School Effectiveness and School Improvement

An International Journal of Research, Policy and Practice

ISSN: 0924-3453 (Print) 1744-5124 (Online) Journal homepage: http://www.tandfonline.com/loi/nses20

School size and student achievement: a longitudinal analysis

Anna J. Egalite & Brian Kisida

To cite this article: Anna J. Egalite & Brian Kisida (2016): School size and student achievement: a longitudinal analysis, School Effectiveness and School Improvement, DOI: 10.1080/09243453.2016.1190385

To link to this article: http://dx.doi.org/10.1080/09243453.2016.1190385

Published online: 06 Jun 2016.

SCHOOL EFFECTIVENESS AND SCHOOL IMPROVEMENT, 2016
http://dx.doi.org/10.1080/09243453.2016.1190385

Routledge
Taylor & Francis Group

School size and student achievement: a longitudinal analysis

Anna J. Egalite[a] and Brian Kisida[b*]

[a]Department of Educational Leadership, Policy, and Human Development, College of Education, North Carolina State University, Raleigh, NC, USA; [b]Department of Education Reform, University of Arkansas, Fayetteville, AR, USA

ABSTRACT
Numerous initiatives by private philanthropies and the US government have supported school size reduction policies as an educational reform intended to improve student outcomes. Empirical evidence to support these claims, however, is underdeveloped. In this article, we draw on information from a longitudinal dataset provided by the Northwest Evaluation Association covering more than 1 million students in 4 US states. Employing a student fixed effects strategy, we estimate how a student's achievement changes as (s)he moves between schools of different sizes. We find evidence that students' academic achievement in math and reading declines as school size increases. The negative effects of large schools appear to matter most in higher grades, which is also when schools tend to be the largest.

ARTICLE HISTORY
Received 5 January 2015
Accepted 4 May 2016

KEYWORDS
Academic achievement; school size; student fixed effects

Introduction

Decades of reform efforts have focused on raising student achievement in the United States. This has been driven, in part, by wide gaps in the quality of education within the United States, as well as the United States' poor showings in international comparisons (Hanushek, Peterson, & Woessmann, 2014). Among these strategies are structural and organizational reforms that target school size policies. School size reduction has been broadly supported by politicians, academics, foundations, educators, and parents as a potentially promising reform for American public schools (Barker & Gump, 1964). In particular, organizations such as the Bill and Melinda Gates Foundation, the Annenberg Foundation, the Carnegie Corporation, and the Pew Charitable Trusts have argued that small schools could improve student outcomes through a variety of mediators, such as community building and increased accountability. By 2002, The Gates Foundation alone had invested over $250 million in grants to facilitate the growth of small schools nationwide, but have since retreated from this reform to focus on other strategies (Vander Ark, 2002).

Compared to other educational policies, the research base of studies attempting to quantify the effects of school size on student achievement is relatively large (for thorough reviews, see Andrews, Duncombe, & Yinger, 2002; Fox, 1981; Leithwood & Jantzi, 2009; Luyten, Hendriks, & Scheerens, 2014), yet much of the research relies upon cross-sectional observations that fail to control for unobserved student characteristics.

In this paper, we examine the relationship between school size and student achievement in both elementary and secondary schools using a research design that focuses on

CONTACT Anna J. Egalite Anna_Egalite@ncsu.edu
*Brian Kisida's current affiliation is The Department of Economics and the Truman School of Public Affairs at the University of Missouri, Columbia, Missouri, USA

individual changes in student achievement as a student moves between schools of varying sizes.

We address two specific research questions:

(1) Is school size significantly related to student achievement outcomes?
(2) Do the achievement effects of school size vary across elementary and secondary school levels?

This study improves upon existing research by addressing two main weaknesses associated with prior studies of school size effects. First, variation in school size is usually not exogenous, and often the observed variation in school size is correlated with other observable and unobservable student traits that may affect student achievement. Our study addresses these estimation challenges by employing a panel dataset that tracks individual students over time and estimates the effect of school size *within* individual students, as opposed to making comparisons across students. This allows us to account for any potentially biasing, time-invariant student characteristics that may be related to the schools they attend, such as underlying ability or motivation. Second, many of the existing studies of school size effects study a single city, such as New York City or Chicago. The effects observed in these locations may not hold in other locations. The database we employ is larger and more geographically comprehensive than those used in previous analyses.

The remainder of the paper is organized as follows: In the next section, we review the existing literature on school size and student achievement and situate our contribution to the literature. We then describe the data for this study and our estimation strategy. Next, we present the results, followed by a discussion of the potential limitations. We conclude with a discussion of the policy relevance of our findings.

