KJEP 12:2 (2015), pp. 197-223

Do school accountability and autonomy affect PISA achievement? Evidence from South Korea*

Pilnam Yi
Hongik University, Korea

Abstract

School accountability and autonomy are the catchwords of global education reform movements. South Korea has also intensified school accountability and autonomy, particularly since 2008. This study examines the effects of school accountability and autonomy on Korean students' math achievement in PISA. Multilevel analyses revealed an insignificant effect of school accountability on math scores and the probability of scoring below the basic proficiency level in 2003 and 2012. In addition, a quantile regression indicated no effect of school accountability on math achievement at different achievement levels. However, school autonomy over curricula and assessment had a positive relationship with math achievement, particularly that of lower-performing students in 2012. The findings suggest that alternative accountability policies deserve consideration in the context of Korean education.

Keywords: school accountability, autonomy, PISA, student achievement, proficiency level

* This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A58019266). I wish to express my deepest gratitude to In-Soo Shin and three anonymous reviewers for their helpful comments and to Hyun-Young Jo for research assistance.

Introduction

Over the past decades, accountability in education, which combines school accountability with student assessment, has been the centerpiece of school reforms across countries. In particular, performance-based accountability policies accompanied by increased school autonomy have spread across national educational systems, and OECD PISA has played a considerable role in the policy discourse (Meyer, Tröhler, Labaree, & Hutt, 2014; Ozga, 2013). Scholars have expressed grave concerns about the impact of PISA on national education policymaking and OECD's transnational influences, often referred to as global governance of education (Meyer & Benavot, 2013). On the other hand, PISA surveys and the accumulated data have provided researchers with ample opportunities for in-depth secondary analyses. This study intends to reveal the effects of school accountability and autonomy on Korean students' achievement using PISA, which claims to "monitor the development of national education systems by looking closely at outcomes over time" (Kirsch et al., 2002).

Although studies have widely examined the effects and side-effects of a performance-based accountability policy at the single-state, cross-state, and international level, their results remain mixed. Some proponents of high-stakes external accountability contend that only consequential accountability measures will be effective in improving student achievement and that the public reporting of school performance without sanctions and rewards would not increase achievement (Dee & Jacob, 2011; Hanushek & Raymond, 2005). Others evidenced that public reporting of school performance would be a cost-effective policy to improve low-performing schools (Burgess, Wilson, & Worth, 2013). Nevertheless, some others found no significant effect of external accountability policies such as the No Child Left Behind Act (NCLB; Lee & Reeves, 2012).

By contrast, concerns about the negative consequences of high-stakes accountability, such as the exclusion of low-performing students from tests, teaching to the rating, narrowed curriculum and increased retention and drop-out rates, have been consistently raised (Figlio & Ladd, 2008). Under the principal–agent theory, teachers, schools, and even the state governments are likely to have an incentive to show opportunistic behaviors, and high-stakes accountability only intensifies their incentives to behave strategically. Thus, many critics of performance-based accountability argue that test-based external accountability alone would not lead to improved, genuine, and long-term student learning outcomes and that internal accountability buttressed by professional accountability and organizational capacity should precede external accountability (Elmore, 2004; Newmann, King, & Rigdon, 1997; O'Day, 2002).

South Korea has been one of the top-scoring countries in PISA. Despite its consistent high performance in various international achievement tests, the Korean government has pursued education reform plans under the influence of global convergence towards performance accountability and school autonomy. Particularly since 2008, Korea has instituted school accountability measures under the Lee Myung-bak government, which highly stressed autonomy and accountability in education. During

his tenure, discourses on choice and diversification prevailed in educational policy making (Sung, 2011), and school accountability policies including the National Assessment of Educational Achievement (NAEA), public reporting of school performance, and financial incentives linked to school performance were implemented.

This study aims to investigate empirically the effects of school accountability and autonomy on student achievement using Korean PISA data for 2003 and 2012. In particular, we examine the effects of school accountability and autonomy on student achievement before and after 2008. Considering the high-stakes nature of the Korean NAEA, in that it is a population test and that school results are publicly reported, we used PISA. This dataset serves as an independent measure of the NAEA by the same logic as US empirical studies' use of the National Assessment of Educational Progress (NAEP; Carnoy & Loeb, 2002). This study thus sheds light on the effectiveness of school accountability and autonomy by adding Korean evidence to the international knowledge base on this highly contested topic.

Literature review

School accountability and autonomy as key school reform measures

In recent years, policy makers and research circles have concerned themselves with performance-based, standards-based, or test-based accountability, in which school performance is evaluated using student performance measures (Figlio & Loeb, 2011). Although there are various conceptions of educational accountability, such as political, legal, bureaucratic, professional, and market accountability (Darling-Hammond, 2007), performance-based accountability has increasingly driven educational reform policies in several educational systems.

In general, school accountability systems include three elements: state-wide student tests, public reporting of school performance, and rewards or sanctions based on some measure of school performance or improvement (Kane & Staiger, 2002). In other words, schools are externally accountable for student performance, and accordingly are rewarded or penalized. High-stakes school accountability occurs when students' results are used as rewards or sanctions for teachers and/or principals. In this sense, high school exit exams have a high-stakes nature for students and thus serve as a tool for student accountability (Lee & Wong, 2004). In contrast to external accountability, the concept of internal accountability holds teachers accountable for student learning in line with personal responsibility and shared expectations, combined with certain consequences (Elmore, 2004). Elmore and Fuhrman (2001) argue that internal accountability should precede external accountability in school reforms.

The global reform trend that emphasizes school accountability has encapsulated East Asian countries. The high performance of Asian countries such as Singapore, Japan, and Hong Kong in various international achievement tests has drawn significant

attention from the West. These countries have long had high-stakes national tests from the perspective of individual students, particularly at secondary exit levels or college entrance stages (Bishop, 1998); however, over the past decade, they have started to introduce performance-based school accountability to hold schools accountable for students' test results. Singapore has an accountability system for reporting schools' performance, including a ranking based on the results of national exams as well as a self-assessment called the School Excellence Model and associated awards (Ng, 2010). Japan started a national test for 6th and 9th grade students in 2007, which was scaled down to a sample of 30% of the students in 2010, but became mandatory once again in 2013 (Takayama, 2013). In Hong Kong, the school accountability framework comprises self-evaluation and external reviews, while standards are monitored through the Territory-wide System Assessment, in which a sample of 3rd, 6th, and 9th grade students from every school should participate. Hong Kong has also carried out student assessment reforms from high-stakes public examinations to assessment of learning, which include school-based, criterion-referenced assessment (M. M. C. Mok, 2007). Hence, although there exist variations in the specific accountability measures across countries, policy convergence towards school accountability is distinct.

In conjunction with school accountability, school autonomy has received considerable policy attention worldwide, in the belief that the decentralization of decision-making enhances the quality of schooling (Hanushek, Link, & Woessmann, 2013; Maslowski, Scheerens, & Luyten, 2007). Indeed, school accountability and autonomy are the catchwords of global education reform movements (Schutz, West, & Wossmann, 2007). The argument for decentralization is that local decision-makers have a better understanding of local needs and thus make better resource decisions, thereby raising productiveness and innovation. High-performing Asian countries have also pursued educational decentralization policies (K. Mok, 2003; Sui-chu Ho, 2006). One approach to decentralization is to set up independent or autonomous schools that are operated with more discretion and autonomy at the school level than regular schools, while another strategy is to devolve previously centralized power to local schools, regardless of school type. Singapore has taken both approaches, while Hong Kong has implemented a school-based management model called the School Management Initiative since 1991 (Cheng, 2009). To summarize, global education reform movements towards school accountability and autonomy prevail, even in high-performing Asian countries, and it is important to examine the effects of these highly popular institutional features.

