students attending the same school share common characteristics, such as school resources and teacher characteristics. Likewise, a simple random sample of 4,000 students is more likely to cover the diversity of population better than a sample of 100 schools with 40 students observed within each school (OECD 2009a, 63). This is particularly true for countries where between-school differences are large. For countries such as Austria, Belgium, Germany, and Hungary, where more than 50 percent of student performance differences are accounted for at the school-level, uncertainties associated with any population parameters are larger than the northern European countries where between-school variances are smaller.

Likewise, computing the sampling variance based on the sampling variance formula that assumes simple random sampling is inappropriate, as the standard error of the population estimate is larger for a two-stage sampling than for a simple random sampling. Particularly for PISA, which samples the primary sampling units (i.e., schools) proportionally to their sizes and adds stratification variables in the sample design, appropriate methods should be utilized for estimating sampling variances. In sum, any statistical approach to calculate the sampling variance that does not take into account the two-stage, stratified sample design, is likely to substantially underestimate standard errors.

One method for estimating the sampling variances from complex sample designs is using resampling or replication methods (OECD 2009a, 67). The replication approach "consists of estimating the variance of a population parameter of interest by using a large number of somewhat different subsamples (or somewhat different sampling weights) to calculate the parameter of interest. The variability among the resulting estimates is used to estimate the true sampling error of the initial or full-sample estimate" (OECD 2009a, 67). There are three main types of replication methods for two-stage samples: (1) the Jackknife, with two variants, one for unstratified samples and another one for stratified samples; (2) the Balanced Repeated Replication (BRR) and its variant, Fay's modification; and (3) the Bootstrap. PISA uses the Balanced Repeated Replication (BRR) with Fay's modification.[10]

Similar to the Jackknife method for stratified two-stage sample designs, the BRR method consists of systematically pairing sampled schools within each stratum and generating replicates. The BRR method selects one school within each pseudo-stratum at random to have its weight set to 0; the weights of the remaining schools in the pseudo-stratum are then doubled. The Fay's variant to the BRR method suggests that instead of multiplying the school weights by a factor of 0 or 2, the weights are multiplied by a deflating factor *k.* The value of the first deflating factor, *k*, is between 0 and 1 and the second inflating factor is equal to 2 minus *k*. The statistic of interest is computed based

[10] Refer to PISA Data Analysis Manual (2009), p. 72-74, for further discussion on the method.

on the whole sample and then again based on each replicate sample. The replicate estimates are then computed to the whole sample estimate to obtain the sampling variance.[11]

As is suggested by PISA, the student-level final weight and 80 replicate weights[12] are used to compute the standard error of population estimate, including mean, standard deviation, correlation coefficients, and multi-level regression coefficients. The general formula for computing the standard error for a given statistic θ with this method is given by the following formula (OECD 2009, 74):

Equation 2 Computing the sampling variance with the Fay's variant of the Balanced Repeated Replication (BRR)

$$\sigma^2_{(\hat{\theta})} = \frac{1}{(G(1-K))^2}\sum_{i=1}^{G}(\hat{\theta}_i - \hat{\theta})^2$$

Where G is the replicate samples, in this case, 80 replicates, and k is the deflating weight factor. For PISA, the deflating weight factor of 0.5, denoted as *k*, is used for computing the standard error. Hence, the equation:

$$\sigma^2_{(\hat{\theta})} = \frac{1}{(80(1-0.5))^2}\sum_{i=1}^{80}(\hat{\theta}_i - \hat{\theta})^2$$

[11] Refer to PISA Data Analysis Manual (2009), p. 72-75 for the formulas for calculating sampling variance for the Jackknife (for both stratified and unstratified samples) and the BRR method.
[12] The final student weight, W_FSTUWT, and the replicate weights, W_FSTR1-W_FSTR80, are used to estimate standard errors. I used SAS® macros to conduct the analyses.

Because PISA provides plausible values for outcome measures, computing sampling variance becomes more complex. Population statistics are estimated using each plausible value separately and then averaging each plausible value statistic.[13] Mathematically, if θ is the population assessment mean score and θ_j is the mean score of the sample computed on one plausible value, then:

Equation 3 Secondary Analyses with Plausible Values

$$\theta = \frac{1}{M}\sum_{i=1}^{M}\theta_j$$

Where M is the number of plausible values.

Plausible values are also used to calculate measurement variance, denoted as the imputation variance. This measurement variance involves the uncertainty in the estimate of θ due to the lack of precision in the measurement test (OECD 2009a, 100). Mathematically, the imputation variance can be described as follows:

Equation 4 Imputation Variance

[13] Averaging the plausible values at the student-level, i.e. by computing in the dataset the mean of the five plausible values at the student level and then computing the statistic of interest once using the average PV value can result in biased estimate of standard errors of population parameters (OECD 2009, p. 100).

$$B_M = \frac{1}{M-1}\sum_{i=1}^{M}(\theta_j - \theta)^2$$

Finally, total variance is computed by combining the sampling variance and the imputation variance.

Equation 5 Total Sampling Variance

$$V=\left[\frac{1}{(G(1-K))^2}\sum_{i=1}^{G}(\bar{\hat{\theta}}_i - \hat{\theta})^2\right]+\left[(1+\frac{1}{M})B_M\right]$$

Standard error is then derived by taking the square root of the variance.[14]

In each country, 35 students from each school are asked to participate and if fewer than 35 students attend a school, the entire school population is included (OECD 2009a, 51). The resulting sample included in the 2009 PISA study originally consisted of 475,460 students from 17,145 schools from 65 participating countries. The major focus of this dissertation lies on the cross-national comparison between the United States and South Korea. The pooled OECD data is used as a benchmark.

Missing Data

In the United States, 165 schools were selected to participate and 5,145 students were randomly selected from these schools. In South Korea, 4,930 students were selected from the 157 schools. France and Austria were removed from the OECD sample due to

[14] Population estimates and their standard errors are calculated using PISA SAS® macros.

missing school-level data. The resulting sample size of OECD-33 is 287,566 students from 1,535 schools.

Most variables included in the analysis have a range of missing data. Most of the variables have less than five percent of missing values, and only a couple of variables, e.g. the number of outside-of-school lesson hours, have missing values for more than 10 percent of the cases.

While complete case analysis is a widely used method, it might produce biased estimates if individuals with complete data differ systematically from the target population (Schafer 1999). Also reduction in the sample size due to missing data can be a problem. It is recommended that an imputation method be used in some cases to account for the missing data. Imputation method generally replaces missing values with some other values, e.g. the sample mean. While the single imputation method is likely to produce biased estimates of coefficients, only three variables in the model have more than five percent of the data missing and therefore, the bias is considered negligible (OECD 2012, 317).

In this dissertation, single imputation method is used to account for the missing data. For individual-level variables, country means were used to replace the missing values, and for the school-level variables, school means were used to replace the missing data.

Measures

Outcome Variable

Rather than using one simple score to represent a student's skills and ability, PISA uses plausible values for reporting student academic performance (OECD 2009a, 36). PISA notes that the methodology of plausible values consists of (OECD 2009a, 95):

(1) Mathematically computing distributions (denoted as posterior distributions) around the reported values; and

(2) Assigning to each observation a set of random values drawn from the posterior distributions.

In other words, plausible values can be "defined as random values from the posterior distributions" (OECD 2009a, 95).PISA reports five plausible values for each subject tested for each student.

Using plausible values to report student performance helps to reduce error in making inferences about the target population. Particularly in education, measures encompass substantial measurement errors because: (1) the concept to be measured is broader; (2) measures can be affected by the students' mental and/or physical conditions on the day of the assessment; and, (3) they can be affected by conditions in which students are tested (OECD 2009a, 96).

Student-level Variables

- *Student Gender* includes a dichotomous variable that indicates student gender. (MALE=1; female=0)
- *Family Background Factors*: include the PISA index of economic, social, and cultural status (ESCS). The index is created based on the parental occupational

status (HISEI), parental education (PARED) and home possession (HOMEPOS), which is based on proxy measures of wealth (WEALTH), home educational resources (HEDRES), and cultural possessions (CULTPOSS). The variable WEALTH is derived based on household possessions such as ‘a room of your own’, ‘a dishwasher’, ‘cellular phones’, ‘televisions’, ‘computers’, and ‘cars’, and the variable HEDRES is derived based on items such as ‘a quiet place to study at home’, ‘a desk to study at’, ‘educational software’, ‘books to help with your school work’, ‘technical reference books’ and ‘a dictionary’. The variable CULTPOSS is derived based on home possessions such as ‘classical literature’, ‘books of poetry’, and ‘works of art’. The ESCS index is mainly used in the analysis to control for family’s SES. The ‘ESCS’ variable was scaled to have a mean of 0 and a standard deviation of 1 at the student level for OECD countries. A dichotomous variable that indicates whether the student lives in a two-parent household (PARENTS=1; parents=0 (not two-parent household) is included in the model.

- *Students’ Approaches to Learning* factors: include two scale variables: Use of memorization strategies (MEMOR) and use of control strategies (CSTRAT). The variable MEMOR is derived based on questions such as “when I study, I try to memorize everything that is covered in the text” and “when I study, I read the text over and over again”; and the variable CSTRAT is derived based on questions such as “when I study, I start by figuring out what exactly I need to learn” and “when I study and I don’t understand something, I look for additional information to clarify this" (OECD 2012b). Students who use "control" strategies attempt to expand their understanding of the material rather than simply memorizeing it. These variables are standardized with an OECD average of 0 and an OECD standard deviation of 1. The higher number indicates that students are more likely to use the learning approach than their OECD counterparts with a lower scale score.

- *Instructional Time*: the number of hours per week of in-school math (MATHHRS) and language lessons (LANGHRS) are included in the model.
- *Shadow Education*: The 2009 PISA survey asks students the number of hours per week they attend outside-of-school lessons, denoted 'shadow education'. In the question, the word "typically" is used to account for seasonal variability in students' participation oin shadow education. For instance, some students might attend shadow education only during a long vacation or they might attend shadow education during the academic semester to receive supplementary help. PISA does not ask separate questions on the provider, e.g. school or private provider. The number of out-of-school hours spent a week in math (SHADOWMATH) and the number of out-of-school hours spent a week in test language (SHADOWLANG) are included in the model.

<u>School-level Variables</u>

- *School demographics and characteristics Factors*: school size (SCHSIZE) and school location [URBANSCH=1 (more than 100,000 population); urbansch=0 (less than 100,000 people)].
- *External Influence Factors:* includes a dichotomous variable measuring school sector (PUBLIC=1; private=0). Percent of government funding (GOVFUND) is included in the model. A dichotomous variable that measures whether there is some parental pressure on academic standards (P_PRESS=1; no-p_press=0) is included in the model.
- *Government Influence on School Management (proxy measure for centralized and decentralized controls):* PISA indices that measure whether school has greater responsibility towards resource allocation (RESPRES) and curriculum management (RESPCURR) are included in the model to measure the degree of school independence. The RESPRES variable is derived based on six items measuring the school principal's report on who has considerable responsibility for tasks regarding management of resource allocation, which includes "selecting teachers for hire", "firing teachers", establishing teachers' starting salaries", "determining teachers' salaries increases", "formulating the school budget", "deciding on budget allocations within the school". The RESPCURR variable is derived based on four items measuring the school principal's report concerning the responsibility for curriculum and assessment, including "establishing student assessment policies", "choosing which textbooks are used",

"determining curriculum content", and "deciding which courses are offered"(OECD 2012b).

- *School resources factors*: The ratio of computers available for 15-yr olds compared to the school size (IRATCOMP), teacher shortage (TCSHORT), pupil-teacher ratio (STRATIO), quality of school resources (SCMATEDU), and availability of extra-curricular activities (EXCURACT) are included in the model as proxy measures of school resources.
- *Class size*: The number of students in language class is used to measure class size (CSIZE).
- *Teacher-related factors*: include percent of certified teachers (PCCERT) at school and the percent of qualified teachers (PCQUAL) at school. A dichotomous variable of whether student achievement data are used to evaluate teacher performance (TCHEVAL=1; tcheval=0) is also included in the model.
- *Accountability Measures*: Dichotomous variables are included as proxy measures for accountability. Whether student achievement data are posted publicly (SCHEVAL=1; scheval=0) and whether the student achievement data are used to evaluate principal performance (PRINCIPEVAL=1; principeval=0) are included in the model.

Methods

To examine the relationships among student background and family SES characteristics, school factors, and student assessment scores (math and reading), this research employed two-level hierarchical linear modeling (HLM). The modeling is useful because PISA is a hierarchical data that randomly selects students nested within schools from participating countries. Multi-level modeling, by acknowledging potential interaction between student-level factors and school-level factors, takes into account the relative variation in the outcome measures, between students within the same school and between schools (OECD 2009b).

For instance, the socioeconomic background of students may impact how students are assigned to schools in some countries. As seen from Figure 3, in Country 1 students' attendance in a particular school does not depend on their socio-economic background, and there is no social segregation. However, in Country 4, schools do not cover the range of socio-economic backgrounds that exist at the population level: Students from more advantaged socio-economic backgrounds attend School 1, whereas students from more disadvantaged socio-economic backgrounds generally attend School 4 (OECD 2009b, 206). Multi-level sampling, which retains the nested structure of the data, allows researchers to examine potential interaction between student-level factors, e.g. socio-economic background, and school-level factors, e.g. parental pressure on school academic standards.

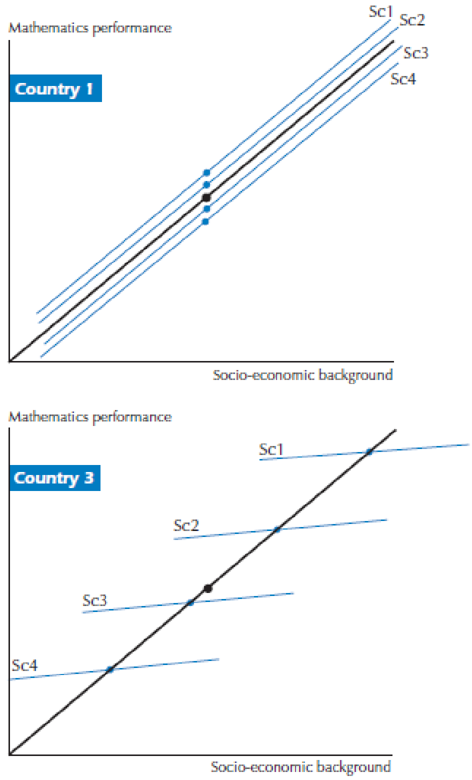

Figure 3 Simple linear regression analysis vs. multiple regression analysis
Source: OECD (2009), "Multilevel Analyses", in PISA Data Analysis Manual: SAS *Second Edition,* OECD Publishing

Because the nested structure of data violates the assumption of independent error terms, the HLM is useful in analyzing the multi-level data by taking into account the correlated errors (Bryk and Raudenbush 1992). The study uses the two-level HLM: level 1 consists of student-level measures and level 2 consists of school-level measures.

The multilevel modeling is used particularly to examine variations that occur at the school-level. For HLM, the school effects will be treated as a random effect so that variance within schools can be estimated (the random intercept model).

Level 1: Individual-Level Model

I first use a within-group (i.e. within-school) model that specifies the relationships among critical individual-level factors and outcome variables. In this study, achievement outcomes are estimated using PISA test scores as proxy variables.

Equation 6 HLM Student-level Model

$$Y_{ij} = \beta_{0j} + \beta_{1j}\,(\text{student gender}) + \beta_{2j}\,(\text{family's economic, social, cultural status})$$
$$+ \beta_{3j}\,(\text{two-parents}) + \beta_{4j}(\text{learning approaches}) + \beta_{5j}(\text{in-school lesson hours})$$
$$+ \beta_{6j}(\text{shadow education lesson hours}) + \varepsilon_{ij}$$

Where Y_{ij} is the assessment score (either math or reading) of student *i* in school *j*. β's measure the pooled within-school measures of student-level variables. Student gender is included at the student-level model. Family socioeconomic status is measured by parental education, parental occupational status, family wealth, home educational resources, e.g. number of books, and home cultural possessions. Whether the student lives in a two-parent household (including step-parents and guardians) is also considered. In order to control for study habits (and learning attitudes to a certain extent), I also included variables that indicate students' approaches to learning. The number of hours/week of <in-school> lessons is included at the student-level. Finally, the number of hours/week of <outside-of-school> lesson participation in math and language, denoted shadow education participation, are included in student-level model. While some schools in the U.S. provide the supplementary educational services (SES), which is better known as an after-school tutoring program, because the shadow education variables

(SHADOWMATH and SHADOWLANG) have both within- and between-school variations, it is included as a student-level variable. ε_{ij} is random error assumed to be distributed normally, with a mean of 0 and a variance of θ. In this model, the within-school regression coefficients are allowed to vary across schools (Raudenbush and Bryk 1986).