Literature review

Evaluations of school size have examined a broad set of outcomes, including student promotion (Fowler & Walberg, 1991), attendance (Darling-Hammond, Ancess, & Ort, 2002; Fowler & Walberg, 1991; Kuziemko, 2006), social relations, (Lee, Bryk, & Smith, 1993), student engagement (Hess & Cytrynbaum, 2002; Kahne, Sporte, & Easton, 2005; Klem & Connell, 2004; Wasley et al., 2000), the quality and variety of courses offered (Lee & Smith, 1995; Monk, 1987; Monk & Haller, 1993), student safety (Rubie-Davies & Townsend, 2007), and the number and type of colleges to which students apply (Schneider, Wyse, & Keesler, 2007).

A particularly convincing set of studies have examined student attainment outcomes using experimental or quasi-experimental methods. Bloom and Unterman (2014) examine results from naturally occurring randomized lotteries and show that attending smaller schools in New York City is associated with a 9.5 percentage point increase in students' graduation rates. Bloom and Unterman's focus on newly created small schools that are also schools of choice, however, makes it difficult to determine if school size, not other related factors, generated the positive outcomes they find. Schwartz, Stiefel, and Wiswall (2012) use an instrumental variables approach to examine the same question in New York City over a longer time period. They show that attending a small high school created during the 1990s is actually associated with a *reduction* in graduation rates, whereas attending a small high school created after 2002 is associated with an *increase* in graduation rates. This disparity may be associated with additional financial and network support provided by non-profits, as well as a competitive application process implemented by the New York City Department of Education in the second time period.

The bulk of studies on the small schools movement assess student test score outcomes. They generally find that student performance declines as school size grows (Deller & Rudnicki, 1993; Eberts, Schwartz, & Stone, 1990; Fowler & Walberg, 1991; Heck & Mayor, 1993; Huang & Howley, 1993; Lee & Loeb, 2000; Lee & Smith, 1995; Walberg & Walberg, 1994; Wasley et al., 2000). Other scholars have noted, however, that the bulk of studies with results favouring smaller schools come from American evaluations, while studies from other countries tend to show mixed or positive effects from larger schools (Luyten, 2014; Scheerens, Hendriks, & Luyten, 2014).

Most existing studies of school size and student achievement rely on statistical controls to address observable differences between the students who attend small schools and those who do not. This approach cannot fully account for unobserved differences in student characteristics, such as differences in student ability, parental motivation, or other relevant characteristics that may simultaneously influence the choice of a small school and student achievement outcomes. As a result, the researchers cannot rule out the possibility that the observed effects of attending a small school are simply the result of pre-existing student differences.

More rigorous studies have used an instrumental variables approach to estimate the effects of school size on student achievement. Barrow, Claessens, and Schanzenbach (2013) use such an approach to capture the impact of attending new small high schools created in Chicago between 2002 and 2006. Using residential proximity as an instrument to predict attendance at a small school, they find no effect on student test scores. Schwartz et al. (2012) also use residential proximity to predict student attendance at a small school. The authors show that attending newer small schools is associated with an increased probability of scoring high on the Regents examinations. Kuziemko (2006) takes advantage of enrolment shocks provided by school openings, closings, and mergers to identify changes in school size. The sample includes third- and sixth-grade math and language outcomes for students in 96 schools in Indiana, tracked over a 3-year time period. Kuziemko shows that a 1 standard deviation (*SD*) increase in school enrolment is associated with a .15 *SD* decrease in math scores.

Absent in the literature on school size has been a student fixed effects specification that examines school size longitudinally using panel data. We use this quasi-experimental methodology because it allows us to statistically control for both observed and time-invariant unobserved student characteristics. With this approach, the effect of school size on student achievement is generated by variation within the same student over time. So long as students sort into schools on the basis of fixed characteristics, this specification allows us to generate rigorous estimates of the relationship between school size and student achievement.

Methods

Sample

We use a large dataset provided by the Northwest Evaluation Association (NWEA) that reports student math and reading achievement on the NWEA Measures of Academic Progress (MAP) assessment in Grades 2 through 10. MAP tests are untimed, computer-adaptive tests. The math tests contain approximately 50 items, and the reading tests contain approximately 40 items. This student-level dataset contains observable characteristics for over one million students in 2,679 unique schools from 2007 through 2011. Data come from four diverse states representing different regions of the United States: the Pacific Northwest, New England, the Rocky Mountains, and the Southeast. These states must remain anonymous due to confidentiality agreements already in place with the testing company.

In addition to student demographic information, the dataset includes student-level test scores on the math and reading MAP assessments. A unique student identifier allows us to track students as they switch between schools. Scores have been standardized by grade/year to have a mean of zero and a standard deviation of 1 to ease interpretation of analyses. The data also include unique school identification codes, merged with data from the National Center for Education Statistic's Common Core of Data. This allows us to match students to specific schools over time. The relevant summary statistics for the merged data files appear in Table 1. The majority of students in these schools are White, followed by Hispanic and Black students. Roughly half of the students are

Table 1. Summary statistics.

