School consolidation and student achievement[1]

Monique De Haan[2] Edwin Leuven[3] Hessel Oosterbeek[4]

[1]We thank David Figlio, Tarjei Havnes, two anonymous referees, and seminar participants in various places for their helpful comments. Previous versions of this paper circulated under different titles.

[2]University of Oslo. Also affiliated with CESifo, ESOP and the Frisch centre.
E-mail: moniqued@econ.uio.no

[3]University of Oslo. Also affiliated with CEPR, CESifo, ESOP, IZA and Statistics Norway.
E-mail: edwin.leuven@econ.uio.no

[4]University of Amsterdam, TIER, Tinbergen Institute, CESifo and FLACSO-Ecuador.
E-mail: h.oosterbeek@uva.nl

Abstract

What is the effect of school consolidation on student achievement? Theory gives little guidance because possibly positive effects from larger school size can be offset by negative effects from reduced choice and competition. We investigate these issues empirically by analyzing the effects on students' achievement of a consolidation reform that took place in Dutch primary education in the mid 1990s. The reform was implemented by increasing the minimum required school size, leading to an increase in actual school size and a reduction in the number of schools. For identification we exploit variation between municipalities. We find that an increase in the minimum required school size of 10% has a small positive effect on student achievement of 0.72% of a standard deviation. Further analysis indicates that this effect can be mainly attributed to the increase in actual school size; reduced competition and choice do not seem to have harmed student achievement. We also find no evidence that the consolidation effect is driven by reduced school segregation or the elimination of small schools that were – given their size – underperforming.

JEL-codes: I21, I22, H75, D40

Keywords: School size, school choice, school consolidation

1 Introduction

Due to increased urban concentration and an aging population, many local policymakers and school administrators are confronted with shrinking student numbers in their municipalities and schools (e.g. Ritsema van Eck et al., 2013; Schlömer, 2013). They then face the question whether to continue operating their small schools or to merge several small schools into one larger school. Both options have potential advantages and disadvantages. Consolidation reduces the fixed costs, may reduce school segregation, and the increased school size may involve economies of scale. Operating several small schools, on the other hand, gives students more schools to choose from and may enhance competition between schools. Optimal consolidation decisions therefore require knowledge of the effects of consolidation on student outcomes.

The analysis of school consolidation effects is challenging due to endogeneity concerns: schools that merge or are closed are likely to differ from schools that do not. Some schools that merge may do so because they are not performing very well. Other schools may merge because they have an ambitious and inspiring school principal. A few previous studies have addressed endogeneity issues when examining the impact of school consolidation on student outcomes. Using variation in the timing of school (district) consolidation across states in the US, Berry and West (2010) find that students educated in states with smaller schools obtained higher returns to education and completed more years of schooling.[1] Liu et al. (2010) combine a difference-in-differences approach with propensity score matching methods and find that primary school mergers that took place in 2002 in China did not harm students' performance measured in 2006.[2]

The current paper exploits a consolidation reform that took place in the Netherlands. The reform was announced in 1992, implemented in 1994 and became fully operative in 1996. It increased average school size from 160 to 200 students and reduced the number of primary schools in the Netherlands by about 15% in just a few years. The consolidation process was implemented in the form of a revision of the minimum required school size rules. Before the

[1] Other studies that investigate school (district) consolidation in the US include Berry (2006); Brasington (2003, 1999, 1997).

[2] Brasington (1997) relies on the inclusion of covariates to address endogeneity concerns.

reform, the minimum required school size was a step function of the number of inhabitants in a municipality. After the reform, the minimum required school size was a smooth function of student density in a municipality. This reform caused substantial variation in the change of the minimum required school size across municipalities.[3]

In our empirical strategy we calculate the predicted change in the minimum required school size in a municipality using the number of inhabitants and student density at baseline (in 1992). Our estimations include municipality fixed effects so that we essentially exploit predicted changes in the minimum required school size over time that are triggered by the change in rules. Variation in this change comes from differences between municipalities in the ratio of student density and the number of inhabitants in 1992. The identifying assumption is that the change in a municipality's minimum required school size and the change in the average residual student outcomes in that municipality are uncorrelated. To weaken this identifying assumption trends in student outcomes are allowed to vary across groups of municipalities of different size. We report results from falsification tests to corroborate the validity of our conditional mean independence assumption.