Effects of school accountability and autonomy on student achievement

From a theoretical viewpoint, the effect of school accountability can be understood based on a principal–agent model, in which a principal (the parent) commissions an agent (the head of a school) to perform a service (the education of the child) on his or her behalf (Wößmann, Lüdemann, Schütz, & West, 2007). Because of the asymmetric

information problem regarding the efforts of schools and teachers, school accountability policies such as setting clear performance standards on external exams and providing performance information can provide incentives to raise student achievement.

Empirical evidence on the effect of accountability policies on students' achievement is predominantly found in the context of the United States, particularly since standard-based school reforms and performance-based accountability policies, for example, NCLB, were aggressively pursued in the US. Some researchers performed cross-state analyses using the NAEP results to identify the effects of NCLB on student achievement (Carnoy & Loeb, 2002; Dee & Jacob, 2011; Hanushek & Raymond, 2005; Lee & Reeves, 2012; Lee & Wong, 2004), whereas others conducted single-state studies to examine the more nuanced effect of state high-stakes accountability policy (Chiang, 2009; Jacob, 2005; Ladd, 1999; Reback, 2008; Rouse, Hannaway, Goldhaber, & Figlio, 2007; Springer, 2008).

According to the US empirical literature, evidence is rather mixed: some studies report significantly positive effects of external accountability, particularly in math achievement (Carnoy & Loeb, 2002; Dee & Jacob, 2011; Hanushek & Raymond, 2005), while others suggest that NCLB did not generate sustainable and generalizable policy effects, with little equity enhancement (Lee & Reeves, 2012; Lee & Wong, 2004). Lee (2008) conducted a meta-analysis and concluded that the high-stakes accountability policy showed a modestly positive effect on average achievement. According to Figlio and Ladd (2008), an accountability policy seemed more effective in math than in reading, with a modest effect size, but it hardly closed the achievement gap, particularly between White and Black students. Another important theme in US studies is that schools under external threats and incentives mainly focused on below-proficient students (Reback, 2008; Springer, 2008), and threat-induced improvement of math test scores persisted for a few years (Chiang, 2009). Jacob (2005) suggested that the observed achievement gains were driven by increases in test-specific skills, while Rouse et al. (2007) attributed a significant portion of test score gains to substantive changes in instructional policies and practices.

Some cross-country studies have provided positive evidence of the effect of accountability on student achievement in PISA and other international tests. Hanushek and Woessmann (2010) reviewed a vast array of economic literature on the international differences in student achievement. Their review suggests that accountability and autonomy are important institutional features that contribute to higher student performance. For example, students in systems with external exit examinations are more likely to perform better (Bishop, 1998; Wößmann, 2007). These international comparative studies highlight the effect of central exit exams, which serve as a student accountability measure, on student achievement. In addition, accountability measures aimed at teachers, such as the internal and external monitoring of teachers' lessons, and those aimed at schools, such as assessments used to compare school performance with district or national performance, have been effective in increasing student performance (Wößmann et al., 2007).

The effect of school autonomy on achievement remains mixed, however. Hanushek,

Link, and Woessmann (2013) suggested that autonomy might be conducive to student achievement in well-developed systems but detrimental in low-performing systems, providing panel data estimation results from five cycles of PISA data as evidence. On the contrary, Maslowski, Scheerens, and Luyten (2007), by using PISA 2000 data, suggested that few factors related to autonomy appear to enhance the reading literacy scores of students. Methodological differences may have led to these divergent conclusions. For example, the former study applied a panel estimation with 42 sample countries participating in PISA, whereas the latter used hierarchical linear modeling with cross-sectional data on 25 OECD countries.

Theory suggests that external exams and school autonomy are complementary, implying that one is only beneficial if the other is also in place (Wößmann et al., 2007). School autonomy is expected to lead to increased efficiency when agents conform to the objectives set by principals. However, there is danger of opportunism by decentralized decision-makers who have more local knowledge. Information asymmetries between principals and agents can be bridged, for instance, by external exams. In a series of studies, Wößmann and colleagues showed that school autonomy is effective when combined with external accountability (Hanushek et al., 2013; Wößmann, 2005, 2007; Wößmann et al., 2007).

In summary, the US-based literature presents a modestly positive effect of high-stakes accountability policy on student achievement, while international cross-country studies show a significant positive effect of central exit examinations and a positive effect of school autonomy, especially when combined with accountability.

Korean policy context

School accountability and autonomy policies

From an international and comparative perspective, Korea has been a strong performer in various international student achievement tests. Under the context of global education reforms toward performance-based accountability and decentralization of decision-making, Korea has also intensified school accountability measures and sought diversification and decentralization since the *5.31 Education Reform Plan* in 1995. In particular, the performance-based accountability system was distinctly strengthened under President Lee's administration. The three elements of the school accountability system were instituted one after another in 2008 under Lee's government.

First, the Korean NAEA, which was previously administered as a sample test to students in the 6th, 9th, and 10th grade from 1998 to 2007, was converted into a national test for the entire student population in 2008. While a sample test continued to be administered by the Korea Institute for Curriculum and Evaluation (KICE), in 2008 it was accompanied by the population test under the Local Education Authorities

(LEA). In 2009, the NAEA expanded to cover all students in the 6th, 9th, and 10th grade.[1)]

Second, the public reporting of school performance was initiated in 2010. Every primary and secondary school in Korea was required to publish information on students' proficiency levels (below basic, basic, and above basic) in reading, math, and English on the publicly accessible school information database.[2)] Since 2011, value-added school performance has been reported as well.[3)]

Third, in 2010, the government initiated the School for Improvement program, under which it selected low-performing schools for additional financial support to increase student achievement on the basis of the NAEA results. In addition, NAEA-based school performance is being widely used for the formal evaluation of teachers, schools, and LEAs. For example, performance indicators include the proportions of students below basic proficiency at the LEA level and value-added student performance, which are linked to financial incentives for teachers and principals. In short, sanction threats are not conspicuous, whereas monetary rewards are linked to school performance in the Korean policy context.

In the policy implementation process, the school accountability policies raised serious controversies over the negative repercussions such as league tables, intense competition, teaching to the test, and discouraging low-performing students (Sung & Kang, 2012). The NAEA, called *iljegosa* by the Korean media, was refused by progressive superintendents of some LEAs and the Korean Teachers Union. Some teachers who disagreed with national testing were reported to be making a stand by instructing their students to skip exams (B. Kim, Choi, & Baik, 2010). In other words, variations in external accountability exist at the school level because of infidelity of implementation by some principals and teachers.

In accordance with school accountability policies, the government put forward a masterplan for improving school autonomy in 2008 (Korean Ministry of Education [KMOE], 2008). Subsequently, school autonomy over curriculum and staffing significantly increased. For example, schools can now use their discretion to increase or decrease class hours per subject within 20% of total class hours. In addition, school principals can decide how to compose 20% of the teaching body, for example by inviting teachers from other schools. Furthermore, school diversification has intensified in the past decade with many autonomous high schools established, thereby fundamentally dismantling the high school equalization policy under which students were randomly assigned to schools near their residence.

Since the policy changes in 2008, several empirical studies have been conducted on the effect of accountability on Korean students' achievement, although they present mixed results. Park, Kim, and Sung (2010) explored the effects of school accountability measures in the 2006 PISA on Korean students' science achievement and found no significant relationship. Similarly, Park, Kim, Oh, Jeong, and Kim (2014) investigated the association of internal and external accountability measures of principals with Korean students' math and science achievement in the 2011 TIMSS and found no significant relationship. By contrast, K.-H. Kim et al. (2012) examined the effect of the

School for Improvement program during 2009–2010 and 2010–2011 using a regression discontinuity approach and found a significantly positive program effect on students' academic performance in the NAEA. Specifically, in 2010, the below-basic student proportion decreased by 2–3.5% points across subjects (English, reading, and mathematics) for middle school and 3–5% points for high school, and in 2011, it decreased by 1–2% points for middle school and 2–4% points for high school. However, as this study used the NAEA results, the estimates could have been upwardly biased due to test score inflation. W. Kim (2012) examined how educational achievement gaps have changed in the 2000s, focusing on the effects of school policy changes such as school accountability and autonomy, and suggested that increasing differentiation and school autonomy seemed to have a considerable impact.