Level 2: School-Level Model

At the second-level, a between-group (i.e. between school) model is used to measure the differences in policies and practices among schools. In the multilevel modeling, the coefficients derived from the individual-level equation are used as dependent variables at the school level.

Equation 7 HLM School-level Models

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\,(\text{school demographics}) + \gamma_{02}\,(\text{school type and external influence}) + \gamma_{03}\,(\text{school resources proxy variables}) + \gamma_{04}\,(\text{teacher-related factors}) + \gamma_{05}\,(\text{policy measures}) + \upsilon_{0j}$$

$$\beta_{1j} = \gamma_{10},\ \beta_{2j} = \gamma_{20},\ \beta_{3j} = \gamma_{30},\ \beta_{4j} = \gamma_{40},\ \beta_{5j} = \gamma_{50}$$

Where β_{0j} is the adjusted mean outcome scores of school j and the other β's are regression coefficients of student-level predictors. At the school-level, measures of school demographics, school type and other external influence factors, school resources, and teacher-related factors are included. υ_j is random effects for specific schools and it is assumed to be normally distributed with a mean of 0 and a variance of ψ.

In the random intercept model, β_{pj} is treated as fixed effects and β_{0j} is considered random and allowed to vary across schools. This model predicts the student performance by taking the average performance of the student's school, and the school performance is predicted by the grand mean, β_{0j} (OECD 2009b, 207). This model has two random components: (1) the variance of ε_{ij}, denoted θ, and (2) the variance of υ_j, denoted ψ (OECD 2009b, 213).

Intra-class correlation (ρ), which measures how schools differ in their student average performance, is reported by computing the percentage of the total variance that is accounted for by the school (OECD 2009b). Mathematically, the between-school variance, ψ, (or the variance of υ_j) is divided by the total variance (or sum of the between-school variance (ψ) and the within-school variance (θ)).

Equation 8 Intraclass correlation

$$\rho = \frac{\sigma^2{}_{between-school}}{\sigma^2{}_{between-school} + \sigma^2{}_{within-school}} = \frac{\psi_0{}^2}{\psi_0{}^2 + \sigma^2}$$

According to PISA provided intraclass correlation statistics for OECD countries in 2006, Korea had a higher intra-class correlation; (0.35) than the U.S., (0.23). Adding student-level variables can have an impact on between-school variance, if schools differ in the mean and range of students with regard to the student-level variables, and/or the within-school regression coefficient of the student-level variable differs from zero (OECD 2009b, 215).

Coefficient of determination (R^2) computes the proportional reduction in prediction error variance comparing the model without covariates (the null model) with the model of interest (Rabe-Hesketh and Skrondal 2012, 134–135). The formula for HLM model is as follows:

Equation 9 Coefficient of Determination (R2)

$$R^2 = \frac{\hat{\psi}_0 + \hat{\theta}_0 - (\hat{\psi}_1 + \hat{\theta}_1)}{\hat{\psi}_0 + \hat{\theta}_0}$$

Where $\hat{\psi}_0$ and $\hat{\theta}_0$ are the estimates for the null model, and $\hat{\psi}_1$ and $\hat{\theta}_1$ are the estimates for the model of interest (Rabe-Hesketh and Skrondal 2012).

CHAPTER FIVE: RESULTS

Descriptive Statistics

Table 1 presents the descriptive statistics for Korea and the United States. This table shows the number of observations before conducting the imputation procedure to address the missing values, and the mean/proportion and standard deviation of variables included in the model are presented for Korea and the United States. Full descriptive statistics, including number of observations, minimum, and maximum can be found in the Appendix B-1 and B-2. Variable names and their descriptions are presented in the Appendix A.

This dissertation focuses on students' mathematics and reading assessment scores. In 2009, the mean math score for the 33 OECD countries in the sample was 488 with a standard deviation of 97. Descriptive statistics for OECD-33 is presented in the Appendix B3. For South Korea, the mean math score was 546 with standard deviation of 89, and for the U.S., the mean math score was 487 with a standard deviation of 91. The mean reading score for the 33 OECD countries in the sample was 492 with a standard deviation of 98 points. As for reading scores, the mean reading for Korea was 539 and the standard deviation of 79 points, and the mean reading score for the U.S. was 500 with a standard deviation of 97 points.

The difference in mean scores between Korean and U.S. students is smaller for reading, 39 points, than for math, 59 points, and both differences are statistically significant. In math, U.S. students scored 1 point higher than the OECD-33 average, and in reading, U.S. students 8 points higher than the OECD-33 average.

According to Raudenbush and Kim, it is important that researchers examine variations in achievement scores along with the national means. They noted that interpreting national means is difficult without knowing how much of the variation in the outcome lies within countries (Raudenbush and Kim 2002, 270). Because the variability of student achievement within a nation is another important indicator of overall inequality of educational outcomes, the dispersion of achievement scores can provide a better picture of how countries compare in their student performance.

Distributions of scores clearly show that students at the top of the U.S. achievement distribution achieve similarly to students at the top of the Korean distribution in reading. In fact, the U.S. students at the top 99^{th} percentile scored 704 points on average and this is slightly higher than the average score of Korea students at the same percentile, which is 699 points. The between-country achievement gap for U.S. and Korean students is larger at lower distributions than at higher distributions for reading (See Figure 4). In math, however, the U.S. students fall well below the Korean students at any distribution (See Figure 5).

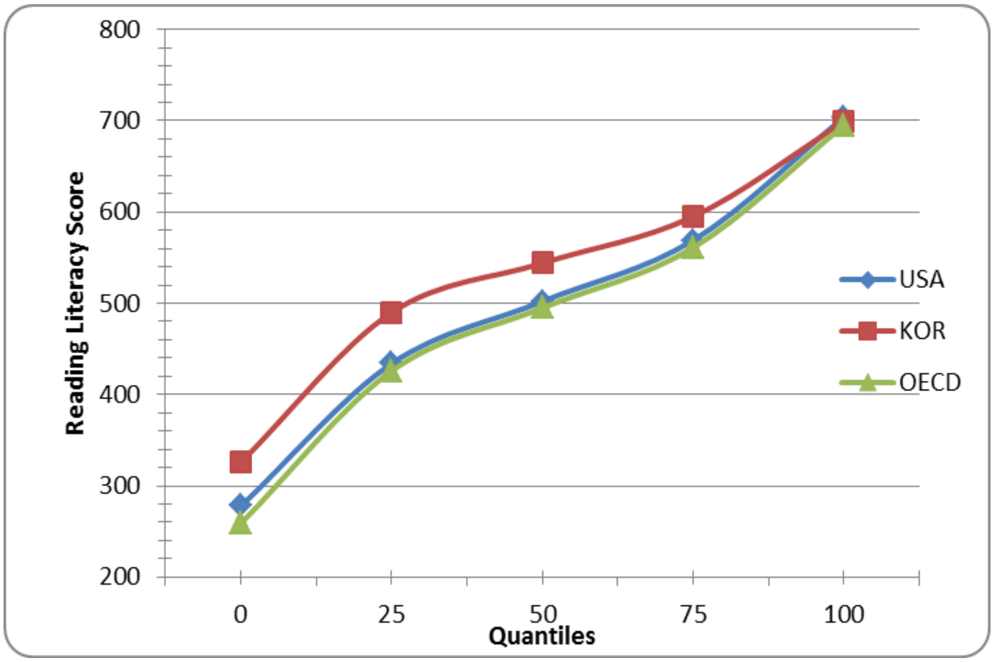

Figure 4 Reading score distribution for U.S., Korea, and OECD-33

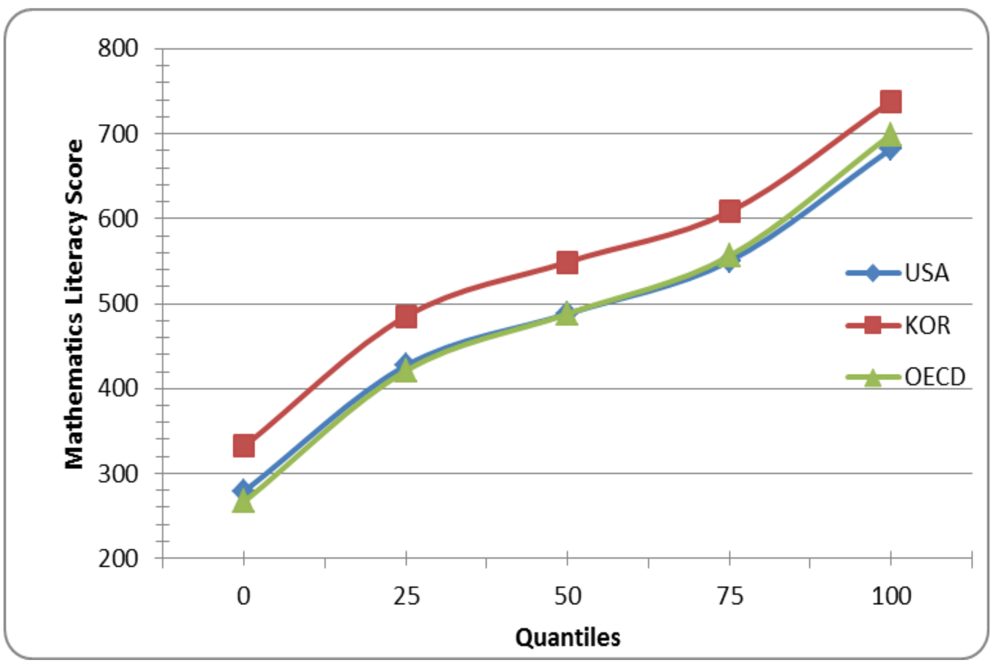

Figure 5 Math score distribution for U.S., Korea, and OECD-33

In addition, distributions of scores clearly show that the within-country achievement disparity between high achievers and low achievers is different for Korea and the United States. Particularly for reading, the within-country disparity is smaller for Korean students than for U.S. students. For instance, the difference in average score of students at the 75th percentile and the 25th percentile is 135 points for U.S. students, but the difference is only 105 points for Korean students. The within-country disparity for math score is similar for U.S. students and Korean students.

In sum, Korean students scored higher on math assessment than U.S. students at all distribution levels. Yet U.S. students at the highest distribution level, 99th percentile, scored higher than Korean students in reading assessment. But the achievement gap between high achievers and low achievers in reading is large for U.S. students.

Table 1 Descriptive Statistics for KOR and USA

	KOR		USA	
VAR NAME	**Mean/ Proportion**	**Std. Dev.**	**Mean/ Proportion**	**Std. Dev.**
Outcome Variables				
MATH	546.2	89.2	487.4	90.8
READING	539.3	79.2	499.8	96.6
Student-level Variables				
MALE	0.52	0.50	0.51	0.50
ESCS (within)	-0.15	0.82	0.17	0.93
HISEI	50.09	13.36	52.03	16.11
PARED	13.73	2.36	13.65	2.55

HEDRES	-0.15	0.94	-0.26	0.86
CULTPOSS	0.52	0.95	-0.30	0.96
PARENTS	0.85	0.36	0.72	0.45
Approaches to Learning				
MEMOR	0.08	0.95	-0.04	1.10
CSTRAT	-0.27	1.00	-0.04	1.12
Instructional Time				
MATHHRS	3.38	0.87	4.38	1.73
LANGHRS	3.28	0.82	4.36	1.74
SHADOWMATH	3.07	2.58	0.58	1.51
SHADOWLANG	2.15	2.13	0.42	1.34
School-level Variables				
School Demographics/ External Influence Variables				
URBANSCH	0.86	0.35	0.39	0.49
SCHSIZE **(SSIZE=SCHSIZE/100)**	1156.87	438.81	1369.13	848.70
PUBLIC	0.83	0.38	0.93	0.25
GOVFUND	47.62	24.54	90.40	26.64
STUDFEE	47.47	24.87	8.81	24.42
RESPRES	-0.44	0.75	0.40	0.92
RESPCURR	0.79	0.78	-0.20	0.93
P_PRESS	0.12	0.33	0.34	0.47
Teacher-related Variables				
PCCERT	97.84	10.77	95.16	14.19
PCQUAL	98.23	8.59	96.31	16.03
TCHEVAL	0.77	0.42	0.81	0.39
Accountability Measures				

SCHEVAL	0.33	0.47	0.89	0.31
PRINCIPEVAL	0.28	0.45	0.62	0.48
School Resources Variables				
IRATCOMP	0.43	0.43	0.73	0.48
TCSHORT	-0.02	0.92	-0.45	0.81
STRATIO	17.21	3.55	16.05	5.27
SCMATEDU	0.06	0.81	0.51	1.04
EXCURACT	1.01	0.82	1.02	0.78
CSIZE	35.89	5.11	24.48	6.78

The mean value for family ESCS (economic, social, and cultural status) is -0.15 for Korea and 0.17 for the United States. And yet, mean values for HEDRES (home educational resources) and CULTPOSS (cultural possessions) are higher for Korea than for the United States. The value of cultural possession is particularly high for Korean households. The mean value of CULTPOSS for Korean households is 0.52, whereas it is -0.30 for the U.S. households. On average, about 85 percent of Korean students answered that they live in a two-parent household and about 72 percent of U.S. students answered likewise.

On average, Korean students (m=0.08) are more likely to use the memorization strategies (MEMOR) than U.S. students (m= -0.04). While Korean (m=-0.27) and U.S. students (m=-0.04) are less likely to use control strategies than their OECD counterparts (m=-0.03), compared to U.S. students, Korean students are less likely to use the control strategies. The approaches to learning variables are derived based on students' approaches to reading and writing tasks, but I assume that students use similar approaches to learning subjects such as math and science as well.

On average, Korean students spend about 3.38 hours per week on math and about 3.28 hours per week on language lessons. U.S. students spend more on math and language lessons than Korean students; on average, U.S. students spend 4.38 hours per week on math and 4.36 hours per week on language.

The mean hours/week students spend in shadow education is also higher for Korean students than for U.S. students. On average, Korean students spend about 3.07 hours per week on math shadow education and 2.15 hours on language shadow education. On the other hand, U.S. students spend about 0.58 hours per week on math shadow education and 0.42 hours per week on language shadow education.

About 86 percent of Korean schools in the sample are located in a city with a population of 100,000 or more. In contrast, about 39 percent of U.S. schools are located in a city of this size. About 83 percent of Korean schools are public, but schools, on average, receive only about 48 percent of school funding from the government, and the rest is financed by student fees (48 percent). U.S. schools are mostly public (93 percent) and are financed by the government (90 percent).

The variables, RESPRES (responsibility of school on resource allocation), RESPCURR (responsibility of school on curriculum management), and P_PRESS (parental pressure on academic standards of school), are derived based on answers from the principal's questionnaire. Korean principals answered that schools have high responsibilities (m= 0.79) on curriculum management, yet lower responsibilities (m= -

0.44) on resource allocation. In contrast, U.S. schools have low responsibility (m= -0.20) on curriculum management and high responsibility (m=0.40) on resource allocation. About 12 percent of Korean school principals answered that their school receives some pressure from parents on the academic standards, and about 34 percent of U.S. principals answered likewise.

About 98 percent of Korean teachers are certified and about 98 percent of Korean teachers are qualified. A separate test was conducted to examine multicollinearity for the variables included in the model, but these two variables are not found to be collinear. About 95 percent of U.S. teachers are certified and about 96 percent of U.S. teachers are qualified to teach at schools. About 77 percent of Korean schools answered that they use student achievement data to evaluate teacher performance. About 81 percent of U.S. schools answered that they use the student achievement data to evaluate teacher performance.

Only about 33 percent of schools publicly post student achievement data (SCHEVAL) in Korea. In contrast, about 89 percent of U.S. schools publicly post student achievement data. This is consistent with the current U.S. education policy (NCLB Act), which requires schools to publicly post the student achievement data by ethnic/racial groups. Similarly, about 62 percent of U.S. schools use student achievement data to evaluate principal's performance, whereas only about 28 percent of Korean schools use the student data to evaluate principal's performance.

School resources proxy variables, SCMATEDU (quality of school materials), IRATCOMP (ratio of computers for 15-year-olds to school size), TCSHORT (teacher shortage), STRATIO (pupil-teacher ratio), and EXCURACT (availability of extracurricular activities), suggest that the quality of educational resources is higher for U.S. schools. For instance, the mean value of SCMATEDU for U.S. schools is 0.51, compared to 0.06 for Korean schools. The size of language class is larger for Korean schools than for U.S. schools. On average, about 36 Korean students study in one language class, whereas about 25 U.S. students study in one language class. The pupil-teacher ratio is similar for both countries.