We focus on two cohorts of students: The last cohort of students who finished primary school before the policy was announced, and the first cohort of students who enrolled in primary school after the policy was fully implemented. By excluding students who were exposed to the transition phase, we are focusing on the systemic effects of school consolidation on students with full exposure.[4] We find that the consolidation reform led to increased student achievement on a nationwide exit exam. The average increase of the minimum required school size from 62 to 101 students is estimated to have increased test scores by 3.5% of a standard deviation.

We also investigate pathways through which the consolidation reform resulted in a positive impact on students' test scores. School consolidation increases average school size and reduces the number of schools. These two channels may have opposite effects on achievement. An

[3]To avoid confusion, our analysis deals with the effects of school consolidation; it is not informative about school district consolidation.

[4]This contrasts with a growing literature that focusses on estimating the effect of school mergers and school closures on student achievement for students who were exposed to the transition phase of a merger or school closure (Liu et al. 2010; Engberg et al. 2011; Brummet 2012; Beuchert et al. 2015). In one of our specifications, we do, however, also include cohorts that were exposed to the transition stage.

increase in school size is – at least in some size-range – likely to improve student outcomes through scale effects. A larger scale makes it for example possible to have a full-time principal, hire a remedial teacher, or better smooth teacher absence. A certain scale is also required to avoid the placement of students from multiple grades in the same class. Beyond a certain size further increases in school size may affect student outcomes negatively. Large schools are often believed to be anonymous and to reduce the involvement of parents.

Much of the literature on the effects of school size on student outcomes is correlational and the results are mixed. In their review of the literature on school size, Andrews et al. (2002) conclude that "there is little convincing evidence in the United States on how consolidation actually affects school districts in the long-run." Kuziemko (2006) also notes the lack of consensus in the school size literature and explains this by "the empirical weakness that the existing papers share", namely omitted variable bias. A recent study that addresses this is the recent study by Schwartz et al. (2013), who study the effect of small high schools in New York City on student outcomes using distances to schools as instrumental variables. While their OLS estimates point to positive effects of both new and existing small schools on outcomes, their IV estimates indicate that this can only be interpreted as causal for the new small schools. The IV estimate of the effect of old small schools is even significantly negative. New small schools are schools founded under the Small School Movement and are not only different in size than large schools but have also significantly more funding, making it unclear whether the better performance of students in these schools is due to size or to other inputs.

While school size increases, school consolidation also reduces the number of schools. A reduction in the number of schools can influence students outcomes negatively. Fewer available schools restricts parents' opportunities to find a school that fits the specific needs of their child. Moreover, having fewer schools also implies that schools face less competitive pressure to perform as well as they can. While these theoretical arguments in favor of more choice and competition are strong, the empirical evidence is mixed. Studies that point to positive effects of increased choice and competition include Hoxby (2000a), Gibbons et al. (2008) and Böhlmark and Lindahl (2013). Studies that find zero or even negative effects include Cullen et al. (2006), Hsieh and Urquiola (2006) and Rothstein (2007).

School consolidation therefore involves a potential trade-off between the impact of school size and the number of schools.[5] This trade-off has not been investigated before in the empirical literature and, as mentioned by Berry and West (2010, fn.9), school consolidation estimates will capture the sum of these effects. To make progress in understanding this trade-off we estimate a 2SLS model where both variables are treated as endogenous regressors. To identify the separate effects of school size and the number of schools we exploit changes in the minimum required school size and municipality level variation in fertility as instruments. The first instrument results from the consolidation reform. The second instrument builds on Hoxby (2000b), and assumes that variation in fertility in a municipality over time has no independent effect on student achievement. The results of this analysis indicate that the positive reform effect can be mainly attributed to positive effects of increased average school size. Reduced competition and choice do not seem to have harmed student achievement.