Variable	Mean	Standard Deviation	Minimum Value	Maximum Value
Average RIT math score	214.62	22.35	115.44	310.29
Average RIT reading score	207.59	20.04	125.76	269.30
School enrolment	617	398	52	3,297
Average school enrolment of White students	0.59	0.28	0.00	1.00
Average school enrolment of Black students	0.05	0.07	0.00	0.48
Average school enrolment of Hispanic students	0.30	0.28	0.00	0.99
Average school enrolment receiving federally subsidized lunch	0.49	0.25	0.00	0.99
Average student grade level	5.82	2.40	2	10
Number of years present in the data	3.08	1.38	1.00	5.00

Note: Summary statistics calculated for baseline year, 2007. *N* = 1,429 schools. RIT: Rasch UnIT.
Source: Data come from the Northwest Evaluation Association.

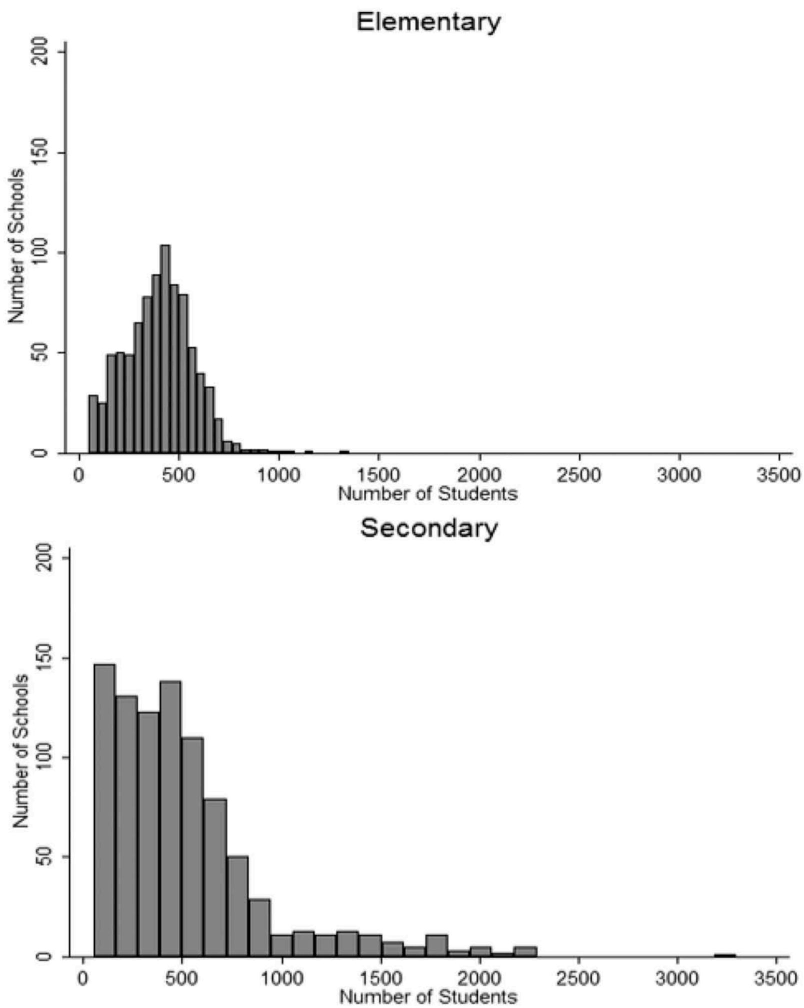

Figure 1. Histogram of the size distribution calculated for baseline year, 2007 (*N*= 1,429 schools).

eligible for the federally subsidized lunch program, and the average grade level is 5.8. Students were nearly evenly distributed across Grades 2 to 10.

There are 2,679 unique schools in our database (Figure 1). School size has a positively skewed distribution. There are a few very large schools (with more than 2,000 students), several very small schools (enrolling fewer than 300 students), with the median school size around 530. Secondary schools have a much higher standard deviation (456 students, as compared to 154).

Empirical strategy

Our primary approach uses student fixed effects to account for unobserved heterogeneity at the student level. This approach implicitly compares students to themselves at different points in time, thus controlling for any time-invariant student characteristics that may be related to student achievement. Because schools are nearly always changing size, every student experienced some amount of variation in school size. Drastic changes in school size, however, are mostly driven by students changing from one school to another. In our sample, 51% of the students observed at baseline switched schools at least once across the observed time period. This is the main source of variation that informs our estimations. Students who switched schools are not substantively different than non-switchers when compared on most observable characteristics (see Appendix 1 Table A1). We do observe that "ever-switchers" have slightly lower test scores, but our analysis accounts for the lower starting point of switchers by measuring changes in individual student achievement over time.