HLM Results

Table 2 presents results for null models for Korea and the U.S. on math and reading scores. The intercepts in the null models indicate the mean math and reading scores of U.S. and Korean students. The slight differences in the mean scores are due to the changes in the number of observations after the imputation method to address the missing data issue. The number of observations is 4,989 for Korean students and 5,233 for U.S. students.

Not only do the null models reveal the proportion of variance at the student- and school-levels, but the information also is essential for comparing between models. According to the null model for Korea, the intraclass correlation (ρ) is 0.38. The intraclass correlation, which measures how schools differ in student average performance (the proportion of variance at the school-level), intraclass correlation indicates that about 38 percent of the variance in math achievement is at the school-level and the rest is at the student-level. About 33 percent of variance in reading achievement is at the school-level. As for the U.S., about 30 percent of variance in math score is at the school-level, whereas

70 percent of variance is at the student level. In reading, 24 percent of variance in reading score is at the school-level. In sum, in both Korea and the U.S., the within-school variance is greater than the between-school variance, and yet, compared to the U.S. (ρ=0.30 for math, ρ=0.24 for reading), the proportion of total variance explained by school-level variance is greater for Korea (ρ=0.38 for math, ρ=0.33 for reading).

Table 2 Null Model for KOR and USA, by subject

	MATHEMATICS		READING	
	KOR	**USA**	**KOR**	**USA**
Intercept	545.1** (0.60)	484.8**[15] (0.71)	538.7** (0.70)	497.1** (0.78)
Random Effects				
School-level SD ($\sqrt{\psi}$)	54.7	49.6	45.3	47.3
Student-level SD ($\sqrt{\theta}$)	70	76.4	64.5	84
ρ	0.38	0.30	0.33	0.24
N	4,989	5,233	4,989	5,233

The high intraclass correlation for Korea is unexpected because the South Korean education system is highly standardized and uniform as a result of the Middle School and High School Equalization Policy. The degree of school differentiation can be explained by the tracking system that begins at the upper secondary education. The 15-year-old

[15] This dissertation does not examine data by ethnic groups, but the U.S. has a country-specific question that asks the participating students' ethnic/racial background. According to the data, Asian students, on average, scored 524 points, white students scored 515 points, black students scored 423 points, and Hispanic students scored 453 points on math in 2009.

students sampled in the PISA data are enrolled in high schools, which are tracked into "academic" and "vocational" schools. Among 4,989 students sampled in Korea, 317 students answered that they are enrolled in middle school, and the remaining 4,672 students are enrolled in high school: among them, 3,567 students attend academic schools and 1,105 students attend vocational schools. Because the tracking is largely determined by students' level of achievement, high schools are generally segregated based on the students' academic abilities.

Null models for Korean students in the academic and vocational schools, run separately to examine the mean scores and the intraclass correlation revealed that Korean students who attend academic high schools, on average, scored 571 points on math and 561 points on reading. In contrast, students who attend vocational high schools scored, on average, 481 points on math and 483 points on reading (See Appendix C-1). When the tracking system is considered, the intraclass correlation decreases by a sizeable amount: for academic high schools, the rho is 0.25 and for vocational schools, the rho is 0.16. The results show that the greater degree of school differentiation for Korea is mainly due to the tracking system that begins at the upper secondary education. This dissertation discusses the general Korean student sample, but for the interested audience, separate models for Korea for academic and vocational high schools are included in Appendices C-2 and C-3.

(1) What role do family SES and 'social/cultural capital' factors play in achievement outcomes for Korean and U.S. students? Is there a difference in the relative importance of family background factors on within-country achievement outcomes between Korea and the United States?

The relationship between family SES and student achievement was carefully examined to answer the first research question. For both Korea and the United States, family SES plays a significant role in explaining student achievement, but the magnitude of the effect[16] is stronger for the U.S. than for Korea.

Table 3 shows two models: Model 1 presents results using the summary composite SES variable, ESCS; Model 2 examines the results using each of the individual family SES components. The purpose of this comparison is to examine whether the composite measure is a fair summary of the effects of individual SES components.

Table 3 Family SES and Assessment scores, by country

	MODEL (1)				MODEL (2)			
	Math		Reading		Math		Reading	
	KOR	USA	KOR	USA	KOR	USA	KOR	USA
Intercept	547.8** (0.59)	481.3** (0.74)	540.9** (0.72)	493.3** (0.82)	491.1** (10.73)	414.4** (7.09)	491.6** (9.42)	428.6** (7.89)
Composite Family SES Variable								
ESCS	17.3** (1.57)	26.8* (1.65)	14.3** (1.43)	28.5** (1.75)	-	-	-	-
Individual Components of Family SES								
HISEI	-	-	-	-	0.5** (0.12)	0.8** (0.11)	0.4** (0.09)	0.9** (0.11)
PARED	-	-	-	-	2.1* (0.82)	2.5** (0.53)	1.6* (0.68)	2.0** (0.58)
WEALTH	-	-	-	-	-2.2 (1.68)	4.8** (1.74)	-6.1** (1.55)	0.4 (1.80)
HEDRES	-	-	-	-	11.4** (1.98)	3.8* (1.48)	13.0** (1.66)	7.9** (1.59)
CULT-POSS	-	-	-	-	1.5 (1.83)	11.6** (1.35)	4.8** (1.74)	13.8** (1.56)

[16] The term 'effect' in these analyses is used to indicate 'relationship' and it does not imply a 'causal relationship'. In the case of SES variables, however, it is well-established that these are causal relationships.

Random Effects								
School-level SD ($\sqrt{\psi}$)	49.4	40	41	36.7	48.8	40	39.6	36.8
Student-level SD($\sqrt{\theta}$)	69.2	74	63.8	81.7	68.8	73.7	62.9	80.9
ρ	0.34	0.23	0.29	0.17	0.33	0.23	0.28	0.17
R^2	0.08	0.15	0.07	0.14	0.10	0.15	0.11	0.15

*Note: ** indicates p-value <0.01 and * indicates p-value <0.05*
Note 2: The number of observations (N: KOR=4,989 USA=5,233) remains the same across different models, since single imputation method was implemented to address the missing data

Willms used a socioeconomic gradient to describe the relationship between a social outcome and socioeconomic status for individuals in a specific jurisdiction, such as a school or a country (Willms 2010, 1014). Socioeconomic gradient is comprised of 'the level', 'the slope' and 'the strength' of the outcome-SES relationship. The level of socioeconomic gradient is measured by the SES-adjusted mean of student achievement. A hypothetical student with a mean SES is expected to have a math score of 548 points in Korea and 481 points in the United States. The ESCS variable is centered around the OECD mean (an OECD mean of 0 and a standard deviation of 1), and thus the intercepts represent the mean values of math and reading for Korea and the U.S. of students who have the family SES at the OECD mean.

While the level of achievement is higher for Korea than for the U.S., the SES slope is steeper for the U.S. than for Korea. The expected math score for Korean students increases by 17.3 points for a one-standard-deviation increase in the ESCS scale, whereas the expected math score for U.S. students increases by 27 points for a one-standard-deviation increase in the ESCS scale. The difference in achievement between the two countries is larger for students with lower family SES than students with higher family SES (See Figure 6).

The strength of the gradient, measured by the proportion of variance in the outcome explained by family SES, is stronger for the U.S. than for Korea (R^2 in Table 3). The ESCS scale explains about 8 percent of variation in Korean students' math outcome, whereas the ESCS scale explains about 15 percent of variance in U.S. students' achievement outcome.

This dissertation confirms the earlier findings reported by Willms that considerable amount of variation in the outcome measure is associated with family SES in the United States, and it is likely that a disproportionate number of poor achievers are among low SES students (Willms 2010). This study also confirms the finding that the family SES explains smaller proportion of total variance in achievement outcomes for the Korean model than the U.S. model. In other words, the effect of family SES on student achievement outcomes is stronger for the U.S. than for Korea, and this finding indicates that while the association between family SES and achievement is strong in the U.S., it is less likely that low achievers in Korea are among low SES students in Korea. The finding is similar for reading performances of students in Korea and the United States.

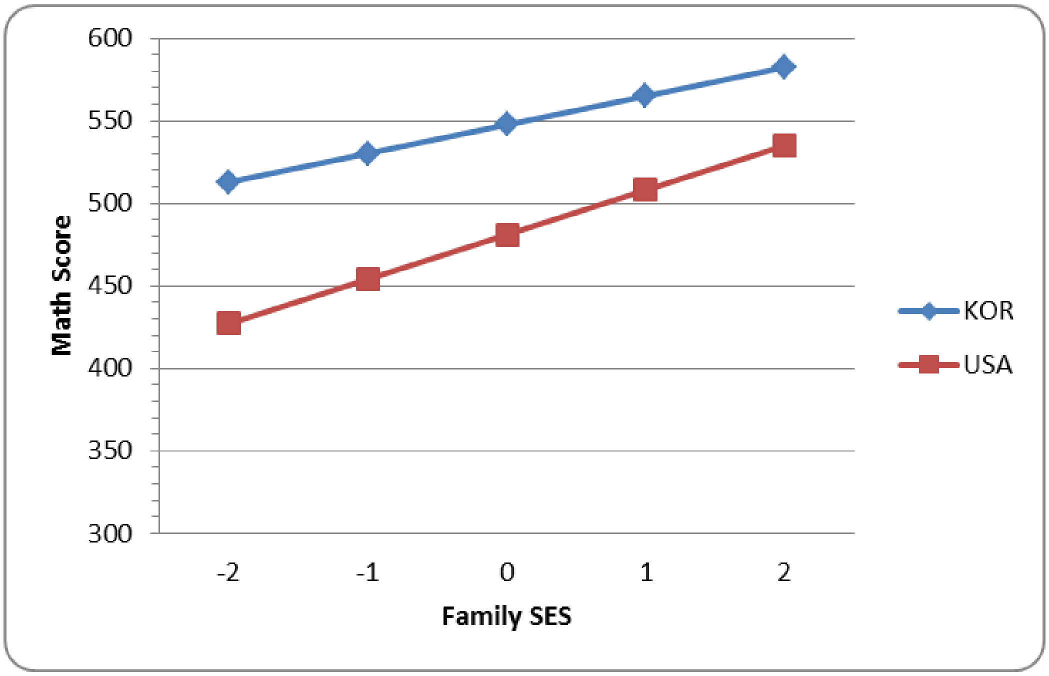

Figure 6 Socioeconomic Gradients for USA and KOR

To examine whether individual components of family SES have different importance and significance on within-country achievement outcomes in Korea and the U.S., a separate analysis was conducted with the five individual SES components. Results were presented side-by-side with results from the SES composite variable, after other student characteristics were considered. In other words, the Model 2 in Table 3 shows the HLM results for coefficients of individual SES variables, HISEI (highest parental occupational level), PARED (highest parental education), WEALTH (household wealth), HEDRES (home educational resources), and CULTPOSS (home cultural possession), which make up the composite ESCS variable. Because these variables have not been standardized (or centered around the mean), the intercept represents a student with zero

HISEI, PARED, WEALTH, HEDRES, and CULTPOSS, and thus the intercepts in Table 3, Model 2 are not very meaningful.

As for the individual components of family SES factor, parental occupation and parental education have a statistically significant relationship with assessment scores in both Korea and the United States. And yet, their substantive significance on scores are relatively smaller compared to other SES components. For instance, a one-year increase in PARED is associated with a 2.1-point increase in math score in Korea and a 2.5-point increase in the United States. In Korea, WEALTH is negatively associated with math and reading scores, and the relationship is not statistically significant for math. On the other hand, the relationship is statistically significant for math in the U.S., but not for reading. The HEDRES variable is positive and statistically significant in math and reading for Korean students; a one-standard-deviation increase in HEDRES is associated with about an 11-point increase in math score and a 13-point increase in reading score. CULTPOSS is not statistically significant for math, but statistically significant for reading for Korean students. A one-standard-deviation increase in CULTPOSS is associated with a 5-point increase in a Korean student's reading score. As for U.S. students, both HEDRES and CULTPOSS are statistically and substantively significantly related with math and reading scores. A one-standard-deviation increase in HEDRES is associated with a 3.8-point increase in math and an 8-point increase in reading; and a one-standard-deviation increase in CULTPOSS is associated with a 12-point increase in math and a 14-point increase in the reading score of a U.S. student, holding all other family SES components constant.

Table 4 presents the HLM results for all of the student background variables using the summary ESCS measure. These variables include the student gender, student's family ESCS, the family structure (whether the student lives in a two-parent household or not), and the student's approaches to learning.

Table 4 Student background factors and the assessment scores, by country

	MATHEMATICS		READING	
	KOR	USA	KOR	USA
Intercept	545.5** (3.51)	458.0** (1.99)	563.6** (3.07)	490.8** (2.04)
MALE	13.0** (3.64)	19.2** (2.32)	-29.4** (3.12)	-25.5** (2.23)
MEMOR	-6.8** (1.57)	-21.1** (1.47)	1.7 (1.33)	-21.7** (1.62)
CSTRAT	24.5** (1.84)	22.1** (1.22)	20.9** (1.49)	26.9** (1.51)
PARENTS	2.2 (3.61)	20.1** (2.46)	-3.0 (3.19)	22.1** (1.64)
ESCS	12.2** (1.40)	21.9** (1.63)	9.2** (1.29)	22.1** (1.64)
Random Effects				
School-level SD ($\sqrt{\boldsymbol{\psi}}$)	45.4	35.6	34.9	32
Student-level SD ($\sqrt{\boldsymbol{\theta}}$)	66.3	70.5	59.7	76.7
ρ	0.32	0.20	0.25	0.15
R^2	0.18	0.25	0.23	0.26
*Notes: ** indicates p-value <0.01, * indicates p-value <0.05; the number of observations is the same across all models (KOR, N=4,989; USA, N=5,233)*				

Adding the student-level variables in the model affected the intraclass correlation of models for Korea as much as for the United States. Compared to the intraclass correlation (ρ=0.38) presented in the null model (Table 2) for math for Korean students, the intraclass correlation (ρ=0.32) is lower. The same is true for the U.S., yet the magnitude of decline is significantly greater for the U.S. than for Korea; the intraclass correlation for math (ρ=0.20) is lower than the intraclass correlation of the null model for math (ρ=0.30). This indicates that, as discussed in the methodology chapter of this dissertation, schools in Korea and the U.S. differ in the mean and range of students with regard to the student-level variables, but the magnitude of difference is greater in the United States. Furthermore, the impact is greater for reading for both countries than for math; the intraclass correlation for reading for Korea (ρ=0.25) saw a substantive decline from the null model (ρ=0.33), and the intraclass correlation for reading for the U.S. (ρ=0.15) is an almost nine percent decline from the null model (ρ=0.24).

The intercepts now show the estimated achievement scores controlling for all of the student background characteristics. Note that while the Korean math intercept is about the same as the null model (545.1 vs. 545.5), the U.S. intercept has declined to 458 from 485. This means that the difference between U.S. and Korean math scores has actually increased to 87 points. In other words, after controlling for student background, the disparity in math scores between the two countries has increased. Similarly, the reading gap has increased from 42 to 64 points.

Student characteristics and their relationships with assessment have some differences and similarities between Korea and the United States. Male students tend to score higher than female students on math in both countries, and they tend to score lower than female students on reading in both countries.

Memorization strategies generally have negative effects in both countries, but they are much weaker in Korea than in the United States: A one-standard-deviation increase in memorization is associated with a 7-point decrease in math scores in Korea, and it is not statistically significant for reading. In contrast, a one-standard-deviation increase in memorization strategies is associated with a 21-point decline in math score and a 22-point decline in reading scores for a U.S. student, holding other student characteristic variables constant.

In both countries, control strategies have strong positive effects on both reading and math. A one-standard-deviation increase in control strategies score is associated with a 25-point increase in math and a 22-point increase in reading in Korea, and increases of 22 points and 27 points for math and reading in the U.S., respectively.

After controlling for all student characteristic variables, including student's gender, approaches to learning, and family structure (whether the student lives in a two-parent household), the magnitude of the effect of family SES (ESCS) on math and reading scores declines by little, but the variable remains highly statistically significant. A one-standard-deviation increase in ESCS is associated with a 12-point increase in math for Korean students and a 9-point increase in reading for Korean students. Similarly, a one-standard-deviation increase in ESCS is associated with a 22-point increase in math for U.S. students and a 22-point increase in reading scores. Unlike in the U.S., where students who live in a two-parent household tend to score considerably higher on math and reading (about 20-22 points higher in math and reading), in Korea, the relationship is not statistically significant.

In sum, the findings together confirm my hypothesis that family background factors play an important role in explaining achievement outcomes in the U.S. and Korea, but the importance is greater for the U.S. than for Korea. The findings further indicate that in both countries, the within-school variance is greater than the between-school variance. The between-school variance is larger for Korea than for the United States, because of the high school tracking system. When the tracking system is considered, the between-school variance drops significantly lower than that of the United States.