With fewer schools there is less scope to sort into different schools on the basis of social background. Reduced sorting may be beneficial for the achievement of some students. We consider this possibility by estimating the effect of the change in the minimum required school size on an index of relative heterogeneity in terms of students' social background. The results do not support the sorting explanation. We also report evidence that our findings cannot be attributed to the closure of weakly performing small schools.

The remainder of the paper continues as follows. The next section provides information about the Dutch education system and describes the details of the change in the minimum school size rule. Section 3 introduces the data and Section 4 provides details of our estimation strategy. Section 5 presents and discusses the empirical results. Section 6 summarizes and concludes.

[5] Although typically ignored, studies of school competition that exploit variation in the number of schools ideally need to take this trade-off into consideration.

2 Institutional context and reform

2.1 The Dutch education system

Freedom of education is a key feature of the Dutch (primary) education system. This has two elements: Parents can freely choose a school for their child, and there is the right to start new schools and freedom to organize the teaching in schools. Parents' school choice is not restricted by where they live (there are no school catchment areas), or by their income. With the exception of a few cases of orthodox religious schools, primary schools do not select students. Parents can therefore enroll their children in the school of their choice. Currently there are about 7,000 primary schools in the Netherlands, serving around 1.6 million students. For most students the nearest primary school is within walking distance. For 59% (89%) of students the nearest school is less than 500 meters (1 kilometer) from their home (Bunschoten, 2008).

About two thirds of students are enrolled in publicly-funded private schools, the rest in public schools. The main difference between publicly-funded private schools and public schools lies in their governance. Private schools are governed by private school boards and the public schools by municipalities. Both types of schools receive funding from the central government through a "money follows student"-mechanism. The funding of a school is thus based on the number of students enrolled. Schools that receive funding through this voucher system are not allowed to charge school fees. Privately-funded primary schools are virtually non-existent in the Netherlands. The Dutch Education Inspectorate monitors whether schools comply with laws and regulations.

2.2 Minimum school size rules

Primary schools in the Netherlands must have a minimum number of students to receive funding. Until 1994 the minimum required school size depended on the population size of the municipality where the school was located as follows

$$\text{min\_size}(pop_{mt}) = \begin{cases} 50 & \text{if } pop_{mt} < 25,000 \\ 75 & \text{if } 25,000 \leq pop_{mt} < 50,000 \\ 100 & \text{if } 50,000 \leq pop_{mt} < 100,000 \\ 125 & \text{if } pop_{mt} \geq 100,000 \end{cases} \tag{1}$$

where pop_{mt} is the population size of municipality m in year t. If a school fell below the minimal school size for three years in a row, its funding was stopped at the beginning of the next school year in the case of a privately-run school, or it was closed down in the case of a publicly-run school.[6]

In the 1980s there were many small schools, and there were concerns about their ability to provide education of sufficient quality (Ministry of Education, 1990). Moreover, the funding system was such that, in addition to the vouchers, each school also received a lump-sum transfer. Many small schools were thus more expensive than a smaller number of large schools.