Conceptual framework and research questions

In this study, we examine the effects of school accountability and autonomy on Korean students' math achievement in PISA. Adopting a theoretical framework based on the principal–agent model, we hypothesize that increased school accountability and autonomy is likely to increase the level of student achievement and decrease the probability of students falling below basic proficiency levels.

By using four cycles of PISA data, we first test a contrast hypothesis that average Korean PISA achievement in math differed before and after 2008 when school accountability and autonomy significantly increased. Then, we seek to test whether school accountability and autonomy are positively associated with student achievement. In the Korean policy context, accountability policies such as school performance reporting and the School for Improvement program target low-performing schools and students. Hence, examining how accountability and autonomy affect the achievement of students at different achievement levels is worthwhile. The specific research questions are as follows:

1. Does Korean students' math achievement differ across PISA cycles, specifically before and after 2008, when school accountability and autonomy were strengthened? Further, has the proportion of students below the basic proficiency level declined since 2008?
2. Do school accountability and autonomy affect Korean students' math achievement in PISA 2003 and PISA 2012?
3. Do school accountability and autonomy affect the probability of scoring below the basic math proficiency level in PISA 2003 and PISA 2012?
4. How do school accountability and autonomy affect the math achievement of students at different achievement levels?

Methods

Data and sample

We retrieved data from the official OECD PISA website (OECD, 2014a) and selected only Korean students and schools for the analyses. Since 2000, OECD has conducted the PISA survey every three years, in all of which Korea has participated. To conduct a contrast analysis, we used four cycles of the Korean PISA data from 2003 to 2012 to observe the trend in math achievement, particularly before and after 2008, when school accountability and autonomy was distinctly intensified.[4] For subsequent multilevel analyses and quantile regression analyses, we used 2003 and 2012 PISA data, of which the main subject was math, to investigate the effects of school accountability and autonomy on student math achievement, controlling for various student and school characteristics affecting math scores. The PISA survey is designed to test the skills and knowledge of 15-year old students who attend either lower or upper secondary school. We selected students attending a general high school to ensure the comparability of school characteristics between the two different cycles.[5] The sample restrictions resulted in final sample sizes of 3,645 and 3,734 students and 93 and 110 schools for PISA 2003 and 2012, respectively.

Variables

Dependent variables

The dependent variables in this study are math achievement and a dummy variable indicating students below the basic proficiency level in PISA 2003 and 2012.

Measures of school accountability and school autonomy

In this study, school accountability measures the use of student assessment for school performance and teacher evaluation. It includes comparing school performance to external benchmarks and making judgments about teachers' effectiveness (Table 1 presents the specific questions used). School accountability can be measured with additional variables, such as teacher monitoring methods (available in PISA 2003, 2009, and 2012) and consequences of teacher evaluation (available in PISA 2012). As the question items in the PISA school questionnaire vary across cycles, we selected questions that were consistently asked to construct a measure of school accountability related to student assessment.[6]

School autonomy measures one over resource allocation and the other over curricula and assessments. School principals were asked to report whether the teachers, principals, governing board, and regional or local education authorities or national education authority were significantly responsible for allocating resources (e.g., appointing

and dismissing teachers, determining teachers' starting salaries and salary raise, and formulating school budgets and allocating them within the school) and curriculum and instructional assessment within the school (e.g., establishing student assessment policies, choosing textbooks, and determining courses to offer and the content of these courses). We coded 1 for each variable when the school (i.e., teachers, principal, and school's governing board) had considerable responsibilities and 0 otherwise. Then, we averaged the values for each construct. The reliability coefficients (Cronbach's α) of composite variables are presented in Table 1.

Table 1. Variables

Variables	Description
Student-level	
Female	Dummy coded as 1 if female and 0 if male
SES	PISA Index of economic, social, and cultural status (ESCS)
Extra math hours	I do not attend out-of-school-time lessons in math = 1 Less than 2 hours per week = 2 2 or more but less than 4 hours per week = 3 4 or more but less than 6 hours per week = 4 6 or more hours per week = 5
Math self-efficacy	PISA Index of math self-efficacy (MATHEFF)
Student–teacher relationship	PISA Index of teacher–student relationship (STUREL)
Disciplinary climate	PISA Index of disciplinary climate (DISCLIM)
School-level	
Control	Dummy coded as 1 for private schools and 0 for public schools
Mean SES	Aggregated individual ESCS at the school level
Location Urban	Urban school located in a large city with over 1,000,000 people
Suburban	Suburban school located in a town or city with about 15,000–1,000,000 people
Rural	Rural school located in a small town or rural village with less than 15,000 people (reference)
Students per math teacher	Math teacher to student ratio
Teacher morale	PISA Index of teacher morale (TCMORALE)
School climate	Composite variable of school climate, specifically 13 survey items on learning hindrances in school questionnaire were recoded and averaged to express positive school climate. α = .908
School accountability	Average of whether student assessments are used to: a) compare school with district or national performance b) monitor school's annual progress c) judge teachers' effectiveness d) compare one school with another. α = .703

Variables	Description
Autonomy in resource allocation	Average of whether principals, teachers, or a school governing board have considerable responsibilities: a) hiring teachers b) dismissing teachers c) deciding teachers' starting salaries d) determining teachers' salary increases e) formulating school budget, and f) allocating the budget. α = .707
Autonomy in curricula & assessment	Average of whether principals, teachers, or a school governing board have considerable responsibilities: a) disciplinary policies b) assessment policies c) student admission d) choosing textbooks e) determining course contents, and f) deciding which courses to offer. α = .784
Dependent variables	
Math scores	5 plausible values in math
Below basic proficiency	Dummy coded as 1 if a student's math performance was below the basic proficiency level, and 0 otherwise

Source: OECD. (2014a). PISA 2003, 2012.
Note. For extra math hours, PISA 2003 data were converted into the same ordered values as PISA 2012. Five index variables (SES, math self-efficacy, student–teacher relationship, disciplinary climate, and teacher morale) provided by the PISA datasets are used. See PISA 2003 and 2012 technical reports for the scaling procedures of the index variables. Reliability coefficients (Cronbach's α) are provided at the end of description of each composite variable (school climate, school accountability, autonomy in resource allocation, and autonomy in curricula & assessment).

Control variables

In the multilevel model specification, we included control variables that could influence math achievements at the student and school level (see Table 1). On the basis of prior empirical studies, we included gender, socio-economic status (SES), extra math study hours, math self-efficacy, student–teacher relationship, and disciplinary climate as control variables at the student level. Gender and SES are generally strong predictors of student math achievement. Regular math study hours in school may not differ considerably across schools, while pervasive private supplementary tutoring in Korea varies among students, which is likely to affect math achievement. Self-efficacy, which is an essential motivation to learn, is also a strong predictor of achievement (see also Zimmerman, 2000). The student–teacher relationship and disciplinary climate are positively associated with student achievement (Shin, Lee, & Kim, 2009). At the school level, characteristics such as control, location, mean SES, math teacher to student ratio, teacher morale, and school climate were controlled for in line with previous literature

(W. Kim, 2012). Private schools may systematically and operationally differ in student achievement. A higher mean SES is positively associated with student achievement, while the teacher to student ratio is considered to be an input resource. Further, a high level of teacher morale and of school climate positively affects student achievement.

The preliminary analysis revealed considerable missing values on background variables in PISA 2012. For example, the control variables such as extra math hours, math self-efficacy, student–teacher relationship, and school climate had over 30% missing values. Thus, a multiple imputation (m = 5) was applied to prevent potential selection bias and serious reduction in sample size (Weirich et al., 2014).