In addition, the findings show that the association between the family SES and assessment scores is stronger in the U.S. than in Korea. Since students from low-SES are more likely to achieve less than students from higher-SES, the U.S. students can benefit from policies that support the lower-SES students to achieve higher. Finally, among the individual components of family ESCS, cultural possessions (e.g. classic literature, books of poetry, and works of art) and home educational resources (e.g. desk, reference books, dictionaries) are highly substantively associated with student achievement in the U.S., and in Korea, home educational resources is highly and substantively associated with achievement scores.

(2) What role do school characteristic differences play in achievement outcomes of Korean and U.S. students? What is the relative importance of the school-level factors in explaining within-country achievement outcomes in Korea and the United States?

This section examines to what extent school-level variables explain achievement outcomes of Korea and U.S. students. Table 5 presents the HLM results for students' in-school and shadow education lesson hours and school-level variables, after the student-level characteristics and family background factors are controlled.

The coefficient of determination (R^2) was estimated to compare the proportional reduction in prediction error variance of the model without covariates (the null model) with the full models used in the analysis (Rabe-Hesketh and Skrondal 2012, 134). Compared to the null models without any covariates, the student- and school-level variables together explain about 32 percent of total variance in Korean students' math scores and about 33 percent of total variance in their reading scores. Compared to the null models, the variables together explain about 36 percent of total variance in U.S. students' math scores and about 36 percent of total variance in their reading scores.

The coefficients of determination (R^2) of student-level-variables only model presented in Table 4 showed that the student-level variables explain about 18 percent of variation in math scores of Korean students and about 25 percent of variation in math scores of U.S. students. By adding school-level variables, the model for Korea added about 14 percentage points and the model for the U.S. added about 11 percentage points. This indicates that the strength of school-level variables is larger for the Korean model than for the U.S. model.

Table 5 HLM results for Korea and USA, by subject

	MATHEMATICS		READING	
	KOR	USA	KOR	USA
Intercept	486.3** (21.96)	487.9** (11.47)	526.6** (22.02)	529.2** (12.88)
MALE	12.4**	19.3**	-29.2**	-26.0**

	(3.43)	(2.27)	(2.92)	(2.17)
ESCS	9.4** (2.35)	17.2** (1.74)	9.4** (1.76)	17.8** (1.72)
PARENTS	1.4 (3.45)	18.8** (2.53)	-3.3 (3.10)	21.7** (2.71)
MATH(/LANG)-HRS	5.6** (1.52)	5.5** (0.85)	0.3 (1.36)	5.2** (0.87)
SHADOW-(MATH/LANG)	3.9** (0.67)	-8.5** (0.73)	1.0 (0.65)	-13.1** (0.95)
MEMOR	-6.7** (1.54)	-19.3** (1.44)	1.6 (1.30)	-19.1** (1.60)
CSTRAT	23.0** (1.74)	21.1** (1.21)	20.5** (1.41)	25.4** (1.52)
URBANSCH	-11.7** (2.79)	-19.1** (2.08)	-18.1** (2.43)	-12.0** (2.16)
SSIZE	3.6** (0.35)	0.7** (0.16)	2.4** (0.27)	0.6** (0.14)
PUBLIC	4.3** (2.60)	-12.3** (7.47)	-3.6 (2.51)	-54.3** (9.43)
GOVFUND	-0.5 (0.03)	-0.2 (0.05)	-0.4** (0.03)	0.2** (0.06)
RESPRES	15.8** (0.73)	-1.2 (1.11)	10.5** (0.91)	-0.6 (1.08)
RESPCURR	3.4** (1.25)	-1.6 (1.15)	3.7** (0.98)	0.9 (1.24)
P_PRESS	39.6** (3.22)	14.3** (2.46)	31.8** (2.40)	11.6** (2.68)
PCCERT	0.1 (0.12)	-0.5** (0.11)	0.2** (0.06)	-0.4** (0.12)
PCQUAL	0.1 (0.11)	0.1* (0.04)	0.1 (0.10)	-0.0 (0.04)
TCHEVAL	0.8 (1.44)	7.5** (2.56)	-0.2 (1.45)	1.8 (2.93)
SCHEVAL	-4.6* (2.17)	24.2** (3.50)	-6.1** (1.55)	27.9** (3.79)
PRINCIPEVAL	7.6** (1.86)	-14.9** (2.56)	4.2* (2.07)	-11.4** (2.81)
IRATCOMP	-20.6** (3.40)	0.7 (2.36)	-17.0** (2.55)	-5.6** (2.12)
TCSHORT	-0.6 (0.91)	-6.6** (0.99)	-1.8* (0.73)	-6.1** (1.07)
STRATIO	-1.6** (0.43)	-1.6** (0.18)	-1.1** (0.37)	-1.7** (0.19)

SCMATEDU	4.8** (1.04)	0.6 (0.88)	2.8** (0.76)	0.4 (0.99)
EXCURACT	-0.8 (1.01)	6.7** (1.37)	1.1 (1.26)	6.2** (1.44)
CSIZE	0.9 (0.70)	0.6** (0.22)	1.2 (0.57)	0.6** (0.21)
ESCS* SHADOW	0.4 (0.67)	-0.3 (0.89)	-0.4 (0.60)	-1.2 (1.12)
ESCS* P_PRESS	0.7 (4.82)	11.1** (2.77)	0.7 (4.07)	8.4** (3.25)
Random Effects				
School-level SD ($\sqrt{\boldsymbol{\psi}}$)	32.3	23.7	24.6	20.6
Student-level SD ($\sqrt{\boldsymbol{\theta}}$)	65.8	68.9	59.7	74.5
ρ	0.19	0.11	0.15	0.07
R^2	0.32	0.36	0.33	0.36
*Note: ** indicates p-value <0.01, * indicates p-value <0.05; N for KOR=4,989 and N for USA =5,233.*				

While adding the student-background characteristics has widened the math score gaps between Korean and U.S. students (Korea M=545.5 vs. USA M=458), adding the school variables has decreased the intercept of the Korean model and increased the intercept of the U.S. model, resulting in the similar levels of estimated performance of schools in Korea and the United States. After controlling for all student- and school-level variables, the intercept for Korea is 486 points for math, and the intercept for the U.S. is 488 points; for reading, the adjusted score is 527 points for Korea and 529 points for the United States.

After controlling for other student- and school-level variables, the in-school math hour variable is positively and substantively associated with math performance for both Korean and U.S. students, and the impact is of about the same magnitude for both countries. On average, a one-hour increase in math lesson hours is associated with a 5.5-point increase in math scores. The in-school language hour variable also is positively associated with school performance in the U.S., and of about the same magnitude as math. Interestingly, in-school language hours are not associated with reading scores in Korea.

Shadow education in math is positively and significantly associated with Korean students' scores, after other critical variables are considered. However, the variable is negatively associated with U.S. student performance. On average, a one-hour increase in shadow math education is associated with a 4-point increase in math scores for Korean students and about a 9-point decrease for U.S. students. For reading scores, shadow education is not related in Korea but a one-hour increase in the U.S. results in a 13-point decrease in reading scores.

The inverse relationship for Korea and the U.S. deserves a closer look, since this variable is one of the few variables that has a significant relationship with students' academic achievement. Whether the relationship points to causal effects of shadow education on student performance (positive effect in Korea and negative effect in the U.S.), or whether the variable indirectly measures motivation, intellectual curiosity, or parental involvement in education (as shadow education is privately financed) needs to be closely examined. Also, there is a possibility of the issue of self-selection; it could be that in Korea, high-achievers are more likely to take more advantage of the shadow education whereas in the U.S., low-achievers are those that primarily participate in shadow education.

If shadow education is a result, not a cause of, achievement scores, it should not be an independent variable in the model. Accordingly, run as a sensitivity analysis, the same models without the shadow education variable included reveal that most of the coefficients are similar to the model with shadow education. The MATHHRS (in-school math hours) variable is slightly higher for Korea. Interestingly, the intercept is about the same for Korea and the U.S., suggesting that achievement is similar for both countries after controlling for all student and school factors, and it was not a result of including shadow education. The result table is included in Appendix D.

In order to examine whether the effect of shadow education on school performance depends on the family SES, the interaction terms between family SES and shadow education (ESCS*SHADOWMATH, ESCS*SHADOWLANG) is also included in the model. The interactions are not statistically significant for either country, indicating that the relationship is relatively uniform for students from upper and lower SES families.

In terms of school demographics and characteristics, U.S. and Korean schools share similarities. Also, controlling for all other school-level variables, urban schools (located in a city of population 100,000 +) in both Korea and the United States tend to score lower in math and reading than non-urban schools. While the variable is statistically significant for both Korea and the U.S., the magnitude is greater for U.S. schools. Urban schools in the U.S., on average, score about 19 points lower on math than schools located in non-urban areas (12 points in reading), whereas urban schools in Korea score about 12 points lower on math than non-urban schools (18 points in reading), holding all else equal.

School size (SSIZE) is positively associated with student achievement in Korea but not in the United States: a 100-student increase is associated with a 3.6-point increase in school math score and a 2.4-point increase in school reading score in Korea, holding all else constant, but the association is not substantively significant in the United States.

Korean and U.S. schools tend to have different patterns in terms of the relationships between the external factors and the academic performance of students. Public schools (PUBLIC) in Korea perform significantly better than private schools. In the U.S., however, public schools on average have lower academic performance than private schools. Students in public schools (PUBLIC) tend to score about 15 points lower on math and 54 points lower on reading than students in private schools ($p\text{-value}<0.01$).

In terms of the percentage of government funding (GOVFUND), it is not substantively significant in Korea and the United States. School autonomy is associated with school performance differently in Korea and in the United States. Schools that have greater responsibility than other governing bodies, including regional/local education authority or national education authority, on resource allocation (RESPRES) and curriculum management (RESPCURR) have higher math and reading scores in Korea. Particularly, the school responsibility for resource allocation has high substantive significance on school performance. A one-standard-deviation increase in RESPRES is associated with about a 16-point increase in a school's mean math score and an 11-point increase in its reading score, holding all else equal. The results show that Korean schools are standardized, but greater autonomy in terms of resource allocation and curriculum management is associated with higher school performance in Korea. Yet greater school responsibility for resource allocation (RESPRES) and curriculum management (RESPCURR) does not have any statistically or substantively significant relationships with U.S. school performance, either math or reading.

Both Korean and U.S. school performance is associated with the presence of parental pressure (P_PRESS) on a school's academic standard. Korean schools with some parental pressure on achievement tend to score about 40 points higher on math than schools with no parental pressure, holding everything else constant. In reading, the magnitude is about 32 points for schools that receive parental pressure. U.S. schools that receive pressure from parents tend to have a math score about 14 points higher than schools that do not and a reading score about 12 points higher than schools that do not.

When the variable is interacted with family SES (ESCS) to examine whether the effect on school performance depends on the family SES, the variable showed no statistical significance for Korean schools, but it showed highly substantive and statistical significance for U.S. schools. This finding indicates that family pressure is important in Korea, but whether the pressure comes from higher family SES does not matter. And yet, in the U.S., parental pressure is more effective when it comes from the parents with higher family SES.

Even the teacher-related variables have different patterns with school performance in Korea and the United States. The percent of certified teachers (PCQUAL) is not significantly related with the math performance of Korean schools, but, while the effect is not strong, it is statistically and significantly and negatively related with math performance. A one-percent increase in the number of certified teachers is associated with a 0.5-point decline in the mean math score of U.S. schools. On reading, the coefficient of PCCERT is positive and its magnitude, 0.2 points, is also substantively significant on the reading performance of Korean schools. It is negative and statistically significant on reading performance of U.S. schools, but its magnitude of the effect, -0.4 points, is also substantively insignificant. As for the percent of qualified teachers (PCQUAL), it is only statistically significantly related with the math performance of U.S. schools, but even so, the magnitude of the effect, 0.1 points, is not strong. In the U.S., whether student achievement data is used to evaluate teacher performance (TCHEVAL) is positively associated with the school math performance - U.S. schools that use student achievement data to evaluate teacher performance, on average, have a math score about 8 points higher than those that do not. TCHEVAL, however, is not statistically significantly associated with the reading performance. The measure is also not statistically associated with school performance in Korea.

In terms of other accountability measures, the school accountability measure (SCHEVAL) is positively and substantively associated with school performance in the United States. A school that publicly posts student achievement data is likely to have a math score about 24 points higher than a school that does not and a reading score about 28 points higher than a school that does not. The measure is only statistically significant on reading performance in Korea, but in the opposite direction of the U.S., where schools that publicly post student achievement data are likely to have a reading score about 6 points lower, holding all else constant. Similarly, the pattern for PINCIPEVAL is different between the two countries, but in the opposite direction. PRINCIPEVAL (whether the school uses the student achievement data to evaluate a principal's performance) is negatively and statistically related with school performance in the United States (γ_{math}= -14.9, $\gamma_{reading}$=11.4), but positively and statistically significantly related with school performance in Korea (γ_{math}= 7.6, $\gamma_{reading}$=4.2).

School resource proxy variables also have different relationships with school performance in Korea and the United States. School resources were measured by several proxy variables. Ratio of computers to school size (IRATCOMP), teacher shortage (TCSHORT), pupil-teacher ratio (STRATIO), quality of school resources (SCMATEDU), and the availability of extracurricular activities (EXCURACT) are included in the model as proxy variables. IRATCOMP is negatively associated with school performance in Korea, but not statistically significantly related with school performance in the United States. It could be that the ratio of computers is a proxy measure of other school characteristics in Korea, but this finding calls for further research.[17] As it is expected, teacher shortage (TCSHORT) takes on the negative coefficients. This variable is not statistically significantly associated with math scores in Korea, but statistically significant for math scores in the U.S.: a one-standard-deviation increase in teacher shortage is associated with -7 points in the math score of U.S. schools, holding all else constant. The variable is statistically significant on reading in both countries; a one-standard-deviation increase in teacher shortage is associated with a 1.8-

[17] A multicollinearity test was conducted to test its association with other school resource proxy variables, but these variables are not substantively or statistically significantly related with other variables. Yet, a model without the school resource proxies had a different result for URBANSCH – the variable changed the direction for Korea.

point decline in reading performance of Korean schools and a 6-point decline for U.S. schools.

Among the school resources proxy variables, the pupil-teacher ratio (STRATIO) seems to be the only variable with consistent patterns across these two countries. STRATIO is negatively and statistically significantly associated with school performance in Korea and in the United States. Quality of school materials (SCMATEDU), however, is only substantively and statistically associated with school performance in Korea. A one-standard-deviation increase in SCMATEDU is associated with a 5-point increase in a Korean school's math score and a 3-point increase in reading score. The SCMATEDU remains statistically insignificant on school performance in the United States. In contrast, the availability of extracurricular activities (EXCURACT) is statistically and substantively significant on school performance in the U.S., whereas it is not in Korea: a one-standard-deviation increase in EXCURACT is associated with a 7-point increase in math scores and a 6-point reading scores of U.S. schools, holding all else equal.

Class size (CSIZE) is not statistically and substantively associated with school performance in Korea. And yet, it is positively statistically associated with school performance in the United States; a one-student increase in the number of students in the language class is associated with a 0.6 point increase in the math and reading scores of U.S. schools.

In sum, perhaps due to the drastic differences in the way their educational systems are set up, the effects of the school-level variables differ considerably between Korea and the United States. The shadow education effects have the most drastic differences between Korea and the United States. Whether the effect is causal or not deserves a closer examination, as the magnitude of the relationship is significant both in Korea and in the United States. In addition, parental pressure plays a great role in school performance in Korea and the U.S., and educational researchers and policy makers should pay attention to the finding that in the U.S., the effect of parental pressure depends on family SES. Furthermore, results show that school and teacher accountability have a positive relationship with the U.S. school performance. Whether the relationship is causal needs to be further examined. As for Korea, the finding that greater school autonomy is associated with higher school performance deserves close attention from policy makers and educational researchers.

(3) What role do 'shadow education' and other government policies (e.g. accountability, class size, teacher performance evaluation, ability grouping) play in explaining between KOR-US achievement outcomes? Do 'shadow education' and other government policies contribute to or lessen the country differences?

In order to examine how much of the student- and school-level variables are accountable for the cross-country achievement difference, a separate HLM analysis, where both countries are included in the same model, was conducted (See Table 6). Model 1 is a null model without any covariates, but with a dummy variable representing whether or not the school is in Korea. Model 2 includes the same set of variables, and some of the policy variables are interacted with the country dummy variable to examine if the relationship depends on the country of residence. ESCS, PARENTS, MATHHRS/LANGHRS, SHADOWMATH/SHADOWLANG, PUBLIC, RESPRES, RESPCURR, TCHEVAL, SCHEVAL, SCMATEDU, and CSIZE were interacted with the country dummy variable (KOR=1) to examine whether these policies have different relationships in Korea than in the United States, but except for PARENTS, SHDAOWMATH/SHADOWLANG, RESPRES, SCHEVAL, and CSIZE, other variables did not show any statistical significance. Table 6 presents the HLM results of variables with statistically significant interaction terms. The full table is included in Appendix E for interested readers.