Model (1) estimates the effect of school size (treated as a continuous variable) on students' math and reading outcomes on the MAP assessment in a given year, while accounting for other potentially confounding student characteristics. Our initial approach estimates ordinary least squares regressions, taking the form:

$$Y_{ist} = \beta_0 + \beta_1 Z_s + \beta_2 SchEnrolment_{ist} + \beta_3 SchEnrolment^2{}_{ist} + \phi_t + \tau + \lambda_i + \epsilon_{ist} \quad (1)$$

where Y_{ijst} is the standardized test score of student *i*, in school *s*, during year *t*; Z is a vector of observable school characteristics including urban/rural status, charter status, school-level proportion of minority students, school-level proportion of students in poverty, and type of school (e.g., vocational school); *SchEnrolment* is a continuous variable for school enrolment size, and $SchEnrolment^2$ is that term squared, which allows the distribution to have a quadratic form in case changes in school size do not have a uniform effect across the distribution (Creemers & Kyriakides, 2006); ϕ is a fixed effect for school year; τ is an indicator for US state; λ is a student fixed effect that eliminates bias that may result if school size is related to unobservable time-invariant student characteristics; and ϵ is a stochastic error term, clustered at the school level to take into account the spatial correlation from students nested within schools. β_2 is the primary parameter of interest.

Defining a "small" school

In order to dig deeper into the policy ramifications of school size, our second approach creates indicators for school size. There is no consensus in the literature on how to define a "small" school. For example, Lee and Loeb (2000) define small schools as those with fewer than 400 students and large schools as those with more than 750 students. The Gates Foundation recommends no more than 100 students per grade level, corresponding to 400 students for a typical Grades 9 to 12 high school (Vander Ark, 2002). The U.S. Department of Education set a school size limit of 300 students through its *Small Schools Initiative* (U.S. Department of Education, 2006). Finally, Lee and Smith (1997) recommend that the ideal small high school should enrol between 600 to 900 students. For the purposes of this study, we divide school size into quintiles, calculated within elementary or secondary school categories. This regression equation takes the form:

$$Y_{ist} = \delta_0 + \delta_1 Z_s + \delta_2 SizeQuintile_{ist} + \phi_t + \tau + \lambda_i + \mu_{ist} \quad (2)$$

This model is identical to the previous model, except that in this case, *SizeQuintile* is a series of indicator variables for each school size quintile (the third quintile is the omitted category). Model

Table 2. Mean school enrolment sizes, by quintile and school level.

	All Grades (Grades 2–10)			Elementary (Grades 2–5)			Secondary (Grades 6–10)		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Quintile 1	167	52	228	186	54	258	138	52	189
Quintile 2	311	230	369	320	259	369	283	191	350
Quintile 3	429	370	479	413	370	449	435	351	502
Quintile 4	548	480	623	496	450	541	602	503	718
Quintile 5	1,002	624	3,297	651	542	1,351	1,168	719	3,297

Note: Summary statistics calculated for baseline year, 2007. N = 1,429 schools.
Source: Data come from the Northwest Evaluation Association.

(2) is our preferred model. Table 2 presents mean school enrolment sizes by quintile for the 2007 baseline year. While the size quintile a school is assigned to can vary by year, this is uncommon. On average, only 7.8% of schools change quintiles from one year to the next. Thus, most of the variation in school size quintiles experienced by students in the sample is the result of students switching to a new school.

Results

Table 3 presents the results of Model (1), where school enrolment size is included as a continuous independent variable. These tables have been scaled so that a one-unit change in the coefficient represents an increase of 100 students. In Table 3, we see a small, negative relationship between math achievement and increases in secondary school size. Across all grade levels, this equates to a decrease of –.013 *SD* in student math achievement for every 100 student increase in school size. Breaking results apart by grade level, we see there is no significant relationship between school size and students' math achievement in the elementary grades, but in Grades 6 through 10, an increase of 100 students is associated with a decrease of –.014 *SD* in students' math outcomes.

Table 3 also presents the results from Model (1) for reading achievement. When we aggregate our results across all grade levels, there is no statistically significant relationship with reading achievement. When we break out our findings by grade level, once again we see no significant relationship in the elementary grades, but a significant negative relationship between school size and student reading outcomes in Grades 6 through 10, with an effect size of –.008 *SD*.

Tables 4 and 5 report the results from Model (2), which uses five indicators of school size in place of the single continuous measure. We present three sets of estimates, varying by grade level.

Table 3. Effect of school size on student math and reading achievement: continuous measure.