Data analysis

First, we employed a contrast analysis method to compare average math scores before and after 2008. Also, we examined the trend of the proportions of students below the basic proficiency level in math.

Second, a two-level HLM was used to estimate the effects of school accountability and autonomy on students' math achievement, controlling for variables affecting math scores. This multilevel analytical method was chosen because the data have a hierarchical structure with individual students nested within school (Raudenbush & Bryk, 2002). Also, all student- and school-level variables which were not dichotomous were centered around their grand means. The level 1 and level 2 models are as follows:

Level 1 model (student level):

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj}(X_q)_{ij} + e_{ij},\ e_{ij} \sim N(0,\ \sigma^2).$$

Level 2 model (school level):

$$\beta_{0j} = \gamma_{00} + \sum_{s=1}^{Q} \gamma_{0s}(W_s) + u_{0j,}$$

$$\beta_{1j} = \gamma_{10} \ldots\ldots \beta_{qj} = \gamma_{q0}, \quad u_{0j}\ N(0,\ \sigma^2).$$

Third, a two-level hierarchical generalized linear model (HGLM) was used to estimate the effects of school accountability and autonomy on the probability of falling below the basic proficiency level in math, controlling for variables affecting math proficiency. The dependent variable is dichotomous, indicating whether the student's math performance falls below the basic proficiency level. The level 1 and level 2 models are as follows:

Level 1 model (student level):

$$Y_{ij} \sim bin(\boldsymbol{\phi}_{ij}, 1),$$

$$\eta_{ij} = \log\left(\frac{\boldsymbol{\phi}_{ij}}{1 - \boldsymbol{\phi}_{ij}}\right) = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj}(X_q)_{ij}.$$

Level 2 model (school level):

$$\beta_{0j} = \gamma_{00} + \sum_{s=1}^{Q} \gamma_{0s}(W_s) + u_{oj},$$

$$\beta_{1j} = \gamma_{10} \cdots\cdots \beta_{qj} = \gamma_{q0}, \quad u_{0j} \sim N(0, \tau_{00}).$$

Additionally, we employed a quantile regression model developed by Koenker and Bassett Jr. (1978) to reveal the effects of school accountability and autonomy on students with different achievement levels. We used the quantile regression model to determine for which achievement level of students school accountability and autonomy matters. The quantile regression model for this study is as follows:

$$Y_i = X'_i\beta_\theta + u_{\theta i}, Quant_\theta(Y \mid X_i) = X'_i\beta_\theta,$$

where Y_i is a student's math achievement. $Quant_\theta(Y \mid X_i)$ refers to the conditional quantile of Y_i, conditional on the vector of explanatory variables X_i and $\theta \in (0, 1)$. It is assumed that $Quant_\theta(u_i \mid X_i) = 0$. The quantile regression estimates are obtained by minimizing the weighted sum of the absolute values of the errors.

Results

Descriptive analyses

We conducted a contrast analysis to determine whether Korean students' average math achievement in PISA differed before and after 2008. Figure 1 shows the mean math achievement of Korean students over the PISA cycles. Korean students' average math scores continuously increased from 2003 to 2012. The contrast analysis revealed that, on average, Korean students' math scores in 2009 and 2012 increased compared to the previous cycles (p = .000). However, it is uncertain whether the overall improvement can be attributed to the intensification of school accountability and autonomy.

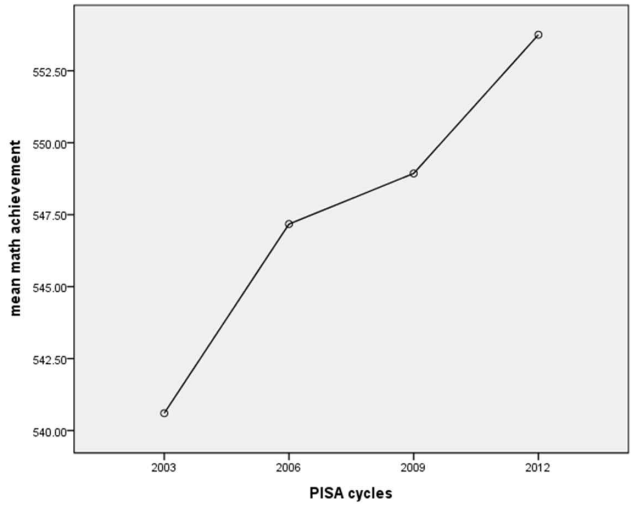

Figure 1. Mean math achievements of Korean students in PISA

Source: OECD. (2014a). PISA 2003, 2006, 2009 and 2012.

Figure 2 shows the trend of Korean students' PISA achievements focusing on the below basic proficiency level.[7] It is notable that the proportions of students below the basic proficiency level in math increased in 2012, contrary to the prior decreasing trend. It is surprising that the proportions of low performers increased despite the school accountability policy and the government's considerable efforts to target low-performing schools.

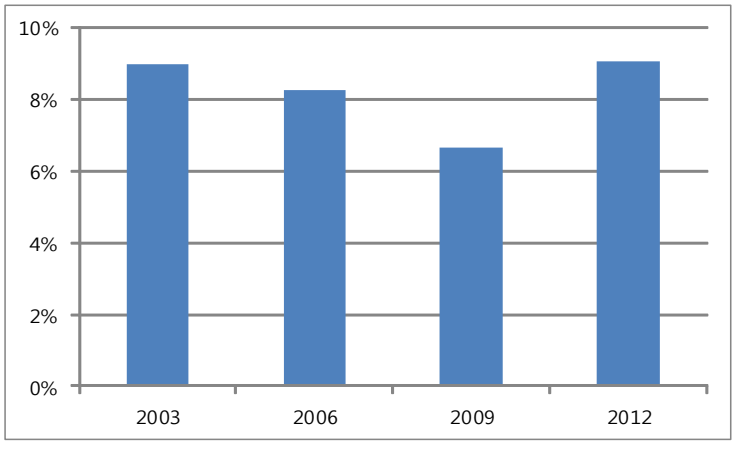

Figure 2. Proportions of below-basic proficiency in math among Korean students in PISA

Source: OECD. (2014a). PISA 2003, 2006, 2009 and 2012.

Table 2 shows the descriptive statistics for the restricted sample of PISA 2003 and 2012. The average math achievement among general high school students increased from 569.77 in 2003 to 574.95 in 2012. At the same time, the proportion of students who scored below the basic proficiency level increased from 3.1% to 5.2%. The proportion of female students is 39.1% and 46.5% in the respective cycles. Extra math hours declined from the scale score of 4.12 in 2003 to 3.08 in 2012. It is noteworthy that private schools comprised 63.3% of 93 general high schools in the 2003 sample, compared to 47.3% of 110 schools included in the 2012 sample.[8] In general, private schools are likely to have a higher degree of administrative and academic autonomy. Therefore, higher autonomy in the 2003 sample could have emanated from the oversampling of private schools. The composite variable of school accountability is .67 for 2003, whereas for 2012, it is .81, which reflects the intensification of school accountability policies introduced since 2008.

Table 2. Descriptive statistics

	PISA 2003		PISA 2012	
Variable	Mean	*SD*	Mean	*SD*
Dependent variables				
Math scores	569.77	79.70	574.95	90.36
Below basic proficiency	0.03	0.17	0.05	0.22
Student-level variables				
Female	0.39	0.49	0.47	0.50
SES	0.09	0.80	0.12	0.72
Extra math hours	4.12	1.17	3.08	1.41
Math self-efficacy	−0.21	0.97	−0.20	1.03
Student–teacher relationship	−0.06	0.84	−0.11	0.87
Disciplinary climate	0.18	0.79	0.24	0.85
School-level variables				
Private	0.63	0.48	0.47	0.50
Mean SES	0.09	0.41	0.12	0.31
Urban	0.50	0.50	0.43	0.50
Suburban	0.45	0.50	0.53	0.50
Students/math teacher	125.58	20.72	113.96	34.83
Teacher morale	−0.26	1.02	−0.17	1.06
School climate	3.41	0.49	3.17	0.49
School accountability	0.67	0.32	0.81	0.28
Autonomy in resource allocation	0.46	0.22	0.32	0.24
Autonomy in curricula & assessment	0.98	0.06	0.88	0.22
Number of observations	3,645		3,734	

Note. The samples are restricted to students attending general high school only. Five datasets drawn by multiple imputation are used.