Table 6 HLM Results for Korea and U.S. achievement differences in math and reading

	MATH		READING	
	NULL (1)	MODEL (2)	NULL (1)	MODEL (2)
INTERCEPT	485.4** (1.55)	444.2** (23.34)	497.5** (1.41)	491.1** (18.69)
KOR-US difference	59.6** (4.84)	-17.4 (31.36)	40.2** (4.25)	-14.3 (27.25)
Student-level Variables				
MALE		18.5** (2.04)		-25.9** (1.96)
ESCS		20.1** (1.51)		19.1** (1.40)
PARENTS		19.7** (2.58)		22.3** (2.75)
KOR-US difference		-13.1** (4.82)		-22.5** (4.38)
MATHHRS (/LANGHRS)		5.5** (0.85)		5.0** (0.86)
SHADOWMATH (/SHADOWLANG)		-8.8** (0.76)		-13.4** (0.97)
KOR-US difference		14.3** (1.24)		15.0** (1.33)
MEMOR		-17.9** (1.19)		-16.6** (1.35)

CSTRAT	21.8** (1.08)	25.2** (1.36)
School-level Variables		
URBANSCH	-15.2** (4.53)	-10.0** (3.80)
SSIZE	0.8** (0.28)	0.7** (0.23)
PUBLIC	5.2 (6.40)	-3.4 (6.15)
GOVFUND	-0.4** (0.08)	-0.2 (0.07)
RESPRES	0.4 (2.71)	1.0 (2.26)
KOR-US difference	16.3** (5.54)	10.9** (4.61)
RESPCURR	-1.0 (2.26)	0.4 (1.70)
P_PRESS	27.3** (4.31)	19.5** (3.53)
PCCERT	-0.4* (0.19)	-0.4* (0.17)
PCQUAL	0.3* (0.14)	0.3* (0.09)
TCHEVAL	-0.5 (3.71)	-0.3 (3.31)
SCHEVAL	20.0** (6.64)	14.7** (4.82)
KOR-US difference	-26.7** (8.37)	-23.5** (7.04)
PRINCIPEVAL	-4.7 (4.59)	-4.4 (3.43)
IRATCOMP	0.9 (4.92)	-2.4 (3.90)
TCSHORT	-3.1 (1.93)	-3.8* (1.63)
STRATIO	-0.9** (0.31)	-1.0** (0.29)
SCMATEDU	1.6 (2.05)	1.9 (1.57)
EXCURACT	2.9 (2.20)	4.4* (1.81)
CSIZE	0.4	0.5*

		(0.23)		(0.22)
KOR-US difference		3.0** (0.92)		2.7** (0.75)
Random Effects				
School-level SD ($\sqrt{\psi}$)	43.3	23.6	41.9	21.3
Student-level SD ($\sqrt{\theta}$)	79.4	70	83.9	73.4
ρ	0.23	0.10	0.20	0.08
R^2	-		-	
*Note: ** indicates p-value <0.01, * indicates p-value <0.05*				

The Korea-US difference is about 60 points for math and about 40 points for reading before any covariates, and the magnitude of the relationship is statistically significant. However, after the student- and school-level controls are included in the model, the Korea-US difference becomes negative (math= -17.4, reading=-14.3) and the difference loses the statistical significance.

Among the variables that have substantive differences across the two countries, the coefficient of PARENTS (whether the student lives in a two-parent household) is 20 points for math and 22 points for reading, and the Korea-US difference is statistically significant and indicates that the slope is less steep for Korea. Holding all else constant, a female Korean student who lives in a two-parent household scores (444.2+19.7-13.1=450.8) 451 points on math, and a similar female U.S. student scores (444.2+19.7=463.9) 464 points on math. On reading, a similar female Korean student scores (491.1+22.3-22.5=490.8) 491 points whereas a U.S. female student scores 513.3 points. As prior research has already established (e.g. Armor 2003), in the U.S., ethnic/racial minorities are more likely to live in a single-parent household and the measure of two-parent household could have been confounded with the measure of racial/ethnic minority.

SHADOWMATH and SHADOWLANG are negatively and statistically significantly related. On average, SHADOWMATH is associated with a 9-point lower math score, holding all else equal, and SHADOWLANG is associated with a 13-point lower reading score, holding all else equal. The Korea-US difference for SHADOWMATH is about 14 points for math and for SHADOWLANG is about 15 points for reading, and these are highly statistically significant. Holding all else equal, a Korean female student who does not attend shadow education is expected to score (444.2-17.4) 426.8 points and a U.S. female student who does not attend shadow education is expected to score 444.2 points. A Korean female student who attends an hour of math shadow education scores (444.2-17.4-8.8+14.3) 426.8 points on math, and a similar U.S. female student who attends an hour of math shadow education scores (444.2-8.8) 435.4 points on math (See Figure 7). Similarly, on reading, a similar female student who does not attend language shadow education scores 491 points on reading, but a female Korean student who attends an hour of language lesson scores (491.1-14.3-13.4+15.0) 478.4 points on reading and a female U.S. student who attends an hour of language lesson scores 477.7 points on reading. The interaction could point to the opposite causal directions in these two countries: Students who attend shadow education in Korea are more likely to achieve higher on assessments than students who do not, but students in the U.S. are more likely to achieve lower if they attend shadow education. However, because the data is cross-sectional and because prior achievement and other unobservable variables (e.g. study attitudes, intelligence) were not controlled for, the direction of the causal relationship could be the reverse. In other words, as it was discussed earlier, high-achieving Korean students might self-select into shadow education participation, achieving even higher than students who do not, but in the U.S., low-achieving students might be the only ones who participate in shadow education.

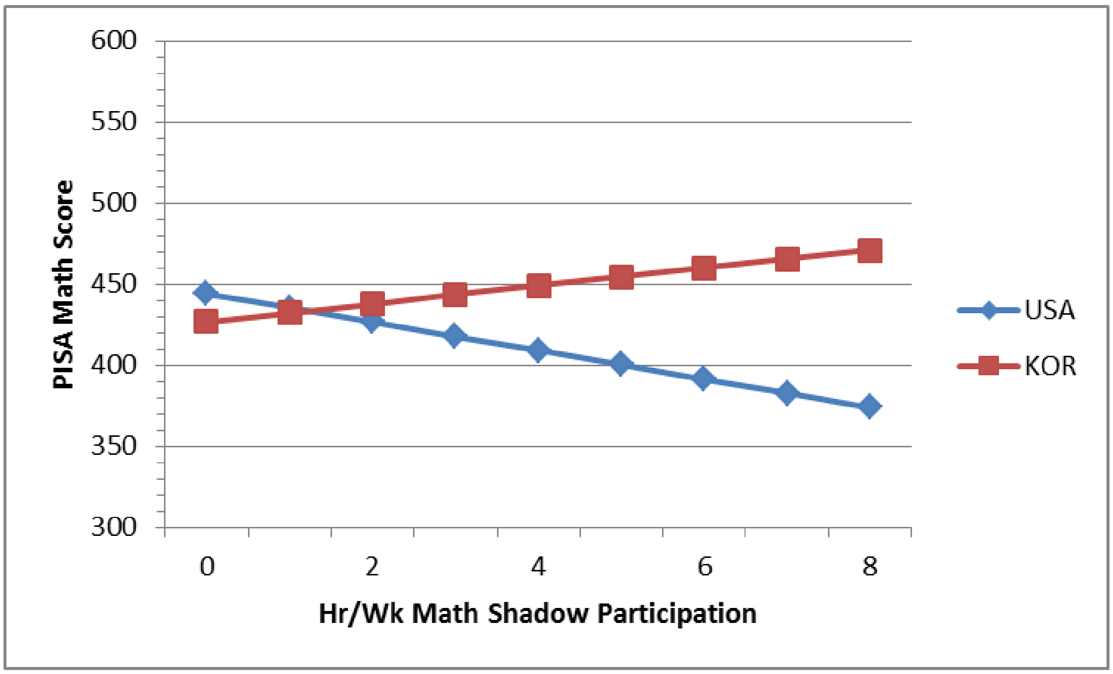

Figure 7 The adjusted math score of a typical student by the number of math shadow education

Greater school autonomy (RESPRES and RESPCURR) does not have statistically significant relationship with student achievement, but the interaction term is statistically significant. The Korea-US difference on RESPRES (school responsibility on resource allocation) is 16 points for math and 11 points for reading, and these are statistically significant.

Whether schools publicly post student achievement data (SCHEVAL) also has an interesting opposite influence on scores between the U.S. and Korea. Holding all else constant, U.S. schools that publicly post achievement data, on average, scores (444.2+20.2) 464.4 points on math, whereas Korean schools that publicly post achievement data, on average, score (444.2+20.0-26.7) 437.5 points on math. The pattern is similar for reading. Similarly, the effect of class size is not statistically significant, but the interaction term is significant, indicating that an increase in class size is likely to be associated with score increase for schools in Korea only.

In sum, after the covariates are included in the model, the Korea-US difference is no longer statistically significant. Before adding any covariates, the difference is close to 60 points for math and over 40 points for reading, but after the covariates are included in the model, the difference takes a different direction and becomes statistically insignificant. Also, while family SES variables are equally important in Korea and the United States, whether the student lives in a two-parent household is more important in the U.S. than in Korea. Greater school autonomy is more important in Korean school performance than U.S. school performance, and accountability measure is associated with positive outcomes in the U.S. while it is negatively associated with schools in Korea. Finally, the shadow education variables (both SHADOWMATH and SHADOWLANG) are statistically significant, but the direction of the relationship is different between Korea and the United States, confirming my hypothesis. In fact, shadow education is the most substantive policy variable that explains the between-country achievement disparities.

Limitations of the Study

In spite of numerous benefits, PISA data has several limitations. Firstly, even when between-country differences in background conditions are adjusted for, social research is limited because not everything in the society can be measured and quantified. Particularly the variables that cannot be observed, such as students' intelligence, parental affection towards the children, students' motivation, and other cultural factors specific to a certain society, are not accounted for in the analysis. Oftentimes, proxy variables that directly or indirectly measure the unobservable variables are used, e.g. school resources variables are accounted for in the dissertation as a way to capture the quality of school. Whether these proxy variables adequately captures the multi-dimensional aspect of the unobservable variable, however, can be debated.

In addition, the cross-sectional nature of the dataset and the lack of prior achievement data make it difficult for researchers to draw any causal relationships based on the results (see Loveless 2002; Porter and Gamoran 2002). As was discussed by Raudenbush and Kim, even with prior achievement data, drawing causal relationships based on the results from a large-scale survey data can be problematic (Raudenbush and Kim 2002).

Another concern raised by international comparative education researchers involves the test materials. PISA tests 15-year-old students on knowledge and skills essential in the adult life, but designing a test that is relevant across the countries can be a challenge. Across the countries, cultures and values vary widely, thus relevant and essential skills and knowledge may also vary widely. Likewise, researchers agree that designing “the content that is common among participating countries is challenging” and that there is an “inherent tension between depth and coverage of topic areas” (Porter and Gamoran 2002, 9). The inevitable reliance on multiple-choice format of testing is also another challenge faced by the designers as the format limits the assessment of students’ higher order skills (Porter and Gamoran 2002).

Even when sampling was conducted properly to ensure representativeness and to minimize errors, survey data has important limitations. Surveys rely on self-reporting information to questions that are subject to personal interpretations, so reliability of the answers is, at times, questioned. Particularly because PISA relies on information from the 15-year-old students across diverse cultures, the comparability of students’ understanding of the questions and the indicators derived based on these answers is being debated (Schneider 2009).

Also, even with the range of variables included in the PISA dataset, it is difficult to adequately capture the increasingly complex educational environment. For instance, a question that asks students’ participation in shadow education captures the general participation of any outside-of-school hour lessons, without detailed information about the scope, size, and structure of these shadow education lessons (Bray 2009). Likewise, it is difficult to tell what the characteristics of the shadow education are that may influence achievement outcomes in Korea and the United States.

Notwithstanding the limitations, researchers generally agree that the quality of comparative international data has improved over the past several years and that the findings of large-scale studies are “worth taking seriously” (Porter and Gamoran 2002, 8). Furthermore, researchers have made consensual agreement that "education in one country can be better understood in comparison to education in other countries" (Porter and Gamoran 2002). The importance of understanding culture and context is noted, as the understanding of cultures and contexts leads to an increased understanding of findings from the large-scale studies and also helps researchers and policy makers in applying finding to a specific country.

CONCLUSION/ POLICY IMPLICATIONS

Conclusion

Using the 2009 PISA data, this dissertation analyzed student- and school-level factors that contribute to student achievement outcomes in Korea and the United States. Some of the findings are summarized below:

1. Family socioeconomic status and background factors play an important role in explaining achievement outcomes in the U.S. and Korea, but the importance of family factors is greater for the U.S. than is for Korea.
2. In both Korea and the U.S., the proportion of variance explained by student factors is larger than the proportion of variance explained by school factors.
3. The intraclass correlation that measures how strongly schools resemble each other is greater for Korea than the U.S. because Korea has adopted the high school tracking system: when the tracking system is considered, the extent to which schools are differentiated is smaller for Korean schools than for U.S. schools.
4. There is a stronger association between family SES and student achievement in the United States than in Korea. Also, after family SES is accounted for, the extent to which U.S. schools are differentiated drops by a significant amount.

5. Korean students on average scored about 60 points higher on math than their counterparts in the U.S., and they scored about 40 points higher on reading than U.S. students, and the differences are statistically significant at the p-value<0.01 level. When family SES is considered, the difference between the mean scores between Korean and U.S. students increased.
6. After controlling for the student- and school-level variables, the Korea-US achievement disparity becomes statistically insignificant.
7. School factors and their relationships with student achievement differ considerably between Korea and the United States. In Korea, school autonomy measures have strong positive relationships with school performance, whereas in the U.S., the associations are not statistically significant. In contrast, in the U.S., school accountability measures, such as posting student achievement data and evaluating teacher performance using student achievement data, have strong positive relationships with school performance, whereas in Korea, the measures have negative relationships with school performance.
8. In both countries, parental pressure on school standards exerts strong influence on school performance. But parental pressure on achievement does not depend on family SES in Korea, whereas it does depend on family SES in the United States.
9. U.S. students who live in a two-parent household are likely to perform significantly higher than their counterparts who do not. The substantive significance of the two-parent household is less significant in Korea.

10. Shadow education in math is the single most substantive policy variable that explains the Korea-US math difference, and shadow education in language is the single most substantive policy variable that explains the Korea-US reading difference. Whether the relationship between shadow education participation and achievement is causal in Korea and/or the United States needs to be further examined. There is a possibility of self-selection, as high-achieving students are more likely to be participating in shadow education in Korea, whereas low-achieving students in the U.S. are more likely to participate in shadow education.

Policy Implications

Availability of international comparative data makes it possible for educational researchers and policy makers to assess how well their students are prepared to perform their roles in a society, compared to their counterparts in the world. Reading news articles that report that the U.S. students rank below the OECD average in major subjects, such math, science, and reading, is distressing for the U.S. citizens and policy makers. And yet, educational researchers do not seem to have a solid hypothesis on what is driving the disparities in achievement between the U.S. students and their counterparts in OECD countries. Parental influence (including family SES), school curriculum and resources, and teacher quality are among the few factors pointed out as the leading reasons for the cross-national achievement disparities between U.S. students and students from other OECD countries, such as Korea, Canada, Finland, and Taiwan.

Particularly, the differences in South Korea and U.S. school performance have received ample attention from policy makers and educational researchers. South Korea's public expenditure in education is markedly low, compared to that of the U.S., and yet, the school performance is greater in Korea than in the United States. While policy makers and educational researchers examined possible reasons for the greater efficiency in the Korean education system, their efforts were not so successful.

This dissertation hypothesizes that there are at least three reasons for the Korea and U.S. achievement differences. Firstly, the decentralized educational system furthers the influence of family SES in the United States, and results in highly differentiated schools. While schools in both Korea and the U.S. are differentiated, U.S. schools are differentiated along the line of family SES to a greater extent than in other OECD countries, such as Korea. In Korea, schools are differentiated as students are tracked into academic and vocational schools starting at the high school level; and yet, the tracking is done largely based on student achievement levels, rather than on the family SES. While the U.S. federal government has expanded the policies and programs to support students from low-income families during the past several decades, the role played by the U.S. government is still very limited. In order to raise the achievement outcomes of low-performing students from lower-SES families, the policies that aim to lessen the link between family SES and achievement should be adopted. These policies may include subsidizing low-income families for cultural and educational possessions in home.