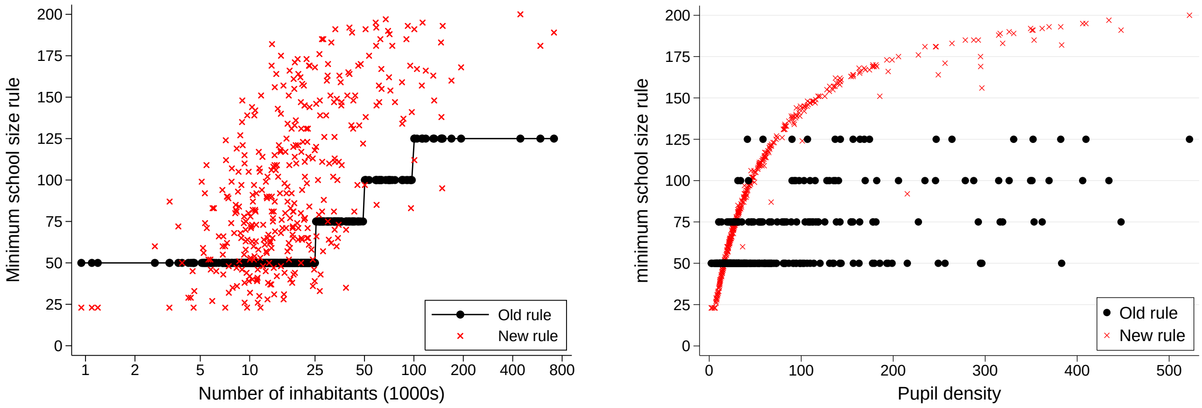
In 1992 the government announced that, from 1994 onwards, the minimum school size rule would no longer be based on population size but on the student density of the municipality, as follows

$$\text{min\_size}(d_{mt}) = \frac{d_{mt}}{0.25 + 0.0045 d_{mt}} \tag{2}$$

where d_{mt}, student density in municipality m in year t, is defined as the number of inhabitants between 4 and 11 years old divided by the size of the municipality in square kilometers.

Figure 1 shows scatter plots of the old and new minimum school size rules for the municipalities in our data. The left panel (a) shows a scatter plot of the old and new rules against the number of inhabitants. The dots connected by the line show the old minimum school size rule, and each dot represents a municipality. The crosses show, for each municipality the minimum school size under the new rule, which ranges from 23 students to 200 students. As can be seen

[6] If a privately-run school stops receiving funding from the government this means in practice that it has to close down. The only source of funding is government funding since schools are not allowed to charge school fees.



(a) **(b)**

Figure 1. Old (a) and new rules (b) by number of inhabitants and student density in 1992

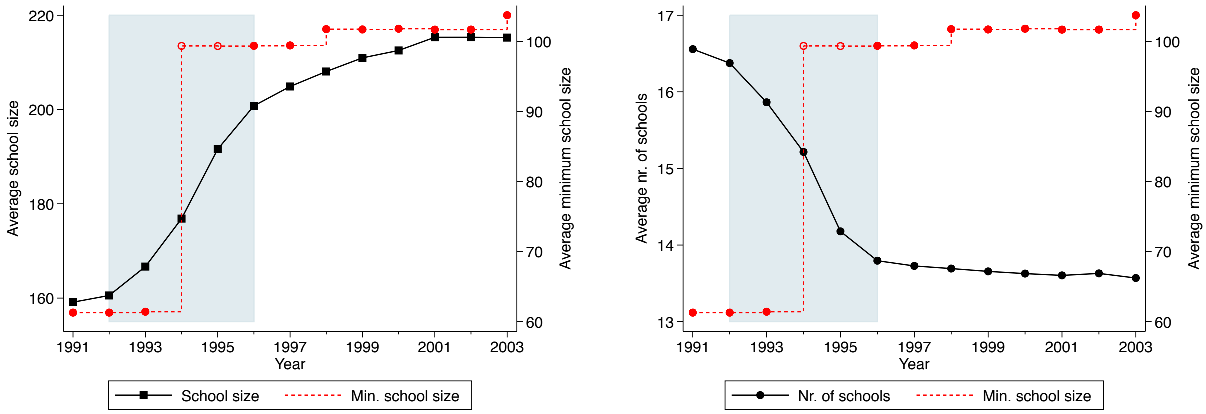
in the first panel, there is substantial variation in the new minimum school size between municipalities with the same number of inhabitants, and thus the same minimum school size, before the announcement of the reform.

The right panel (b) in Figure 1 shows for the different municipalities the old and new rules against student density. The new minimum school size rule, indicated by the crosses, illustrates the relation with student density.[7] Municipalities with the same student density have the same minimum school size after the reform but, as the dots show, the old minimum school size was often very different for municipalities with the same student density.