Multilevel analyses

To account for a nested data structure and disentangle the relationship between accountability & autonomy and math achievement, we conducted HLM analyses. Table 3 presents both unconditional and conditional HLM model results. According to the unconditional model, the between-school variance accounted for 29% of the total variance in PISA 2003 and 34.5% in PISA 2012. This shows that school differences in math scores can be reliably examined using school-level predictors (Raudenbush & Bryk, 2002).

In the conditional model, we examined the effect of accountability and autonomy on achievement, controlling for various student- and school-level predictors. It turned out that school accountability of Korean general high schools were not related to math scores in either PISA 2003 or 2012. However, school autonomy in curricula and assessment was significantly associated with math achievement. Schools with higher autonomy over curricula and assessment, including the course offerings and contents, are likely to increase student math achievement effectively.

In both cycles, the significant predictors of math achievement at the student level included gender, SES, extra math hours, and math self-efficacy. Male students are likely to achieve 14.41 points higher than female students, but this gender gap in math achievement narrowed in 2012 (p = 0.07). The student–teacher relationship variable was a statistically significant predictor of math achievement in 2012, but not in 2003. At the school level, the mean SES and math teacher to student ratio were significantly related to student math achievement in both cycles, and the effect of school mean SES on student math score increased in 2012. School climate was significantly positively associated with math achievement in 2012, but not in 2003.

Table 3. Multilevel results from the two-level HLM analysis of math achievement

	PISA 2003		PISA 2012	
	Unconditional model	Conditional model	Unconditional model	Conditional model
Intercept	569.64 ***	473.53 ***	574.10 ***	489.18 ***
	(4.59)	(48.82)	(5.22)	(27.77)
Fixed effect at student level				
Female		-14.41 ***		-5.22 †
		(2.72)		(2.89)
SES		4.40 **		4.81 **
		(1.42)		(1.77)
Extra math hours		9.75 ***		5.58 ***
		(0.92)		(0.80)
Math self-efficacy		32.94 ***		40.42 ***
		(1.11)		(1.18)
Student–teacher relationship		0.55		4.01 **
		(1.23)		(1.29)
Disciplinary climate		0.03		3.65
		(1.29)		(2.12)

	PISA 2003		PISA 2012	
	Unconditional model	Conditional model	Unconditional model	Conditional model
Fixed effect at school level				
Private		-1.69		0.98
		(5.76)		(7.63)
Mean SES		45.06 ***		56.85 ***
		(7.45)		(8.86)
Urban		4.78		19.43
		(13.10)		(14.04)
Suburban		15.63		21.41
		(12.51)		(14.19)
Students/math teacher		-0.29 **		-0.19 *
		(0.11)		(0.08)
Teacher morale		2.73		2.68
		(2.69)		(2.71)
School climate		6.47		13.76 *
		(5.13)		(5.36)
School accountability		0.19		9.48
		(7.79)		(9.60)
Autonomy in resource allocation		1.71		-4.07
		(11.59)		(17.29)
Autonomy in curricula & assessment		71.06 †		25.78 *
		(42.06)		(11.95)
Random effect				
Level 1 variance	4513.05	3310.65	5383.80	3779.10
Level 2 variance	1842.72	393.61	2839.90	535.98

Note. Five datasets drawn by multiple imputation are used for the conditional model estimation. Standard errors for regression coefficients are in parenthesis.
†$p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Next, we examined whether school accountability and autonomy was related with students' scoring below basic proficiency. Table 4 presents the HGLM analysis results. First, the results indicated that school accountability did not significantly affect the probability of students falling below the basic proficiency level. However, autonomy in curricula and assessment was negatively associated with students scoring below basic proficiency in 2012. That is, we can infer that students are less likely to fall below the basic proficiency level in schools with higher autonomy in curricula and assessment.

Student-level variables were stronger predictors of the probability of below basic proficiency in math than school-level variables. Extra math hours and math self-efficacy had a statistically significant negative relationship with students' probability of scoring below the basic proficiency in math. In 2003, female students were less likely to fall below the basic proficiency level, but this gender difference disappeared in 2012. By contrast, a positive student–teacher relationship contributed to a lower probability of students falling below the basic proficiency level in 2012.

Table 4. Multilevel results from the two-level HGLM analysis of the probability of below-basic proficiency level in math

	PISA 2003			PISA 2012		
	Coefficient	*SE*	*OR*	Coefficient	*SE*	*OR*
Intercept	-1.70	3.56	0.18	-0.18	1.78	0.83
Fixed effect at student level						
Female	-0.50†	0.27	0.61	-0.07	0.20	0.93
SES	-0.13	0.15	0.88	-0.19	0.13	0.83
Extra math hours	-0.38***	0.08	0.68	-0.33**	0.09	0.72
Math self-efficacy	-0.96***	0.13	0.38	-1.19***	0.12	0.31
Student–teacher relationship	-0.11	0.14	0.90	-0.38**	0.13	0.68
Disciplinary climate	-0.06	0.14	0.94	-0.11	0.13	0.90
Fixed effect at school level						
Private	0.34	0.41	1.41	-0.51	0.47	0.60
Mean SES	-1.08†	0.57	0.34	-0.43	0.54	0.65
Urban	-1.12	0.80	0.33	-0.62	0.71	0.54
Suburban	-0.86	0.73	0.42	-0.94	0.76	0.39
Students/math teacher	0.01†	0.01	1.01	0.01	0.00	1.01
Teacher morale	-0.15	0.21	0.86	-0.24	0.16	0.79
School climate	-0.09	0.39	0.91	-0.52	0.37	0.60
School accountability	0.69	0.58	1.98	-0.62	0.52	0.54
Autonomy in resource allocation	0.80	0.88	2.23	0.96	1.11	2.61
Autonomy in curricula & assessment	-3.24	3.20	0.04	-1.41†	0.75	0.24
Random effect						
Estimates	0.91			0.93		
SD	0.19			0.15		

Note. Five datasets drawn by multiple imputation are used for the estimation.
†$p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Quantile regression analysis

In addition, we employed a quantile regression model to reveal the effects of school accountability and autonomy on students of different achievement levels, after controlling for various student and school characteristics. Table 5 presents the quantile regression results. First, school accountability did not have a statistically significant relationship with math scores in 2003 and 2012. However, autonomy in curricula and assessment had a statistically significant association with math achievement at the 0.10, 0.25, 0.50, and 0.90 quantiles in 2003 and at the 0.10–0.75 quantiles in 2012. Except for the 0.90 quantile in 2003, the effect of instructional autonomy was larger with the lower quantiles.