Secondly, Korean educational success is largely attained by the role played by parents, and to a lesser extent by the school factors. The school-level variables together explain a large amount of variation in Korean students' achievement, but most substantive variables at the school-level involve the parental role played in the education system. Parental involvement in education beyond their social/economic status plays a very significant role in Korean education. Korean parents provide home environments that are more conducive to studying for their children. Furthermore, Korean parents invest in their children's education by contributing to school finance (about 47 percent of school financing comes from student fees) and also by providing their children with private education. Statistics Korea reported that 74 percent of Korean middle school students participate in shadow education, and this figure indicates that a large majority of private households spend extra money on their children's education, outside of the formal schooling. Also, parental pressure on school standards is another substantive variable at the school-level. The dissertation finds that parental pressure is associated with a 40-point increase in math performance and a 32-point increase in reading performance in Korea, after controlling for other critical variables. Hence, it is safe to conclude that family effect is as strong in Korea as it is in the United States, and stronger than other school-level factors. Also, it is important to note that in Korea, it is not the family SES effect that is deterministic, but it is the parental role played in education that makes the difference.

Finally, policy makers should be wary of benchmarking other countries' education systems without careful analysis, since what works in one country might not work in another country. This dissertation finds that comparisons between two drastically different education systems solely based on the average assessment scores can be meaningless. For instance, several school resources-related variables are found to be statistically and substantively related with school performance in Korea. The quality of school educational resources has a strong relationship with school performance. And yet, this finding does not indicate that South Korea's centrally controlled education system is more effective in the use of educational resources, because the findings also show that the schools that have greater autonomy in resource allocation and curriculum management tend to have higher math and reading scores than those that do not. This finding raises the question of how much centralization is desirable for pushing for better aligned education and more effective use of educational resources, if at all.

The finding that shows an inverse relationship between the variables and student achievement in the U.S. and Korea deserves a closer look. For instance, school accountability measures, such as posting student achievement data for public viewing and evaluating teacher performance based on student achievement data, showed negative relationship with student achievement in Korea, whereas it showed strong positive association with student achievement in the United States. Because of the cross-sectional nature of the dataset, it is not easy to tell whether the coefficient signifies a causal relationship. Rather, it is more likely that the relationship points to a plausible self-selection issue. In other words, it is possible that the high-quality schools in U.S. post student achievement data more willingly than the lower-quality schools, and not that posting achievement data causes schools to perform better.

Shadow education also has the inverse cross-national relationships. It is likely that Korean students seek extra educational opportunities regardless of their performance in their schools, to excel further; whereas it is more likely for U.S. students and parents to look for outside-of-school educational help only if they are struggling in formal schooling. Because of the lack of prior achievement scores, it is not easy to tell whether higher achievers are more likely to benefit from shadow education in Korea. Similarly, whether the low-achieving students in the U.S. benefits from the remedial lessons remains to be answered. While comparative international datasets provide researchers with a wide range of data and allow for examination of the different patterns in the educational systems, the datasets are limited in some aspects, as they are cross-sectional. Consequently, hypotheses generated from this dissertation need to be further examined using longitudinal, domestic data.

APPENDICES

A: Description of Variables

VAR NAME	SCALE	DESCRIPTION
Outcome Variables		
MATH	Plausible values	
READING	Plausible values	
Student-level Variables		
MALE	Dichotomous	Gender (female=0)
MEMOR	Scale	Students' use of memorization strategies when reading texts
CSTRAT	Scale	Students' use of control strategies when reading texts
ESCS_within	Factor scale	Index of economic, social, cultural status; composed of HISEI (highest parental occupational status), PARED (highest parental educational attainment), WEALTH, HEDRES(home educational resources), and CULTPOS(cultural possessions)
PARENTS	Dichotomous	Student lives in a two-parent household =1; other=0
MATHHRS	Continuous	Number of hours/week students spend in in-school math lessons (range between 0 and 16.67).
LANGHRS	Continuous	Number of hours/week students spend in in-school language lessons (range between 0 and 16.67).
SHADOWMATH	Continuous	Number of hours/week outside-of-school math lessons (range between 0 and 8)
SHADOWLANG	Continuous	Number of hours/week outside-of-school math lessons (range between 0 and 8)
School-level Variables		
URBANSCH	Dichotomous	Urban school (city size of 100,000 people or more)=1; non-urban school=0
ESCS_between	Factor scale	School mean of ESCS

SCHSIZE **(SSIZE=schsize/100)**	Continuous	Number of students enrolled in school
PUBLIC	Dichotomous	School sector (private=0)
GOVFUND	Percent	Percent of school funding financed by the government
STUDFEE	Percent	Percent of school financing from student fees
RESPRES	Scale	Responsibility of the school on resource allocation
RESPCURR	Scale	Responsibility of school on curriculum management
P_PRESS	Dichotomous	Some (or more) parent pressure on school academic standards (none=0)
PCCERT	Percent	Percent of ceritified teachers at school
PCQUAL	Percent	Percent of qualified teachers at school
TCHEVAL	Dichotomous	Use of student achievement data to evaluate teacher performance=1
SCHEVAL	Dichotomous	Use of student achievement data to post publicly=1
PRINCIPEVAL	Dichotomous	Use of student achievement data to evaluate principal performance=1
IRATCOMP	Ratio	Ratio of computers available for 15-yr olds compared to school size
TCSHORT	Scale	Teacher shortage
STRATIO	Ratio	Pupil-teacher ratio
SCMATEDU	Scale	Quality of school educational materials
EXCURACT	Scale	Availability of extra-curricular activities
CSIZE	Continuous	Number of students in a language class

B-1: Descriptive Statistics - Korea

VAR NAME	No. obs.	Mean/Proportion	Std. Dev.	Min.	Max.
Outcome Variables					
MATH	4,989	546.2	89.2	62.95	873.2
READING	4,989	539.3	79.2	150.84	782.78
Student-level variables					
MALE	4,989	0.52	0.50	0	1
ESCS_within	4,982	-0.15	0.82	-3.72	2.38
HISEI	4,908	50.09	13.36	16	88
PARED	4,930	13.73	2.36	3	16
HEDRES	4,980	-0.15	0.94	-4.26	0.98
CULTPOSS	4,974	0.52	0.95	-1.3	1.49
PARENTS	4,959	0.85	0.36	0	1
Approaches to Learning					
MEMOR	4,978	0.08	0.95	-3.02	2.69
CSTRAT	4,978	-0.27	1.00	-3.45	2.50
Instructional Time					
MATHHRS	4,985	3.38	0.87	2	8
LANGHRS	4,985	3.28	0.82	2	8
SHADOWMATH	4,236	3.07	2.58	0	8
SHADOWLANG	4,169	2.15	2.13	0	8
School-level variables					
School demographics/ External influence variables					
URBANSCH	4,989	0.86	0.35	0	1
SCHSIZE	4,989	1156.87	438.81	83	2181
PUBLIC	4,849	0.83	0.38	0	1

GOVFUND	4,744	47.62	24.54	0	100
STUDFEE	4,744	47.47	24.87	0	100
RESPRES	4,989	-0.44	0.75	-0.84	2.45
RESPCURR	4,989	0.79	0.78	-0.91	1.36
P_PRESS	4,954	0.12	0.33	0	1
Teacher-related variables					
PCCERT	4,989	97.84	10.77	6.1	100
PCQUAL	4,815	98.23	8.59	8.7	100
TCHEVAL	4,956	0.77	0.42	0	1
Accountability measures					
SCHEVAL	4,922	0.33	0.47	0	1
PRINCIPEVAL	4,853	0.28	0.45	0	1
School resources variables					
IRATCOMP	4,679	0.43	0.43	0.6	2.12
TCSHORT	4,989	-0.02	0.92	-1.02	3.34
STRATIO	4,954	17.21	3.55	4.58	27.61
SCMATEDU	4,989	0.06	0.81	-1.35	1.93
EXCURACT	4,989	1.01	0.82	-1.28	2.95
CSIZE	4,986	35.89	5.11	14	51

B-2: Descriptive Statistics – USA

VAR NAME	No. obs.	Mean/Proportion	Std. Dev.	Min.	Max.
Outcome Variables					
MATH	5,233	487.4	90.8	139.2	799.8
READING	5,233	499.8	96.6	105.87	855.07
Student-level variables					
MALE	5,233	0.51	0.50	0	1
ESCS	5,190	0.17	0.93	-3.42	2.89
HISEI	4,989	52.03	16.11	16	88
PARED	5,145	13.65	2.55	3	16
HEDRES	5,172	-0.26	0.86	-3.48	0.67
CULTPOSS	5,149	-0.30	0.96	-1.68	1.06
PARENTS	5,153	0.72	0.45	0	1
Approaches to Learning					
MEMOR	5,180	-0.04	1.10	-3.02	2.69
CSTRAT	5,181	-0.04	1.12	-3.45	2.50
Instructional Time					
MATHHRS	4,946	4.38	1.73	0	8
LANGHRS	4,912	4.36	1.74	0	8
SHADOWMATH	4,373	0.58	1.51	0	8
SHADOWLANG	4,374	0.42	1.34	0	8
School-level variables					
School demographics/ External influence variables					
URBANSCH	5,233	0.39	0.49	0	1
SCHSIZE	5,002	1369.13	848.70	100	6694
PUBLIC	5,121	0.93	0.25	0	1

GOVFUND	4,990	90.40	26.64	0	100
STUDFEE	4,295	8.81	24.42	0	98
RESPRES	5,233	0.40	0.92	-0.75	-1.37
RESPCURR	5,233	-0.20	0.93	-1.37	1.36
P_PRESS	5,194	0.34	0.47	0	1
Teacher-related variables					
PCCERT	4,653	95.16	14.19	17.6	100
PCQUAL	4,834	96.31	16.03	0	100
TCHEVAL	5,163	0.81	0.39	0	1
Accountability measures					
SCHEVAL	5,233	0.89	0.31	0	1
PRINCIPEVAL	5,230	0.62	0.48	0	1
School resources variables					
IRATCOMP	4,961	0.73	0.48	0.02	2.5
TCSHORT	5,210	-0.45	0.81	-1.02	3.34
STRATIO	4,710	16.05	5.27	3.54	40.36
SCMATEDU	5,233	0.51	1.04	-1.99	1.93
EXCURACT	5,232	1.02	0.78	-1.56	2.95
CSIZE	4,870	24.48	6.73	1	50

B-3: Descriptive Statistics- OECD

VAR NAME	No. obs.	Mean/Proportion	Std. Dev.	Min.	Max.
Outcome Variables					
MATH	298,454	488.4	97.2	3.67	953.27
READING	298,454	491.5	97.5	6.65	904.5
Student-level Variables					
MALE	298,452	0.51	0.50	0	1
MEMOR	283,450	-0.04	1.03	-3.02	2.69
CSTRAT	283,336	-0.03	1.05	-3.45	2.5
ESCS	294,641	-0.15	1.08	-6.04	3.53
HISEI	285,389	48.63	16.47	16	90
PARED	289,473	12.99	3.22	3	18
HEDRES	295,356	-0.18	1.03	-4.60	2.13
CULTPOSS	293,571	-0.11	0.99	-1.92	1.82
PARENTS	278,962	0.81	0.39	0	1
MATHHRS	273,686	3.84	1.57	0	16.67
LANGHRS	273,360	3.89	1.58	0	16.67
SHADOWMATH	217,686	1.12	1.96	0	8
SHADOWLANG	215,229	0.74	1.61	0	8
School-level Variables					
URBANSCH	289,483	0.44	0.50	0	1
ESCS_between	294,641	-0.12	0.36	-3.87	0.93
SCHSIZE	284,654	971.81	776.84	2	11268
PUBLIC	284,785	0.92	0.27	0	1

GOVFUND	271,098	78.97	31.64	0	100
STUDFEE	240,538	17.63	28.99	0	100
RESPRES	290,036	-0.04	0.96	-0.84	2.45
RESPCURR	290,036	-0.07	1.01	-1.37	1.36
P_PRESS	286,540	0.23	0.42	0	1
PCCERT	248,273	83.20	33.81	0	100
PCQUAL	252,703	86.60	29.70	0	100
TCHEVAL	264,436	0.71	0.45	0	1
SCHEVAL	287,764	0.49	0.50	0	1
PRINCIPEVAL	285,739	0.43	0.50	0	1
IRATCOMP	272,753	0.55	0.44	0	2.5
TCSHORT	288,599	-0.03	1.09	-1.02	3.34
STRATIO	268,570	16.12	11.33	0.33	723
SCMATEDU	289,020	0.07	1.12	-3.39	1.93
EXCURACT	288,632	0.46	0.98	-3.29	2.95
CSIZE	280,000	27.19	8.98	1	90

C-1: Null Model– for Academic versus Vocational High Schools

MATHEMATICS					
	KOR	Academic HS	Vocational HS	Middle School	USA
Intercept	545.1** (0.60)	571** (0.58)	480.6** (1.37)	507.2** (4.03)	484.8** (0.71)
School-level SD ($\sqrt{\psi}$)	54.7	39	30.8	50.3	49.6
Student-level SD ($\sqrt{\theta}$)	70	68.1	70.3	96	76.4
ρ	0.38	0.25	0.16	0.22	0.30
N	4,989	3,567	1,105	317	5,233

READING					
	KOR	Academic HS	Vocational HS	Middle School	USA
Intercept	538.7** (0.70)	560.6** (0.70)	482.7** (1.41)	508.8** (3.02)	497.1** (0.78)
School-level SD ($\sqrt{\psi}$)	45.3	28.4	33.6	41.3	47.3
Student-level SD ($\sqrt{\theta}$)	64.5	61.6	67.6	89.3	84
ρ	0.33	0.18	0.20	0.18	0.24
N	4,989	3,567	1,105	317	5,233

C-2: SES-Adjusted Model– Academic versus Vocational Schools

	MATHEMATICS			READING		
	KOR	**Academic HS**	**Vocational HS**	**KOR**	**Academic HS**	**Vocational HS**
Intercept	547.8** (0.59)	570.7** (0.59)	487.7** (3.11)	540.9** (0.72)	560.4** (0.71)	487.6** (2.60)
Family SES	17.3** (1.57)	17.7** (1.67)	10.4** (4.19)	14.3** (1.43)	14.6** (1.52)	7.1** (3.34)
School-level SD ($\sqrt{\boldsymbol{\psi}}$)	49	36	29	41	26	32
Student-level SD ($\sqrt{\boldsymbol{\theta}}$)	69	67	70	64	61	67
ρ	0.34	0.22	0.14	0.29	0.15	0.19
R^2	0.09	0.06	0.03	0.07	0.04	0.03
N	4,989	3,567	1,105	4,989	3,567	1,105

C-3: Full Model– Academic versus Vocational Schools

	MATHEMATICS			READING		
	KOR	**Academic HS**	**Vocational HS**	**KOR**	**Academic HS**	**Vocational HS**
Intercept	486.3** (21.96)	462.1** (32.49)	449.9** (44.47)	526.6** (22.02)	507.4** (21.00)	567.9** (47.82)
MALE	12.4** (3.43)	17.0** (3.65)	8.5 (7.67)	-29.2** (2.92)	-24.0** (2.88)	-34.8** (5.44)
ESCS	9.4** (2.35)	12.8** (3.38)	0.1 (3.67)	9.4** (1.76)	12.6** (2.39)	-0.9 (3.08)
PARENTS	1.4 (3.45)	1.3 (4.04)	3.1 (6.14)	-3.3 (3.10)	-3.9 (3.91)	-1.1 (5.32)
MATH (/LANG-) HRS	5.6** (1.52)	2.1 (1.18)	21.6** (7.28)	0.3 (1.36)	-1.5 (1.15)	34.9** (5.26)
SHADOW-(MATH/LANG)	3.9** (0.67)	3.3** (0.63)	5.4* (2.46)	1.0 (0.65)	0.5 (0.61)	3.5 (2.97)
MEMOR	-6.7** (1.54)	-8.6 (1.63)	-3.4 (2.90)	1.6 (1.30)	0.3 (1.32)	3.7 (2.48)
CSTRAT	23.0** (1.74)	22.8** (1.93)	20.9** (4.02)	20.5** (1.41)	19.9** (1.48)	18.5** (3.16)
URBANSCH	-11.7** (2.79)	6.8* (3.31)	2.1 (9.35)	-18.1** (2.43)	-5.9 (3.39)	18.0 (10.76)
SSIZE	3.6** (0.35)	2.7** (0.35)	-0.3 (1.03)	2.4** (0.27)	1.2** (0.28)	-1.5 (1.75)
PUBLIC	4.3** (2.60)	-0.9 (2.95)	-2.1 (10.39)	-3.6 (2.51)	-7.1** (2.22)	8.0 (12.30)
GOVFUND	-0.5 (0.03)	-0.1* (0.04)	0.0 (0.18)	-0.4** (0.03)	-0.1* (0.04)	0.0 (0.15)
RESPRES	15.8** (0.73)	8.5** (1.03)	28.0** (7.21)	10.5** (0.91)	3.7** (1.11)	30.5** (6.28)
RESPCURR	3.4** (1.25)	4.6** (1.09)	-4.7 (3.67)	3.7** (0.98)	6.0** (1.11)	-9.6** (3.23)
P_PRESS	39.6** (3.22)	24.2** (2.87)	-32.3* (13.22)	31.8** (2.40)	17.2** (2.71)	-35.7** (9.73)
PCCERT	0.1 (0.12)	-0.0 (0.27)	0.0 (0.22)	0.2** (0.06)	0.1 (0.16)	1.1** (0.24)