	Math			Reading		
	All Grades	Elementary (Gs2–5)	Secondary (Gs6–10)	All Grades	Elementary (Gs2–5)	Secondary (Gs6–10)
School Size	–.013***	.001	–.014***	–.005	.009	–.008**
	(.003)	(.006)	(.004)	(.003)	(.005)	(.003)
School Size Squared	.000***	.000	.000***	.000	–.000	.000**
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Adj. R-squared	.845	.825	.876	.804	.815	.804
Obs	2,280,837	1,118,479	1,162,358	2,256,530	1,101,323	1,145,207
Students	1,047,478	602,484	635,545	1,034,862	595,715	625,298

Note: Dependent variable is a student's standardized score on the NWEA MAP math or reading test. Models include controls for year, state, grade, urban/rural status, charter status, school-level proportion of minority students, school-level proportion of students in poverty, and type of school (e.g., vocational school). Standard errors in parentheses, clustered at the school level to account for the spatial correlation of students nested within schools. Results have been scaled so that a one-unit change in the coefficient represents an increase of 100 students. Because the data are longitudinal, some students appear in the elementary analysis and the high school analysis.
$^{**}p < .05$. $^{***}p < .01$, (two-tailed tests).

Table 4. Effect of school size on student math achievement: quintile indicators.

	All Grades	Elementary (Gs2–5)	Secondary (Gs6–10)
Quintile 1	.012	–.001	.003
	(.012)	(.018)	(.019)
Quintile 2	.004	–.004	.010
	(.007)	(.014)	(.013)
Quintile 3	(omitted)	(omitted)	(omitted)
Quintile 4	–.002	.002	–.028
	(.007)	(.012)	(.011)
Quintile 5	–.046***	.010	–.052***
	(.009)	(.014)	(.014)
R-squared	.845	.825	.876
Observations	2,280,837	1,118,479	1,162,358
Students	1,047,478	602,484	635,545

Note: Dependent variable is a student's standardized score on the NWEA MAP math test. Models include controls for year, state, grade, urban/rural status, charter status, school-level proportion of minority students, school-level proportion of students in poverty, and type of school (e.g., vocational school). Standard errors in parentheses, clustered at the school level to account for the spatial correlation of students nested within schools. Because the data are longitudinal, some students appear in the elementary analysis and the high school analysis.
$^{**}p < .05$. $^{***}p < .01$ (two-tailed tests).

Table 5. Effect of school size on student reading achievement: quintile indicators.

	All Grades	Elementary (Gs2–5)	Secondary (Gs6–10)
Quintile 1	–.002	–.011	–.005
	(.012)	(.016)	(.021)
Quintile 2	.002	–.008	.005
	(.007)	(.011)	(.014)
Quintile 3	(omitted)	(omitted)	(omitted)
Quintile 4	–.008	–.004	–.025**
	(.003)	(.010)	(.010)
Quintile 5	–.022**	.007	–.029**
	(.009)	(.012)	(.014)
R-squared	.803	.815	.804
Observations	2,246,530	1,101,320	1,145,207
Students	1,034,862	595,715	625,298

Note: Dependent variable is a student's standardized score on the NWEA MAP reading test. Models include controls for year, state, grade, urban/rural status, charter status, school-level proportion of minority students, school-level proportion of students in poverty, and type of school (e.g., vocational school). Standard errors in parentheses, clustered at the school level to account for the spatial correlation of students nested within schools. Because the data are longitudinal, some students appear in the elementary analysis and the high school analysis.
$^{**}p < .05$.

The first column displays results from models that incorporate all grade levels in the sample; the second column displays results from elementary school models that only include students in Grades 2 through 5. The third column displays results from middle/high school models that include only students in Grades 6 through 10.

Looking at the first column, we see significant negative achievement effects for students in schools in the largest quintiles of school enrolment size, with an effect size of –.046 *SD* for the largest schools. When we examine the elementary grades in isolation, none of the coefficients are statistically significant. Looking at the secondary grades, we see large negative effects for the largest school size quintile with an effect size of –.052 *SD*. Clearly, the negative effects observed in the full sample are being driven by the larger high schools.

Table 5 displays the results of the same model for reading achievement. In the aggregate model, our estimate of the relationship between school size and reading achievement reveals a significant negative coefficient of –.022 *SD* in the largest schools. In the elementary grades, none of the coefficients are statistically significant. Finally, in the secondary grades (6 through 10), the largest schools have a significant, negative effect size of −.025 *SD* in the fourth quintile and −.029 *SD* in the fifth quintile.

Our results reveal two key findings, which point to the importance of school size as a contributing factor to students' academic growth. First, school size is significantly negatively related to student achievement in both math and reading. Across all grades, we find that attending a large school with an enrolment greater than 624 students is significantly negatively related to student academic achievement, relative to attending a school with the median enrolment. Second, these impacts vary by grade level. School size has the largest impact on student learning in Grades 6 to 10, with student achievement significantly declining in schools that enroll more than 712 students.