Table 5. Quantile regression results for math achievement

	PISA 2003					PISA 2012				
	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Female	-11.25 **	-14.17 ***	-19.97 ***	-23.50 ***	-19.55 ***	-0.09	-2.82	-9.72 *	-11.96 ***	-13.82 ***
	(3.97)	(2.93)	(2.49)	(2.91)	(3.72)	(5.29)	(3.90)	(3.93)	(3.21)	(3.72)
SES	1.30	2.18	4.95 *	4.06 †	5.03 †	4.55	2.23	4.68 †	3.28	1.38
	(2.29)	(2.47)	(2.09)	(2.21)	(2.59)	(3.54)	(3.19)	(2.49)	(2.61)	(2.55)
Extra math hours	11.35 ***	12.66 ***	10.98 ***	8.71 ***	8.45 ***	8.00 ***	6.93 ***	6.05 ***	4.28 ***	2.63 *
	(1.96)	(1.29)	(1.49)	(1.27)	(1.49)	(1.86)	(1.35)	(1.09)	(1.03)	(1.27)
Math self-efficacy	40.75 ***	38.36 ***	36.68 ***	33.80 ***	31.11 ***	48.55 ***	47.76 ***	44.45 ***	41.09 ***	39.90 ***
	(2.00)	(1.77)	(1.97)	(1.53)	(2.32)	(2.61)	(1.70)	(1.68)	(1.99)	(2.09)
Student–teacher relationship	1.21	2.07	0.13	-0.77	0.11	6.02 *	5.92 *	3.65	4.46 †	5.36 †
	(1.93)	(1.75)	(1.53)	(1.46)	(2.06)	(2.87)	(2.23)	(2.49)	(2.27)	(3.07)
Disciplinary climate	-0.17	2.48	2.25	2.64	-0.98	8.54 *	7.41 ***	5.20 *	5.32 *	4.85
	(2.40)	(1.77)	(2.01)	(1.82)	(2.46)	(3.28)	(1.99)	(2.08)	(2.39)	(2.85)
Private	-1.20	-2.35	-0.38	-4.29	3.34	4.87	5.28	-0.52	-3.03	-3.05
	(3.95)	(3.83)	(3.28)	(2.99)	(4.97)	(6.18)	(5.83)	(4.39)	(4.89)	(5.39)
Mean SES	39.33 ***	41.73 ***	45.41 ***	48.75 ***	49.16 ***	39.83 ***	44.87 ***	54.32 ***	59.21 ***	63.37 ***
	(7.02)	(4.26)	(4.48)	(4.45)	(6.71)	(8.55)	(6.94)	(5.76)	(5.77)	(7.09)
Urban	3.39	5.57	8.75	4.50	-26.21 *	10.64	19.98 *	24.72 **	16.13 †	12.72
	(10.94)	(8.83)	(7.11)	(9.79)	(12.59)	(13.07)	(9.41)	(8.17)	(9.46)	(12.45)
Suburban	15.66	14.70 †	21.97 **	14.67	-11.72	11.17	20.91 *	25.56 **	17.37 †	12.85
	(10.48)	(7.99)	(6.47)	(9.39)	(12.27)	(13.30)	(9.55)	(7.97)	(9.20)	(12.44)
Math teacher to student ratio	-0.28 *	-0.27 ***	-0.26 ***	-0.33 **	-0.23 †	-0.13 †	-0.21 ***	-0.19 ***	-0.13 **	-0.08
	(0.12)	(0.07)	(0.07)	(0.10)	(0.12)	(0.07)	(0.04)	(0.04)	(0.04)	(0.08)
Teacher morale	7.16 **	3.84 *	2.75 **	0.36	-0.23	6.99 *	2.78	-0.93	0.95	-0.67
	(2.70)	(1.93)	(1.35)	(1.58)	(1.85)	(2.56)	(1.84)	(1.80)	(1.81)	(2.11)
School climate	11.83 **	8.16 *	6.81 *	6.05 *	4.12	11.89 *	14.67 **	16.85 ***	13.60 **	14.55 ***
	(3.98)	(4.06)	(2.79)	(2.90)	(3.68)	(4.88)	(4.15)	(4.32)	(4.25)	(3.90)
School accountability	-6.63	-4.10	1.37	-1.96	1.56	12.14	7.54	10.04	-0.76	0.64
	(6.92)	(4.83)	(5.17)	(4.97)	(6.47)	(8.09)	(8.01)	(7.36)	(6.41)	(7.15)
Autonomy in resource allocation	-6.82	-2.02	-3.00	14.61 *	2.82	-15.94	-11.76	-5.29	-0.62	0.75
	(10.71)	(8.15)	(7.08)	(5.79)	(8.69)	(12.73)	(10.71)	(10.40)	(10.00)	(13.91)
Autonomy in curricula & assessment	124.79 **	108.17 **	80.90 *	35.78	83.55 *	38.34 **	35.67 ***	28.43 **	25.78 **	16.59
	(43.46)	(33.14)	(33.34)	(29.20)	(39.33)	(12.48)	(8.78)	(9.71)	(7.52)	(10.23)
Intercept	328.48 ***	383.93 ***	454.64 ***	560.60 ***	567.29 ***	396.55 ***	436.45 ***	479.00 ***	548.44 ***	589.78 ***
	(49.10)	(35.98)	(36.57)	(34.88)	(45.50)	(25.71)	(23.39)	(20.82)	(20.00)	(22.70)

Note. Five datasets drawn by multiple imputation are used for the estimation. Standard errors for regression coefficients, which were obtained via bootstrapping, are in parenthesis.
†$p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Second, although other individual and school characteristics were used as covariates, some warrant consideration in comparison with the HLM results (see Table 3). For example, female students were expected to achieve lower scores than male students at all quantiles in 2003, but this was not the case at the 0.10 and 0.25 quantiles in 2012. This result is aligned with the HGLM results (see Table 4). Extra math hours and math self-efficacy had a significantly positive relationship with the math scores of all quantiles, and the effects were larger for the lower quantiles. By contrast, school mean SES also had a significantly positive association with math scores of all the quantiles, but the effects were larger for the higher quantiles. Teacher morale, which was not a significant predictor in the HLM analysis, was found to have a statistically significant relationship with math scores at the 0.10, 0.25, and 0.50 quantile in 2003 and 0.10 quantile in 2012. This indicates that teachers with individual responsibilities and enthusiasm can have a positive impact on low-performing students' achievement. School climate had a statistically positive relationship at the 0.10–0.75 quantile in 2003 and at all quantiles in 2012.

Discussion

In this study, we examined the trend of Korean students' math achievement in PISA over four cycles, particularly before and after 2008, when school accountability and autonomy was distinctly strengthened. Then, we investigated the effects of school accountability and autonomy on math achievement in PISA 2003 and 2012, controlling for various student and school characteristics.

The HLM analysis results indicate that school accountability was not associated with student math achievement in either PISA 2003 or 2012. The HGLM analysis results also reveal that school accountability did not have a statistical relationship with the probability of scoring below the basic proficiency level. The quantile regression analysis result did not differ from those of the multilevel analyses: the effect of school accountability on students at different math achievement levels was not found. These findings are consistent with those of previous empirical studies that used TIMSS 2011 or PISA 2006 Korean data (Park et al., 2010, 2014).

By contrast, school autonomy in curricula and assessment had a statistically positive association with math achievements at the .05 level and a negative association with the probability of scoring below the basic proficiency level at the .10 level in PISA 2012. Particularly, the instructional autonomy was more likely to affect lower-performing students than higher-performing ones as the size of coefficients decreased at higher quantiles.

The results indicate that school accountability systems in Korea, originally targeted at low-performing students, may not have been as effective as anticipated. Rather, school autonomy over curricula and assessment had a positive relationship with math achievement and seemed to have a stronger positive effect on students at lower

achievement levels.

How do we interpret the insignificant effect of school accountability and significant effect of instructional autonomy, particularly in 2012? First, Korea's school accountability system introduced in 2008 can be regarded as relatively less punitive since sanctions based on the NAEA results are not as consequential. Moreover, there are no serious sanction threats in Korea, such as school choice and replacement of principals, as in the case of some American states with a strong accountability system (Carnoy & Loeb, 2002). Second, public information about the proportions of students' proficiency levels at the NAEA reveals the annual school performance on the national test, but the general public may not be as attentive to the test results as expected. In fact, Korean general high schools' reputation is predominantly determined by the number of their graduates' admissions into prestigious universities, to which parents pay more attention. Thus, we need to take an excellence-oriented educational context into consideration to understand the mechanism of school accountability. Third, government programs for low-performing schools may not have been effective in improving students' knowledge and skills that PISA measured. Although it was reported that the School for Improvement program effectively decreased below-basic student proportions (K.-H. Kim et al., 2012), the data used for the analyses were drawn from the NAEA results. In other words, the government program may have boosted the observed student performance at the NAEA, which may not have translated into generalizable students' learning outcomes, as in the case of the PISA achievements explored in this study. Finally and most importantly, schools with higher autonomy over curricula and assessment are likely to provide individualized learning opportunities to students in need, thereby contributing more to improving low-performing students' math achievement. For example, the proportion of schools that introduced within-school curriculum differentiation as a form of ability grouping increased from 66.3% in 2007 to 82.2% in 2010 (KMOE, 2010).