PCQUAL	0.1 (0.11)	0.3 (0.22)	-1.0 (0.63)	0.1 (0.10)	0.2 (0.14)	-3.3** (0.68)
TCHEVAL	0.8 (1.44)	-5.3 (2.96)	-20.0* (8.38)	-0.2 (1.45)	-6.3** (1.87)	-24.0** (6.97)
SCHEVAL	-4.6 (2.17)	-1.3 (2.40)	18.0* (8.29)	-6.1** (1.55)	-2.2 (2.08)	8.2 (8.16)
PRINCIPEVAL	7.6** (1.86)	17.9** (2.14)	-6.9 (9.17)	4.2* (2.07)	9.6** (2.19)	-17.7** (6.32)
IRATCOMP	-20.6** (3.40)	47** (6.32)	8.1 (9.51)	-17.0** (2.55)	37.2** (4.48)	-3.5 (7.22)
TCSHORT	-0.6 (0.91)	-0.5 (1.20)	12.6** (4.66)	-1.8* (0.73)	-1.9* (0.92)	12.6* (5.62)
STRATIO	-1.6** (0.43)	1.3 (0.72)	-0.4 (1.71)	-1.1** (0.37)	1.2* (0.60)	0.8 (1.85)
SCMATEDU	4.8** (1.04)	2.5 (1.42)	0.2 (6.06)	2.8** (0.76)	-2.7* (1.14)	9.6 (5.93)
EXCURACT	-0.8 (1.01)	4.1** (0.76)	4.3 (5.15)	1.1 (1.26)	3.8** (1.04)	13.0** (3.70)
CSIZE	0.9 (0.70)	-0.5 (0.76)	2.4 (1.33)	1.2 (0.57)	0.2 (0.52)	1.8 (1.04)
ESCS* SHADOW	0.4 (0.67)	-0.2 (0.84)	2.7 (1.65)	-0.4 (0.60)	-1.0 (0.69)	2.0 (1.77)
ESCS* P_PRESS	0.7 (4.82)	-0.5 (4.31)	-9.2 (11.86)	0.7 (4.07)	-0.6 (3.88)	-14.5 (7.59)
School-level SD ($\sqrt{\boldsymbol{\psi}}$)	32	25	14	25	17	4
Student-level SD ($\sqrt{\boldsymbol{\theta}}$)	66	64	66	60	58	61
ρ	0.19	0.13	0.04	0.15	0.08	0.01
N	4,989	3,567	1,105	4,989	3,567	1,105

D: HLM Results without ‘shadow education’ variables, by country

	MATHEMATICS		READING	
	KOR	**USA**	**KOR**	**USA**
Intercept	484.8** (22.10)	486.4** (11.78)	527.7** (21.99)	527.6** (12.92)
MALE	12.8** (3.53)	19.1** (2.27)	-29.2** (2.95)	-25.7** (2.18)
ESCS	11.7** (1.51)	17.0** (1.71)	8.7** (1.36)	17.6** (1.69)
PARENTS	1.9 (3.60)	19.7** (2.47)	-3.2 (3.16)	23.0** (2.67)
MATH(/LANG)-HRS	6.3** (1.47)	5.8** (0.87)	0.4 (1.36)	5.5** (0.89)
SHADOW-(MATH/LANG)	-	-	-	-
MEMOR	-6.9** (1.55)	-20.5** (1.47)	1.6 (1.30)	-21.1** (1.59)
CSTRAT	24.3** (1.82)	21.4** (1.21)	20.7** (1.47)	26.1** (1.50)
URBANSCH	-12.0** (2.81)	-20.8** (2.05)	-18.3** (2.44)	-15.2** (2.23)
SSIZE	3.6** (0.36)	0.7** (0.16)	2.4** (0.27)	0.7** (0.14)
PUBLIC	5.1** (2.65)	-14.8** (7.31)	-3.4 (2.55)	-57.5** (9.25)
GOVFUND	-0.5 (0.03)	-0.2 (0.05)	-0.4** (0.03)	0.2** (0.06)
RESPRES	16.4** (0.72)	-1.5 (1.13)	10.6** (0.91)	-0.5 (1.11)
RESPCURR	3.3** (1.26)	-1.2 (1.13)	3.6** (1.00)	0.9 (1.24)
P_PRESS	40.1** (3.23)	15.4** (2.46)	32.0** (2.39)	13.1** (2.66)
PCCERT	0.1 (0.12)	-0.6** (0.11)	0.2** (0.06)	-0.4** (0.12)
PCQUAL	0.2 (0.12)	0.1* (0.04)	0.1 (0.10)	-0.0 (0.04)

TCHEVAL	1.1 (1.43)	7.4** (2.59)	0.0 (1.41)	1.7 (3.03)
SCHEVAL	-4.8 (2.17)	25.0** (3.50)	-6.3** (1.53)	29.5** (3.78)
PRINCIPEVAL	7.8** (1.86)	-15.3** (2.56)	4.4* (2.05)	-11.9** (2.83)
IRATCOMP	-22.7** (3.14)	0.9 (2.38)	-17.8** (2.39)	-5.9** (2.17)
TCSHORT	-0.4 (0.91)	-6.9** (0.99)	-1.9* (0.73)	-6.7** (1.10)
STRATIO	-1.4** (0.43)	-1.7** (0.18)	-1.1** (0.37)	-1.9** (0.19)
SCMATEDU	5.5** (1.06)	0.3 (0.88)	2.9** (0.76)	-0.2 (0.99)
EXCURACT	-0.3 (1.01)	7.1** (1.38)	1.1 (1.27)	6.5** (1.44)
CSIZE	1.0 (0.71)	0.6** (0.22)	1.2 (0.58)	0.7** (0.22)
ESCS* P_PRESS	0.7 (4.35)	11.6** (2.71)	0.7 (4.05)	9.3** (3.23)
Random Effects				
School-level SD($\sqrt{\psi}$)	**33.1**	**25.1**	**24.8**	**22**
Student-level SD ($\sqrt{\theta}$)	**66.3**	**69.8**	**59.7**	**76.1**
ρ	**0.20**	**0.11**	**0.15**	**0.08**
R^2	**0.30**	**0.34**	**0.33**	**0.32**
*Note: ** indicates p-value <0.01, * indicates p-value <0.05*				

E: Results for Korea and U.S. achievement differences in math and reading

	MATH		READING	
	NULL (1)	MODEL (2)	NULL (1)	MODEL (2)
INTERCEPT	485.4** (1.55)	455.1** (22.25)	497.5** (1.41)	504.4** (16.17)
KOR-US difference	59.6** (4.84)	-54.5 (29.37)	40.2** (4.25)	-55.6 (31.59)
Student-level Variables				
MALE		18.4** (2.03)		-26.3** (1.95)
ESCS		20.4** (1.66)		19.6** (1.58)
KOR-US difference		-2.3 (2.75)		-4.1 (2.27)
PARENTS		19.5** (2.62)		21.9** (2.77)
KOR-US difference		-12.3** (4.72)		-20.1** (4.49)
MATHHRS (/LANGHRS)		5.5** (0.88)		5.2** (0.87)
KOR-US difference		0.3 (4.95)		-3.8 (4.39)
SHADOWMATH (/SHADOWLANG)		-8.8** (0.75)		-13.4** (0.97)
KOR-US difference		14.3** (1.24)		15.2** (1.35)
MEMOR		-18.0** (1.19)		-16.5** (1.35)
CSTRAT		21.8** (1.08)		25.1** (1.35)
School-level Variables				
URBANSCH		-16.6** (4.46)		-11.7** (3.74)
SSIZE		0.9** (0.27)		0.8** (0.23)
PUBLIC		-14.4 (12.65)		-35.0* (11.07)
KOR-US difference		25.3		40.4*

	(14.77)	(12.71)
GOVFUND	-0.3** (0.10)	-0.2* (0.08)
RESPRES	1.6 (2.72)	0.9 (2.32)
KOR-US difference	15.3** (5.27)	12.3** (4.96)
RESPCURR	-4.1 (2.70)	-0.7 (1.99)
KOR-US difference	8.9 (4.75)	4.2 (3.84)
P_PRESS	22.7** (4.88)	14.0** (4.05)
PCCERT	-0.3 (0.19)	-0.2 (0.15)
PCQUAL	0.3* (0.14)	0.2* (0.09)
TCHEVAL	-2.3 (5.29)	-2.7 (4.04)
KOR-US difference	3.1 (9.54)	6.5 (8.11)
SCHEVAL	22.2** (6.41)	20.0** (4.94)
KOR-US difference	-28.7** (8.54)	-29.0** (7.31)
PRINCIPEVAL	-5.5 (4.71)	-4.7 (3.42)
IRATCOMP	0.6 (5.10)	-3.7 (4.04)
TCSHORT	-3.4 (1.94)	-4.7** (1.59)
STRATIO	-1.2** (0.32)	-1.2** (0.28)
SCMATEDU	1.9 (2.06)	1.7 (1.60)
KOR-US difference	0.4 (4.53)	-0.4 (3.75)
EXCURACT	2.2 (2.21)	3.5* (1.70)
CSIZE	0.4 (0.23)	0.6* (0.22)
KOR-US difference	3.2**	3.1**

		(0.87)		(0.78)
Random Effects				
School-level SD ($\sqrt{\boldsymbol{\psi}}$)	43.3	23.4	41.9	20.7
Student-level SD ($\sqrt{\boldsymbol{\theta}}$)	79.4	70	83.9	73.3
ρ	0.23	0.10	0.20	0.07
R^2	-	0.33	-	0.34
*Note: ** indicates p-value <0.01, * indicates p-value <0.05*				

REFERENCES

Anderson, Lorin. 2004. “Increasing Teacher Effectiveness”. UNESCO International Institute for Educational Planning.

Armor, David J. 2003. *Maximizing Intelligence*. Transaction Publishers.

———. 2007. “Can NCLB Close Achievement Gaps?” In *No Child Left Behind and the Reduction of the Achievement Gap*, edited by Alan Sadovnik, Jennifer O’Day, George Bohrnstedt, and Kathryn Borman. Routledge.

Augenblick, John G., John L. Myers, and Amy B. Anderson. 1997. “Equity and Adequacy in School Funding.” *The Future of Children* 7 (3): 63–78.

Baker, David P., Motoko Akiba, Gerald K. LeTendre, and Alexander W. Wiseman. 2001. “Worldwide Shadow Education: Outside-School Learning, Institutional Quality of Schooling, and Cross-National Mathematics Achievement.” *Educational Evaluation and Policy Analysis* 23 (1): 1–17.

Baker, David P., and Gerald K. LeTendre. 2005. *National Differences, Global Similarities: World Culture and the Future of Schooling*. Stanford University Press.

Baker, David P., Brian Goesling, and Gerald K. LeTendre. 2002. “Socioeconomic Status, School Quality, and National Economic Development: A Cross-National Analysis of the ‘Heyneman-Loxley Effect’ on Mathematics and Science Achievement.” *Comparative Education Review* 46 (3): 291–312.

Becker, Gary S. 1991. *A Treatise on the Family, Enlarged Edition*. Cambridge, MA: Harvard University Press.

Bergeron, Deborah L. 2010. “The Impact of Attendance in Supplemental Educational Services on Students’ Academic Achievement”. PhD diss., George Mason University.

Berliner, D. C. 1990. "What's All This Fuss About Instructional Time?" In *The Nature of Time in Schools*, edited by M. Ben-Peretz and B. Rainer, 3–35. New York: Teachers College Press.

Blau, Peter M., and Otis D. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Boyd, William, and Debra Miretzky, ed. 2003. "American Educational Governance on Trial: Change and Challenges." In The National Society For The Study Of Education.

Bray, Mark. 1999. "The Shadow Education System: Private Tutoring and Its Implications or Planners. Fundamentals of Educational Planning". UNESCO International Institute for Educational Planning. http://www.ingentaconnect.com/content/klu/revi/2000/00000046/00000006/00276636.

———. 2006. "Private Supplementary Tutoring: Comparative Perspectives on Patterns and Implications." *Compare: A Journal of Comparative Education* 36 (4): 515–530.

———. 2009. *Confronting the Shadow Education System : What Government Policies for What Private Tutoring?* Paris: United Nations Educational, Scientific and Cultural Organization : International Institute for Educational Planning.

Brown, Byron W., and Daniel H. Saks. 1986. "Measuring the Effects of Instructional Time on Student Learning: Evidence from the Beginning Teacher Evaluation Study." *American Journal of Education* 94 (4): 480–500.

Bryk, Anthony, Valerie Lee, and Peter Holland. 1993. *Catholic Schools and the Common Good*. Harvard University Press.

Bryk, Anthony, and Stephen Raudenbush. 1988. "Toward a More Appropriate Conceptualization of Research on School Effects: A Three-Level Hierarchical Linear Model." *American Journal of Education* 97 (1): 65–108.

———. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Advanced Quantitative Techniques in the Social Sciences 1. Newbury Park: Sage Publications.

Buchmann, Claudia. 2002. “Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges.” In *Methdological Advances in Cross-National Surveys of Educational Achievement*, 150–197. Washington, DC: National Academy Press.

Buchmann, Claudia, Dennis J. Condron, and Vincent J. Roscigno. 2010. “Shadow Education, American Style: Test Preparation, the SAT and College Enrollment.” *Social Forces* 89 (2): 435–461. doi:10.1353/sof.2010.0105.

Byun, Soo-yong, and Kyung-keun Kim. 2010. “Educational Inequality in South Korea: The Widening Socioeconomic Gap in Student Achievement.” *Research in Sociology of Education* 17: 155–182.

Byun, Soo-yong, and Hyunjoon Park. 2012. “The Academic Success of East Asian American Youth The Role of Shadow Education.” *Sociology of Education* 85 (1): 40–60.

Carroll, John. 1963. “A Model of School Learning.” *Teachers College Record* 64 (8): 723–733.

Chubb, John E., and Terry M. Moe. 1990. *Politics, Markets, and America’s Schools*. Washington, D.C.: Brookings Institution Press.

Chung, Bong Gun. 2002. “Korea’s War on Private Tutoring” presented at the Second International Forum on Education Reform. http://www.worldedreform.com/intercon2/f20.pdf.

Coleman, James, Earnest Campbell, Carol Hobson, James McPartland, Alexander Mood, Frederick Weinfeld, and Robert York. 1966. *Equality of Educational Opportunity*. Washington, D.C.: U.S. Dept. of Health, Education, and Welfare, Office of Education.

Coleman, James, Thomas Hoffer, and Sally Kilgore. 1982. “Cognitive Outcomes in Public and Private Schools.” *Sociology of Education* 55 (2): 65–76.

Corey, Douglas Lyman, Geoffrey Phelps, Deborah Loewenberg Ball, Jenny Demonte, and Delena Harrison. 2012. “Explaining Variation in Instructional Time An Application of Quantile Regression.” *Educational Evaluation and Policy Analysis* 34 (2): 146–163.

Cross, Christopher T. 2004. *Political Education: National Policy Comes of Age*. New York: Teachers College Press.

Dang, Hai-Anh. 2007. "The Determinants and Impact of Private Tutoring Classes in Vietnam." *Economics of Education Review* 26 (6): 683–698.

Darling-Hammond, Linda. 2000. "Teacher Quality and Student Achievement." *Education Policy Analysis Archives* 8 (1): 1–44.

Darling-Hammond, Linda, Ruth Chung Wei, and Christy Marie Johnson. 2009. "Teacher Preparation and Teacher Learning." In *Handbook of Education Policy Research*, edited by Gary Sykes, Barbara Schneider, and David Plank, 613–636. Routledge.

Desimone, Laura, and Daniel Long. 2010. "Teacher Effects and the Achievement Gap: Do Teacher and Teaching Quality Influence the Achievement Gap Between Black and White and HIgh- and Low- SES Students in the Early Grades?" *Teachers College Record* 112 (12): 3024–3073.

Duncan, Otis D. 1961. "A Socioeconomic Index for All Occupations." In *Occupation and Social Status*, edited by A. J. Reiss, 109–161. New York: Free Press.