Discussion

Overall, we identify a negative relationship between increases in school size and student achievement outcomes. Specifically, a school size increase of 100 students is associated with a decline in students' math test scores of approximately 1% of a standard deviation. A further examination of the data reveals that these effects are driven by large middle and high schools (i.e., schools serving Grades 6 through 10), with negative effect sizes increasing to 5% of a standard deviation. It is important to emphasize, however, that the magnitude of these effects is small and the avenues through which they operate unclear. When interpreting these results, policymakers should bear in mind the unique characteristics of schools in their particular community, which may exacerbate or mute the negative influences of school size.

Given the large and comprehensive dataset we have examined here, it is worth considering some general reasons why school size might have a stronger relationship with student achievement outcomes at the secondary school level. It is possible that the self-contained nature of many elementary school classrooms may serve as a protective factor. In these environments, students spend the majority of their time with just one teacher and the same peers, potentially reducing the influence of school size on student outcomes. In a typical secondary school, on the other hand, students are constantly interacting with different teachers and different peers, which may present problems academically and socially as the size of that school increases. Even so, it is possible that individual secondary schools could develop strategies to address the negative potential of increased school size, such as establishing a school-within-a school model (Raywid, 2002).

Another hypothesis is that it may simply be the case that there is a tipping point at which school size begins to have a negative effect on student achievement, and elementary schools rarely pass this threshold. High schools, which are on average larger than elementary schools, are more likely to reach the point at which their size has a negative effect on student achievement.

Nonetheless, our data do not allow us to test these hypotheses about the channels through which school size effects operate, and policymakers should interpret them with that caveat in mind.

Limitations

There are several limitations that should be taken into consideration when interpreting these findings. First, if trends in school enrolment size and student achievement are jointly determined, such as in the case of high-achieving schools that grow in enrolment as parents and families see test scores improving, then a fixed effects approach inadequately addresses this potential bias – though such a bias should show that larger schools are better.[1] At the same time, it could be problematic if, for example, struggling students tended to migrate to larger schools. Such a pattern would make switches to larger schools appear to have a negative effect on student achievement, when in actuality the effect is driven by student sorting.

Second, our data do not allow us to identify when large schools divide into smaller units for instructional and organization purposes, such as a "school-within-a-school" model. It remains to be seen whether such schools can overcome the negative achievement effect of enrolment in a large school. In cases where the data are available to test this hypothesis, researchers should separately measure the achievement impacts for this unique category of schools.

Finally, this study does not test the effect of school size on alternative outcomes such as student well-being or student engagement. These separate, distinct outcomes may be influenced by school size in different ways than the cognitive outcomes examined here (Opdenakker & Van Damme, 2000). For instance, several review studies suggest that school size may matter for non-cognitive outcomes such as social cohesion, safety, well-being, and involvement (Hendriks, 2014; Scheerens et al., 2014).

Conclusion

We find consistent evidence that larger schools have small negative effects on student math outcomes in aggregate models that examine outcomes across all grade levels. These generally small effects are consistent with earlier meta-analyses (Luyten et al., 2014). The results for higher grades in our sample in particular – Grades 6 through 10 – are larger and potentially policy relevant. Two key takeaways are apparent for policymakers deliberating over the efficacy of school size reforms. The first is that, conditional on average achievement and time-invariant characteristics of students, math outcomes are significantly related to the size of a school a student attends. The second key takeaway is that school size matters most in the oldest grades, where schools tend to be larger.

It would be especially informative for future research to unearth the direct causes of school size effects on academic achievement observed here and in other school size studies. There are likely certain characteristics that mediate the effects of school size, such as class size, school climate, or other instructional characteristics (Scheerens et al., 2014). In most existing studies, these "black box" factors are not well identified. To date, only a handful of studies empirically assess the role of mediating variables (Luyten, 2014).

As a policy reform, the once popular movement to create smaller schools in the United States has received less attention in recent years. Our research, however, and other recent studies (e.g., Bloom & Unterman, 2014) suggest that efforts to reduce school size, particularly those serving middle and high school students in the United States, can lead to improvements in student achievement and attainment. As policymakers continually struggle to identify effective school reform strategies from the myriad of available options, the evidence regarding the efficacy of smaller schools continues to grow. If policy follows evidence, the move to decrease school size may gain traction yet again. As yet, however, the specific causes of school size effects are largely unknown. Policymakers and researchers would be wise to advocate for deeper research that provides a clearer understanding of the causes of school size effects before rushing to action.