Policy implications

Korea has been a consistently strong performer with high mean scores and low variances since the inception of PISA (OECD, 2014c). Nevertheless, the global education reform movement toward performance-based accountability led Korea to intensify school accountability and autonomy since 2008. The findings of this study indicate that school accountability was hardly associated with Korean students' math achievement in 2012, whereas school autonomy over curricula and assessment had a positive relationship with math achievement, particularly of lower-performing students.

It is important to consider the Korean educational context to understand the study findings. The Korean education system has long relied on student accountability such as high-stakes college entrance examinations (Park, 2013). Based on a policy trend toward greater diversification and decentralization since the mid-1990s, government-led school reforms have sought to intensify school accountability and autonomy simultaneously.

However, as a system heavily influenced by student accountability, it seems that additional school accountability does not lead to a higher level of achievement or lower basic proficiency in math.

By contrast, school autonomy in curricula and assessment seems to be an effective means of increasing student achievement or decreasing the proportion of low-performing students. A high school equalization policy has been criticized as an obstacle to educational excellence in Korea, and school differentiation at the upper secondary level, such as the establishment of elite schools with special purposes and private autonomous schools, has intensified over the past decade. These schools are likely to enjoy a higher level of autonomy than general high schools. Although it is difficult to analyze the separate effect of these schools because of data limitations,[9] the findings suggest that a higher degree of instructional autonomy in general high schools could improve student achievement.

Our findings are in line with those of the international literature, which has suggested that autonomy might be conducive to student achievement in well-developed systems with central external exams (Hanushek et al., 2013; Wößmann, 2005, 2007). Even before the NAEA as a population test started, Korea had long established external exams, which heavily influenced students' future academic careers. Therefore, we suggest an alternative approach to school accountability in Korea based on our study findings.

Instead of high-stakes school accountability policies such as NCLB in the United States, it is important to devise a fundamental solution to school reforms. Examples include enhancing organizational capacity (Newmann et al., 1997) and emphasizing professional accountability based on mutual trust between principals and agents (Møller, 2009; O'Neill, 2013; Sahlberg, 2008). Moreover, the subsidiary finding that teacher morale is positively related to student math achievement at the lower quantiles suggests that internal accountability rather than externally imposed accountability should be promoted in the Korean educational context.

Limitations and implications for future research

There exist some limitations to this study. First, it examined the effects of school accountability on Korean students' math achievement using PISA 2003 and 2012. Our findings are based on cross-sectional data, which make it difficult to capture a causal relationship between accountability or autonomy and achievement. Individual- or school-level longitudinal data will enable us to research further the causal impact of school accountability and autonomy on student achievement.

The second limitation is the measures of accountability. Within the limits of the PISA surveys, we referenced the question items aimed at student assessment only. However, accountability measures can encompass principal leadership behaviors, quality assurance methods, and consequences of teacher evaluation, among other components. This study used only consistent measures of accountability between PISA

2003 and 2012 for the purpose of comparison; thus, future research needs to use a number of these accountability measures included in PISA 2012.

Furthermore, multiple cycles of international comparative student achievement data have accumulated over the years, which can facilitate not only within-country but also cross-country analyses. More consistent question items across each cycle of data can help observe the detailed changes in educational systems. Future research should be conducted to compare the effectiveness of school accountability and autonomy across countries.

1) In 2010, the policy was amended so that 11th graders instead of 10th graders were required to take the NAEA (http://naea.kice.re.kr).
2) In 2008, the government launched the school information database as per the Special Act on the Information Disclosure of Educational Institutions. For more information, please see http://www.schoolinfo.go.kr.
3) Value-added performance is measured by subtracting actual student scores from expected scores, based on students' prior achievement (NAEA, 2014).
4) PISA scores for math scores are comparable only from 2003 to 2012 (OECD, 2014b, p. 280). For the contrast analysis, we included all Korean participants in PISA.
5) We excluded vocational high school students in our analyses given that they were exempted from taking the NAEA in 2012 (KMOE, 2012).
6) PISA 2012 comprises the most abundant information related to school accountability, for example, measures of quality assurance, principal leadership behaviors, and consequences of teacher appraisal. As this study examines school accountability over time, we used only questions asked consistently across cycles.
7) The sample consists of all Korean participants including students attending lower secondary schools and vocational high schools.
8) In 2003, 628 of the 1,297 (48.4%) general high schools were private, whereas in 2012, there were 664 private schools out of 1,565 general high schools (42.4%) (Korean Ministry of Education & Korean Educational Development Institute, 2003, 2012).
9) The PISA school questionnaire does not ask specific school types, and it is impossible to separate out school effects by specific categories of general high schools in Korea.

Address for correspondence

Pilnam Yi
Assistant Professor
Department of Education
Hongik University
Wowsan-ro 94, Mapo-gu
Seoul, 04066, Korea
Email: pilnamyi@hongik.ac.kr

References

Bishop, J. H. (1998). The effect of curriculum-based external exit exam systems on student achievement. *The Journal of Economic Education*, *29*(2), 171–182.

Burgess, S., Wilson, D., & Worth, J. (2013). A natural experiment in school accountability: The impact of school performance information on pupil progress. *Journal of Public Economics*, *106*, 57–67. doi:10.1016/j.jpubeco.2013.06.005

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, *24*(4), 305–331. doi:10.3102/01623737024004305

Cheng, Y. C. (2009). Hong Kong educational reforms in the last decade: Reform syndrome and new developments. *International Journal of Educational Management*, *23*(1), 65–86. doi:10.1108/09513540910926439

Chiang, H. (2009). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, *93*(9–10), 1045–1057. doi:10.1016/j.jpubeco.2009.06.002

Darling-Hammond, L. (2007). Standards, accountability, and school reform. In C. E. Sleeters (Ed.), *Facing accountability in education* (pp. 78–111). NY: Teachers College Press.

Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, *30*(3), 418–446. doi:10.1002/pam

Elmore, R., & Fuhrman, S. (2001). Holding schools accountable: Is it working? *Phi Delta Kappan*, *83*(1), 67–72. Retrieved from http://repository.upenn.edu/cgi/viewcontent.cgi?article=1007&context=gse_pubs

Elmore, R. F. (2004). *School reform from the inside out: Policy, practice, and performance*. Cambridge: Harvard Education Press.

Figlio, D. N., & Ladd, H. F. (2008). School accountability and student achievement. In H. F. Ladd & E. B. Fiske (Eds.), *Handbook of research in education finance and policy* (pp. 166–182). New York: Routledge.

Figlio, D. N., & Loeb, S. (2011). School accountability. In E. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 383–421). Amsterdam: Elsevier.

Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, *104*, 212–232. doi:10.1016/j.jdeveco.2012.08.002

Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297–327. doi:10.1002/pam.20091

Hanushek, E. A., & Woessmann, L. (2010, April). *The economics of international differences in educational achievement* (Working Paper No. 15949). Cambridge, MA: National Bureau of Economic Research.

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes

testing in the Chicago Public Schools. *Journal of Public Economics*, *89*(5–6), 761–796. doi:10.1016/j.jpubeco.2004.08.004

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, *16*(4), 91–114. doi:10.1257/089533002320950993

Kim, B., Choi, S., & Baik, S. (2010, July 14). The day of "iljegosa", no stress for achievement. *The Kyunghyang Shinmun*.