Earley, Penelope M. 2000. "Finding the Culprit: Federal Policy and Teacher Education." *Educational Policy* 14 (1): 25–39.

Education Law Center. 2012. "Is School Funding Fair? A National Report Card". Rutgers Graduate School of Education. http://www.schoolfundingfairness.org/index.htm.

Farkas, George, and Rachel Durham. 2007. "The Role of Tutoring in Standards-Based Reform." In *Standards-Based Reform and the Poverty Gap: Lessons for No Child Left Behind*. Brookings Institution Press.

Fetler, Mark. 1999. "High School Staff Characteristics and Mathematics Test Results." *Education Policy Analysis Archives* 7 (9): 1–23.

Finn, Jeremy D., and Charles M. Achilles. 1999. "Tennessee's Class Size Study: Findings, Implications, Misconceptions." *Educational Evaluation and Policy Analysis* 21 (2): 97–109.

Floden, Robert. 2010. “Policy on Teacher Preparation: The Demand for More Research.” *New Educator*. http://edwp.educ.msu.edu/new-educator/2011/policy-on-teacher-preparation-the-demand-for-more-research/.

Fuller, Bruce, and Prema Clarke. 1994. “Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules, and Pedagogy.” *Review of Educational Research* 64 (1): 119–157.

Gershoff, Elizabeth T., C. Cybele Raver, J. Lawrence Aber, and Mary Clare Lennon. 2007. “Income Is Not Enough: Incorporating Material Hardship Into Models of Income Associations With Parenting and Child Development.” *Child Development* 78 (1): 70–95.

Gordon, Bryon, and David J. Armor. 2004. “The Effects of Accountability Systems on the Achievement Gap: 2000-2003”. Paper presented at the APPAM Research Conference, October 26, Atlanta, GA.

Greenwald, Rob, Larry Hedges, and Richard Laine. 1996. “The Effect of School Resources on Student Achievement.” *Review of Educational Research* 66 (3): 361–396.

Guthrie, James W. 1997. “School Finance: Fifty Years of Expansion.” *The Future of Children* 7 (3): 24–38.

Hanushek, Eric A. 1986. “The Economics of Schooling: Production and Efficiency in Public Schools.” *Journal of Economic Literature* 24 (3): 1141–1177.

———. 1989. “The Impact of Differential Expenditures on School Performance.” *Educational Researcher* 18 (4): 45–62.

———. 1996. “A More Complete Picture of School Resource Policies.” *Review of Educational Research* 66 (3): 397–409.

———. 1999. “Some Findings From an Independent Investigation of the Tennessee STAR Experiment and From Other Investigations of Class Size Effects.” *Educational Evaluation and Policy Analysis* 21 (2): 143–163.

Hanushek, Eric A., and Margaret E. Raymond. 2004. “The Effect of School Accountability Systems on the Level and Distribution of Student Achievement.” *Journal of the European Economic Association* 2 (2/3): 406–415.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996. "Aggregation and the Estimated Effects of School Resources." *The Review of Economics and Statistics* 78 (4): 611–627.

Hauser, Robert M., and William Sewell. 1986. "Family Effects in Simple Models of Education, Occupational Status, and Earnings: Findings from the Wisconsin and Kalamazoo Studies." *Journal of Labor Economics* 4 (3): S83–S115.

Hedges, Larry, and Rob Greenwald. 1996. "Have Times Changed? The Relation Between School Resources and Student Performance." In *Does Money Matter?: The Effect of School Resources on Student Performance and Adult Success*, edited by Gary T. Burtless. Washington, D.C.: Brookings Institution Press.

Herrnstein, Richard J., and Charles Murray. 1996. *Bell Curve: Intelligence and Class Structure in American Life*. 1st ed. Free Press.

Hess, Frederick. 2003. "Breaking the Mold: Charter Schools, Contract Schools, and Voucher Plans." In *American Educational Governance on Trial: Change and Challenges*, edited by William Boyd and Debra Miretzky. The National Society For The Study Of Education.

Heyneman, Stephen. 1976. "Influences on Academic Achievement: A Comparison of Results from Uganda and More Industrialized Societies." *Sociology of Education* 49 (3): 200–211.

Heyneman, Stephen, and William Loxley. 1982. "Influences on Academic Achievement Across High and Low Income Countries: A Re-Analysis of IEA Data." *Sociology of Education* 55 (1): 13–21.

———. 1983. "The Effect of Primary-School Quality on Academic Achievement Across Twenty-nine High- and Low-Income Countries." *American Journal of Sociology* 88 (6): 1162–1194.

Hill, Paul. 2003. "What's Wrong with Public Education Governance in Big Cities...and How Should It Be Fixed?" In *American Educational Governance on Trial: Change and Challenges*, edited by William Boyd and Debra Miretzky. The National Society For The Study Of Education.

Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephanie Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Baskc Books, Inc.

Kim, Ee-gyeong, Jae-woong Kim, and You-kyung Han. 2009. "Secondary Teacher Policy Research in Asia: Secondary Education and Teacher Quality in the Republic of Korea". UNESCO, Bangkok. http://doc.iiep.unesco.org/cgi-bin/wwwi32.exe/%5Bin=epidoc1.in%5D/?t2000=028067/%28100%29.

Kim, Junyeop, and Hye Sook Shin. 2012. "Planning Special Purpose High Schools and Private Tutoring Needs in Korea: A Longitudinal, Causal Relationship." *KEDI Journal of Educational Policy* 9 (1): 29–49.

Kim, Sunwoong, and Ju-Ho Lee. 2010. "Private Tutoring and Demand for Education in South Korea." *Economic Development and Cultural Change* 58 (2): 259–296.

Konstantopoulos, Spyros, and Geoffrey Borman. 2011. "Family Background and School Effects on Student Achievement: A Multilevel Analysis of the Coleman Data." *Teachers College Record* 113 (1): 97–132.

Korea Policy Portal. 2007. "Forty Years of Korean Education". Hans Media, Seoul.

Lauglo, Jon, and Martin McLean, ed. 1985. *Control of Education: International Perspectives on the Centralization/Decentralization Debate*. Heinemann Educational Publishers.

Lee, Jaekyung. 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". Harvard Education Publishing Group.

———. 2007. "Two Worlds of Private Tutoring: The Prevalence and Causes of After-SChool Mathematics Tutoring in Korea and the United States." *Teachers College Record* 109 (5): 1207–1234.

Lee, Jisoon. 2001. "Education Policy in the Republic of Korea: Building Block or Stumbling Block?" The International Bank for Reconstruction and Development and the World Bank.

Lee, Jong Tae, Yang Boon Kim, and Cho Hee Yoon. 2004. "The Effects of Pre-Class Tutoring on Student Achievement: Challenges and Implications for Public Education in Korea." *KEDI Journal of Educational Policy* 1 (1): 25–42.

Lee, Ju-Ho. 2012. *Positive Changes - The Education Science & Technology Polices of Korea*. Seoul, Korea: Korea Economic Daily.

Lockheed, Marlaine E., and N. T. Longford. 1991. "School Effects on Mathematics Achievement Gain in Thailand." In *Schools, Classrooms and Pupils: International Studies of Schooling from a Multi-level Perspective*, edited by Stephen Raudenbush and J. Douglas Willms, 131–148. San Diego, CA: Academic Press Inc.

Loveless, Tom. 2002. *The Great Curriculum Debate: How Should We Teach Reading and Math?* Washington, D.C.: Brookings Institution Press.

Lubienski, Sara, Christopher Lubienski, and Corrina Crane. 2008. "Achievement Differences and School Type: The Role of School Climate, Teacher Certification, and Instruction." *American Journal of Education* 115 (1): 97–138.

Lubienski, Sarah Theule, and Christopher Lubienski. 2006. "School Sector and Academic Achievement: A Multilevel Analysis of NAEP Mathematics Data." *American Educational Research Journal* 43 (4): 651–698.

Manna, Paul. 2011. *Collision Course: Federal Education Policy Meets State and Local Realities*. Washington, D.C.: CQ Press.

McLanahan, Sara, and Gary Sandefur. 1994. *Growing Up With a Single Parent: What Hurts, What Helps*. Harvard University Press.

McLoyd, Vonnie C. 1998. "Socioeconomic Disadvantage and Child Development." *American Psychologist* 53 (2): 185–204.

MEST. 2013a. "Introduction: Education System." *Ministry of Education, Science, and Technology*. Accessed March 21. http://english.mest.go.kr/web/1692/site/contents/en/en_0203.jsp.

———. 2013b. "Key Policies: Major Tasks in Elementary and Secondary Education." *Ministry of Education, Science, and Technology*. Accessed March 21. http://english.mest.go.kr/web/1707/site/contents/en/en_0275.jsp.

Moser, Michele, and Ross Rubenstein. 2002. “The Equality of Public School District Funding in the United States: A National Status Report.” *Public Administration Review* 62 (1): 63–72.

Munoz, M., A. Potter, and S. Ross. 2008. “Supplemental Educational Services as a Consequence of the NCLB Legislation: Evaluating Its Impact on Student Achievement in a Large Urban District.” *Journal of Education for Students Placed at Risk* 13 (1): 1–25.

Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner City Children*. Cambridge, MA: Ballinger.

National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform*. Washington, D.C.: U.S. Government Printing Office.

OECD. 2009a. *PISA Data Analysis Manual: SPSS.* 2nd ed. Paris: OECD. http://dx.doi.org/10.1787/9789264056275-en.

———. 2009b. “Multilevel Analyses.” In *PISA Data Analysis Manual: SAS*, 2nd ed. OECD Publishing. http://dx.doi.org/10.1787/9789264056251-16-en.

———. 2010. *PISA 2009 Assessment Framework*. Paris: OECD Publishing.

———. 2012a. “Education at a Glance 2012: OECD Indicators”. OECD Publishing. http://dx.doi.org/10.1787/eag-2012-en.

———. 2012b. “Scaling Procedures and Construct Validation of Context Questionnaire Data.” In *PISA 2009 Technical Report*, by OECD, 279–315. OECD Publishing. http://www.oecd-ilibrary.org.mutex.gmu.edu/education/pisa-2009-technical-report/scaling-procedures-and-construct-validation-of-context-questionnaire-data_9789264167872-17-en.

———. 2013. “About PISA.” *OECD Program for International Student Assessment*. Accessed March 21. http://www.oecd.org/pisa/aboutpisa/.

Park, Hyun-Jeong, Jongim Byun, and Soonok Jo. 2012. “Do After-school Programs Matter? A Longitudinal Study on the Effectiveness of Participating in After-school Programs.” *KEDI Journal of Educational Policy* 9 (1): 3–27.

Park, Hyunjoon. 2005. “Cross-national Variation in the Effects of Family Background and Schools on Student Achievement: The Relevance of Institutional and Policy Contexts”. PhD diss., The University of Wisconsin - Madison.

Park, Hyunjoon, Soo-yong Byun, and Kyung-keun Kim. 2011. “Parental Involvement and Students’ Cognitive Outcomes in Korea Focusing on Private Tutoring.” *Sociology of Education* 84 (1): 3–22.

Park, Sungha. 2008. “On College-Entrance Exam Day, All of South Korea Is Put to the Test.” *Wall Street Journal*, sec. World News. http://online.wsj.com/article/SB122644964013219173.html.

Perry, Laura B., and Andrew McConney. 2010. “Does the SES of the School Matter? An Examination of Socioeconomic Status and Student Achievement Using PISA 2003.” *Teachers College Record* 112 (4): 1137–1162.

Peterson, Paul E. 2003. *Our Schools and Our Future: Are We Still at Risk?* 1st ed. Hoover Institution Press.

Porter, Andrew, and Adam Gamoran. 2002. “Progress and Challenges for Large-Scale Studies.” In *Methdological Advances in Cross-National Surveys of Educational Achievement*, edited by Andrew Porter and Adam Gamoran, 3–23. Washington, DC: National Academy Press.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2012. *Multilevel and Longitudinal Modeling Using Stata, Volumes I and II, Third Edition*. 3rd ed. Stata Press.

Raudenbush, Stephen, and Ji-soo Kim. 2002. “Statistical Issues in Analysis of International Comparisons.” In *Methdological Advances in Cross-National Surveys of Educational Achievement*, edited by Andrew Porter and Adam Gamoran, 267–294. Washington, DC: National Academy Press.

Ravitch, Diane. 2010. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books.

Ross, S., A. Potter, J. Paek, D. McKay, W. Sanders, and J. Ashton. 2008. “Implementation and Outcomes of Supplemental Educational Services: The Tennessee State-wide Evaluation Study.” *Journal of Education for Students Placed at Risk* 13 (1): 26–38.

Rothstein, Richard. 2004. *Class And Schools: Using Social, Economic, And Educational Reform To Close The Black-white Achievement Gap*. Economic Policy Institute and Teachers College.

Rowan, Brian, Richard Correnti, and Robert Miller. 2002. "What Large-scale Survey Research Tells Us About Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools". Philadelphia: Consortium for Policy Research in Education.

Schafer, Joseph L. 1999. "Multiple Imputation: a Primer." *Statistical Methods in Medical Research* 8 (1): 3–15.

Schneider, Mark. 2009. "The International PISA Test: A Risky Investment for States." *Education Next*.

Seth, Michael J. 2002. *Education Fever: Society, Politics, and the Pursuit of Schooling in South Korea*. University of Hawaii Press.

Shin, Jongho, Hyunjoo Lee, and Yongnam Kim. 2009. "Student and School Factors Affecting Mathematics Achievement International Comparisons Between Korea, Japan and the USA." *School Psychology International* 30 (5): 520–537.

Sirin, Selcuk R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research* 75 (3): 417–453.

Sousa, Sonia, Eun Jung Park, and David Armor. 2012. "Effects of Family and School Factors on Cross-national Academic Achievement Using the 2009 and 2006 PISA Surveys." *Journal of Comparative Policy Analysis* 14 (5): 1–20.

Statistics Korea. 2010. "Expenditure in Shadow Education." *Korean Statistical Information Service*. http://kosis.kr/abroad/abroad_01List.jsp.

Stevenson, David, and David P. Baker. 1992. "Shadow Education and Allocation in Formal Schooling: Transition to University in Japan." *American Journal of Sociology* 97 (6): 1639–1657.

Stillwell, R., and J. Sable. 2013. "Public School Graduates and Dropouts from the Common Core of Data: School Year 2009-10: First Look (Provisional Data)

(NCES 2013-309)". U.S. Department of Education. Washington, D.C.: National Center for Education Statistics. http://nces.ed.gov/pubsearch.

Summers, Anita A., and Barbara L. Wolfe. 1997. "Do Schools Make a Difference?" *The American Economic Review* 67 (4): 639–652.

The Economist. 2011. "Exams in South Korea: The One-shot Society", December 17. http://www.economist.com/node/21541713.

Timpane, P. Michael. 1974. "Federal Aid to Schools: Its Limited Future." *Law and Contemporary Problems* 38 (3): 493–512.

Vergari, Sandra. 2009. "Charter School Policy Issues and Research Questions." In *Handbook of Education Policy Research*, edited by Gary Sykes, Barbara Schneider, and David Plank, 478–490. Routledge.

Vinovskis, Maris. 2008. *From a Nation at Risk to No Child Left Behind: National Education Goals and the Creation of Federal Education Policy*. Teachers College Press.

Wasik, Barbara A., and Robert E. Slavin. 1993. "Preventing Early Reading Failure with One-to-One Tutoring: A Review of Five Programs." *Reading Research Quarterly* 28 (2) (April 1): 179–200. doi:10.2307/747888.

White, Karl R. 1982. "The Relation Between Socioeconomic Status and Academic Achievement." *Psychological Bulletin* 91 (3): 461–81.

Willms, J. Douglas. 2010. "School Composition and Contextual Effects on Student Outcomes." *Teachers College Record* 112 (4): 1008–1037.

Wong, Kenneth. 1991. "State Reform in Education Finance: Territorial and Social Strategies." *Publius* 21 (3): 125–142.

Wong, Raymond. 1998. "Multidimensional Influences of Family Environment in Education: The Case of Socialist Czechoslovakia." *Sociology of Education* 71 (1): 1–22.

Wood, B. Dan, and Nick A. Theobald. 2003. "Political Responsiveness and Equity in Public Education Finance." *The Journal of Politics* 65 (3): 718–738.

CURRICULUM VITAE

Eun Jung "EJ" Park received a Bachelor of Arts degree in International Affairs from the George Washington University in 2003. In 2007, she received a Master of International Studies degree from Graduate School of International Studies, Yonsei University, in Seoul. Upon graduation, she entered Georgetown Public Policy Institute of Georgetown University to pursue her degree in Master of Public Policy. In 2009, she entered School of Public Policy, George Mason University to pursue her doctoral degree in Public Policy.