Note

1. In our second model, however, we mostly remove this potential bias by coding school size as a series of categorical quintiles instead of treating school size as a continuous variable. As a result, variations in school size for a particular student are estimated when a student changes to a school in a different quintile, but not when a single school changes size.

Acknowledgments

The authors would like to thank Nathan Jensen and the staff at the Kingsbury Center at the Northwest Evaluation Association who provided the data for this project. Helpful comments and discussion were also provided by participants at the Association for Education Finance and Policy conference, March 2013.

Notes on contributors

Anna J. Egalite is an Assistant Professor in the Department of Educational Leadership, Policy, and Human Development in the College of Education at North Carolina State University. She holds a PhD in Education Policy from the University of Arkansas and an MEd from the University of Notre Dame's Alliance for Catholic Education. She completed her postdoctoral fellowship in the Program on Education Policy and Governance at Harvard University. Her research focuses on the evaluation of education policies and programs intended to close racial and economic achievement gaps.

Brian Kisida is an Assistant Research Professor in the Department of Economics and the Truman School of Public Affairs at the University of Missouri. He holds a PhD in Education Policy and an MA in Political Science, both from the University of Arkansas. His academic work emphasizes rigorous methodologies and includes studies of school governance, school partnerships with community organizations, school integration, and art and music education.

References

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: Are we any closer to a consensus? *Economics of Education Review*, *21*, 245–262.

Barker, R. G., & Gump, P. V. (1964). *Big school, small school: High school size and student behavior*. Palo Alto, CA: Stanford University Press.

Barrow, L., Claessens, A., & Schanzenbach, D. W. (2013). *The impact of Chicago's small high school initiative* (Working Paper No. WP-13-20). Retrieved from http://www.ipr.northwestern.edu/publications/docs/workingpapers/2013/IPR-WP-13-20.pdf

Bloom, H. S., & Unterman, R. (2014). Can small high schools of choice improve educational prospects for disadvantaged students? *Journal of Policy Analysis and Management*, *33*, 290–319.

Creemers, B. P. M., & Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement*, *17*, 347–366.

Darling-Hammond, L., Ancess, J., & Ort, S. W. (2002). Reinventing high school: Outcomes of the coalition campus schools project. *American Educational Research Journal*, *39*, 639–673.

Deller, S. C., & Rudnicki, E. (1993). Production efficiency in elementary education: The case of Maine public schools. *Economics of Education Review*, *12*, 45–57.

Eberts, R. W., Schwartz, E. K., & Stone, J. A. (1990). School reform, school size and student achievement. *Economic Review*, *26*(2), 2–15.

Fowler, W.J., Jr., & Walberg, H. J. (1991). School size, characteristics and outcomes. *Educational Evaluation and Policy Analysis*, *13*, 189–202.

Fox, W. F. (1981). Reviewing economies of size in education. *Journal of Education Finance*, *6*, 273–296.

Hanushek, E. A., Peterson, P. E., & Woessmann, L. (2014). *Not just the problems of other people's children: U.S. student performance in global perspective*. Cambridge, MA: The Program on Education Policy and Governance at the Harvard Kennedy School.

Heck, R. H., & Mayor, R. A. (1993). School characteristics, school academic indicators and student outcomes: Implications for policies to improve schools. *Journal of Education Policy*, *8*, 143–154.

Hendriks, M. (2014). Research synthesis of studies published between 1990 and 2012. In H. Luyten, M. Hendriks, & J. Scheerens (Eds.), *School size effects revisited: A qualitative and quantitative review of the research evidence in primary and secondary education* (pp. 41–176). Dordrecht, The Netherlands: Springer.

Hess, G. A., Jr., & Cytrynbaum, S. (2002). The effort to redesign Chicago high schools: Effects on schools and achievement. In V. E. Lee (Ed.), *Reforming Chicago's high schools: Research perspectives on school and system level change* (pp. 19–49). Chicago, IL: Consortium on Chicago School Research.

Huang, G., & Howley, C. B. (1993). Mitigating disadvantage: Effects of small-scale schooling on student achievement in Alaska. *Journal of Research in Rural Education*, *9*, 137–149.

Kahne, J. E., Sporte, S. E., & Easton, J. Q. (2005). Creating small schools in Chicago: An early look at implementation and impact. *Improving Schools*, *8*, 7–22.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, *74*, 262–273.

Kuziemko, I. (2006). Using shocks to school enrolment to estimate the effect of school size on student achievement. *Economics of Education Review*, *25*, 63–75.

Lee, V. E., Bryk, A. S., & Smith, J. B. (1993). The organization of effective secondary schools. *Review of Research in Education*, *19*, 171–267.

Lee, V. E., & Loeb, S. (2000). School size in Chicago elementary schools: Effects on teachers' attitudes and students' achievement. *American Educational Research Journal*, *37*, 3–31.

Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education*, *68*, 241–270.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best and for whom? *Educational Evaluation and Policy Analysis*, *19*, 205–227.

Leithwood, K., & Jantzi, D. (2009). A review of the empirical evidence about school size effects: A policy perspective. *Review of Educational Research*, *79*, 464–490.

Luyten, H. (2014). Quantitative summary of research findings. In H. Luyten, M. Hendriks, & J. Scheerens (Eds.), *School size effects revisited: A qualitative and quantitative review of the research evidence in primary and secondary education* (pp. 177–218). Dordrecht, The Netherlands: Springer.

Luyten, H., Hendriks, M., & Scheerens, J. (Eds.). (2014). *School size effects revisited: A qualitative and quantitative review of the research evidence in primary and secondary education*. Dordrecht, The Netherlands: Springer.

Monk, D. H. (1987). Secondary school size and curriculum comprehensiveness. *Economics of Education Review*, *6*, 137–150.

Monk, D. H., & Haller, E. J. (1993). Predictors of high school academic course offerings: The role of school size. *American Educational Research Journal*, *30*, 3–21.

Opdenakker, M.-C., & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. *School Effectiveness and School Improvement*, *11*, 165–196.

Raywid, M. A. (2002). The policy environments of small schools and schools-within-schools. *Educational Leadership*, *59* (5), 47–51.

Rubie-Davies, C. M., & Townsend, M. A. R. (2007). Fractures in New Zealand elementary school settings. *Journal of School Health*, *77*, 36–40.

Scheerens, J., Hendriks, M., & Luyten, H. (2014). School size effects: Review and conceptual analysis. In H. Luyten, M. Hendriks, & J. Scheerens (Eds.), *School size effects revisited: A qualitative and quantitative review of the research evidence in primary and secondary education* (pp. 7–40). Dordrecht, The Netherlands: Springer.

Schneider, B., Wyse, A. E., & Keesler, V. (2007). Is small really better? Testing some assumptions about high school size. In T. Loveless & F. M. Hess (Eds.), *Brookings papers on education policy 2006/2007* (pp. 15–47). Washington, DC: Brookings Institution Press.

Schwartz, A. E., Stiefel, L., & Wiswall, M. (2012). *Do small schools improve performance in large, urban districts? Causal evidence from New York City* (NYU Wagner Research Paper No. 1986092). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1986092

U.S. Department of Education. (2006). *Smaller learning communities program*. Retrieved from http://www2.ed.gov/programs/slcp/index.html

Vander Ark, T. (2002). The case for small high schools. *Educational Leadership*, *59*(5), 55–59.

Walberg, H. J., & Walberg, H. J., III. (1994). Losing local control. *Educational Researcher*, *23*(5), 19–26.

Wasley, P. A., Fine, M., Gladden, M., Holland, N. E., King, S. P., Mosak, E., & Powell, L. (2000). *Small schools: Great strides*. New York, NY: Bank Street College of Education.

Appendix 1. Examining the characteristics of school switchers

Because of the student fixed effects specification we employ, the variation in school enrolment size and school size quintiles is mostly informed by school switchers – students who change schools over the time period examined. To test if this student group is atypical of the tested students in these four states, Table A1 compares switchers to non-switchers across a variety of observable characteristics at baseline (i.e., school year 2006–2007). We find little evidence that switchers are dramatically different from students who stay in the same school for all years. Specifically, 51% of never-switchers and ever-switchers are male. Approximately 60% of both groups are White. Six percent of never-switchers and 4% of ever-switchers are Black. Twenty-six percent of never switchers are Hispanic, compared to 25% of ever-switchers. We do observe that ever-switchers have slightly lower test scores – a math Rasch Unit (RIT) score of 211, compared to 219 and an English Language Arts RIT score of 205, compared to 211 – but our analysis accounts for the lower starting point of switchers by measuring changes in individual student achievement over time.

Table A1. Comparing observable characteristics of student switchers and non-switchers at baseline.

	Never-Switchers		Ever-Switchers		
	Mean	*SD*	Mean	*SD*	Diff
Male	51.34	49.98	51.07	49.99	0.27**
White	57.24	49.47	60.25	48.94	−3.01***
Black	5.98	23.71	4.07	19.77	1.91***
Hispanic	26.49	44.13	25.10	43.36	1.38***
Math RIT score	218.63	23.65	211.36	20.68	7.27***
ELA RIT score	210.54	20.91	205.26	19.02	5.28***

Note: Difference column displays value and significance level from a two-sample difference in means test; **$p < .05$. ***$p < .01$ (two-tailed tests).

Source: Author's calculations from school size & student achievement matched file, using data provided by the Northwest Evaluation Association for 2007.