Kim, K.-H., Sang, K., Shin, J., Lee, B., Kwon, J., Woo, S., & Ryu, S. (2012). *A comparative analysis of the No Child Left Behind policy between Korea and the US*. Seoul: Korea Institute for Curriculum and Evaluation. (In Korean)

Kim, W. (2012). The trend in socioeconomic inequalities in student achievement: Focusing on the effects of school policy. *Korean Journal of Sociology of Education*, *22*(3), 49–76. (In Korean)

Kirsch, I., De Jong, J., Lafontaine, D., McQueen, J., Mendelovits, J., & Monseur, C. (2002). *Reading for change performance and engagement across countries: Results from PISA 2000*. Paris: OECD Publishing.

Koenker, R., & Bassett Jr., G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33–50. doi:10.2307/1913643

Korean Ministry of Education. (2008). *A masterplan for improving school autonomy*. Seoul: Author.

Korean Ministry of Education. (2010). *A masterplan for school autonomy: School based reform initiative*. Seoul: Author.

Korean Ministry of Education. (2012). *Annual plan for the national assessment of educational achievement 2012*. Seoul: Author.

Korean Ministry of Education., & Korean Educational Development Institute. (2003). *Statistical yearbook of education*. Seoul: Ministry of Education.

Korean Ministry of Education., & Korean Educational Development Institute. (2012). *Statistical yearbook of education*. Seoul: Ministry of Education.

Ladd, H. F. (1999). The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review*, *18*(1), 1–16. doi:10.1016/S0272-7757(97)00044-7

Lee, J. (2008). Is test-driven external accountability effective? Synthesizing the evidence from cross-state causal-comparative and correlational studies. *Review of Educational Research*, *78*(3), 608–644. doi:10.3102/0034654308324427

Lee, J., & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990-2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis*, *34*(2), 209–231. doi:10.3102/0162373711431604

Lee, J., & Wong, K. K. (2004). The impact of accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes. *American Educational Research Journal*, *41*(4), 797–832. doi:10.3102/00028312041004797

Maslowski, R., Scheerens, J., & Luyten, H. (2007). The effect of school autonomy and school internal decentralization on students' reading literacy. *School Effectiveness and School Improvement*, *18*(3), 303–334. doi:10.1080/09243450601147502

Meyer, H., Tröhler, D., Labaree, D. F., & Hutt, E. L. (2014). Accountability: Antecedents, power, and processes. *Teachers College Record*, *116*(9), 1–12.

Meyer, H.-D., & Benavot, A. (Eds.). (2013). *PISA, power, and policy. The emergence of global educational governance*. Oxford, UK: Symposium Books.

Mok, K. (2003). Decentralization and marketization of education in Singapore. *Journal of Educational Administration*, *41*(4), 348–366. doi:10.1108/09578230310481621

Mok, M. M. C. (2007). Quality assurance and school monitoring in Hong Kong. *Educational Research for Policy and Practice*, *6*, 187–204. doi:10.1007/s10671-007-9027-9

Møller, J. (2009). School leadership in an age of accountability: Tensions between managerial and professional accountability. *Journal of Educational Change*, *10*(1), 37–46. doi:10.1007/s10833-008-9078-6

Newmann, F., King, M., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review*, *67*(1), 41–69.

Ng, P. T. (2010). The evolution and nature of school accountability in the Singapore education system. *Educational Assessment, Evaluation and Accountability*, *22*(4), 275-292.

O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, *72*(3), 293–329.

O'Neill, O. (2013). Intelligent accountability in education. *Oxford Review of Education*, *39*(1), 4–16. doi:10.1080/03054985.2013.764761

OECD. (2014a). *Databases for each year the students took the test: 2012,2009,2006,2003* [Data files and code books]. Retrieved from http://www.oecd.org/pisa/pisaproducts/

OECD. (2014b). *PISA 2012 results: What students know and can do (Vol. I, Revised ed., February 2014): Student performance in mathematics, reading and science*. Paris: OECD Publishing. doi:http://dx.doi.org/10.1787/9789264208780-en

OECD. (2014c). *Strong performers and successful reformers in education: Lesson for Korea*. Paris: OECD Publishing.

Ozga, J. (2013). Accountability as a policy technology: Accounting for education performance in Europe. *International Review of Administrative Sciences*, *79*, 292–309. doi:10.1177/0020852313477763

Park, H. (2013). *Re-evaluating education in Japan and Korea: Demystifying stereotypes*. New York, NY: Routledge.

Park, S.-Y., Kim, J., Oh, E., Jeong, D., & Kim, S. (2014). The effects of school principals' accountability mechanism on school outcomes with TIMSS 2011. *Korean Journal of Educational Administration*, *32*(1), 159–185. (In Korean)

Park, S.-Y., Kim, J., & Sung, K. (2010). The effects of school accountability mechanism on student's achievement. *Korean Journal of Education Administration*, *28*(1), 83–102. (In Korean)

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage Publications.

Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, *92*(5–6), 1394–1415. doi:10.1016/j.jpubeco.2007.05.003

Rouse, C. E., Hannaway, J., Goldhaber, D., & Figlio, D. (2007, December). *Feeling the*

Florida heat? How low-performing schools respond to voucher and accountability pressure (Working Paper No. 13681). Cambridge, MA: National Bureau of Economic Research.

Sahlberg, P. (2008). Rethinking accountability in a knowledge society. *Journal of Educational Change*, *11*(1), 45–61. doi:10.1007/s10833-008-9098-2

Schutz, G., West, M. R., & Wossmann, L. (2007). *School accountability, autonomy, choice, and the equity of student achievement: International evidence from PISA 2003* (OECD Education Working Papers No. 14). Paris: OECD Publishing. doi:10.1787/246374511832

Shin, J., Lee, H., & Kim, Y. (2009). Student and school factors affecting mathematics achievement international comparisons between Korea, Japan and the USA. *School Psychology International*, *30*(5), 520–537.

Springer, M. G. (2008). The influence of an NCLB accountability plan on the distribution of student test score gains. *Economics of Education Review*, *27*(5), 556–563. doi:10.1016/j.econedurev.2007.06.004

Sui-chu Ho, E. (2006). Educational decentralization in three Asian societies: Japan, Korea and Hong Kong. *Journal of Educational Administration*, *44*(6), 590–603. doi:10.1108/09578230610704800

Sung, Y.-K. (2011). Cultivating borrowed futures: The politics of neoliberal loanwords in South Korean cross-national policy borrowing. *Comparative Education*, *47*(4), 523–538. doi:10.1080/03050068.2011.555118

Sung, Y.-K., & Kang, M. (2012). The cultural politics of national testing and test result release policy in South Korea: A critical discourse analysis. *Asia Pacific Journal of Education, 32*(1), 53–73.

Takayama, K. (2013). Untangling the global-distant-local knot: The politics of national academic achievement testing in Japan. *Journal of Education Policy*, *28*(5), 657–675. doi:10.1080/02680939.2012.758833

Weirich, S., Haag, N., Hecht, M., Böhme, K., Siegle, T., & Lüdtke, O. (2014). Nested multiple imputation in large-scale assessments. *Large-scale Assessments in Education*, *2*(1), 1–18.

Wößmann, L. (2005). The effect heterogeneity of central examinations: Evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics*, *13*(2), 143–169. doi:10.1080/09645290500031165

Wößmann, L. (2007). International evidence on school competition, autonomy, and accountability: A review. *Peabody Journal of Education*, *82*(2–3), 473–497. doi:10.1080/01619560701313176

Wößmann, L., Lüdemann, E., Schütz, G., & West, M. R. (2007). *School accountability, autonomy, choice, and the level of student achievement: International evidence from PISA 2003* (OECD Education Working Papers No. 13). Paris: OECD Publishing. Retrieved from http://dx.doi.org/10.1787/246402531617

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91.