Although the new rule was implemented in 1994 there was a grace period of two years. Consequently, no schools were forced to close down or stopped receiving funding in the school years 1994/95 and 1995/96. From 1996 onwards, schools that fell below their threshold during at least two out of the previous three years were either closed down (in the case of public schools), or stopped receiving funding (in the case of private schools) from the beginning of the following school year.

On average the minimum required school size increased due to the reform. In many municipalities, this led to an increase in the average actual school size and a reduction in the number of schools. Figure 2a shows the average minimum school size by year and the mean of the average

[7]There are some "outliers" which are due to the fact that if the student density was more that 500 it was set at 500 and when the size of the municipality was smaller than 10 km^2 it was set at 10.



(a) **(b)**

Figure 2. Average school size (a) and number of schools (b) over time

actual school size per municipality by year.[8] Figure 2b shows the average minimum school size by year and the mean number of schools per municipality by year. The vertical axes on the left show the scales of the mean of average actual school size and the mean of the number of schools. The vertical axis on the right shows the mean of the minimum school size.

Until 1993 the average minimum school size was just above 60 students. In 1994, after the implementation of the law, the average minimum school size jumped to about 100. At the same time average actual school size increased and the number of schools declined. In 1991 the average school size in a municipality was on average close to 160 students, but after 1992 average school size rapidly increased to 200 in 1996, after which it increased further at a much slower rate. The average number of schools per municipality was 16.5 in 1991, but after 1992 it declined until 1996, after which it stabilized around an average of 13.5 schools per municipality. The total number of schools went down from 8,362 schools in 1992 to 7,136 schools in 1996, a decline of 15% in a period of four years.[9]

Most schools that were below the new rule in 1994, merged with another school instead of being closed down. Of the 8,362 primary schools in 1992, 2,201 schools were part of a merger

[8]Per municipality we calculate the average actual school size. The figure reports the average municipal average school size by year.

[9]The reform affected privately-run and publicly-run schools similarly. We do not have access to schools' denomination in our micro data, but aggregate statistics show that the share of publicly-run schools remained approximately constant between 1992 and 1997, 35 percent vs. 33.5 percent.

in the five years between 1992 and 1996. Most of these mergers were real mergers and not just administrative mergers as is reflected by the fact that the number of school locations declined to 7,163 in 2003.

3 Data

We use data from various sources. As outcome variable we use standardized test scores. At the end of primary school students take a nationwide exit exam developed by Cito, the national institute for educational testing and measurement. This test determines for a large part the type of secondary school a student will attend after primary school. Although the test is not compulsory, most students take it.[10] The test consists of multiple choice questions covering language, math and information processing. Answer sheets are sent to the testing institute where they are marked. We standardize the scores by year in the total population of test takers, so that results can be interpreted in terms of standard deviation units of the annual test score distribution.

School level data, such as information about school size and the share of minority students, are obtained from the Dutch Ministry of Education. Data at the municipality level are obtained from Statistics Netherlands. The minimum school size rules are taken from official publications of the government.

The main analysis compares two cohorts of students. The first cohort is the one that finished primary school in 1992, just before the reform was announced. The second cohort is the one that enrolled in primary school just after the large reduction in the number of schools and which was therefore not exposed to the transition stage of a merger or school closure. This cohort finished primary school in 2003. In one of the additional specifications we also include data from the transition years (exam years 1999-2002).

Some municipalities (not schools) merged during our observation period. If a municipality merger led to changes in student density it could trigger changes in the minimum required school size. This means that mergers between municipalities might give another source of

[10]Table A1 in the appendix shows that the reform did not affect test taking behavior.

Table 1. Summary statistics

	1992		2003	
	mean	SD	mean	SD
Outcome variable:				
Standardized test score	-0.06	0.53	-0.04	0.46
Number of schools	3,517		4,133	
Municipality Characteristics:				
Minimum required school size	62.2	21.1	101.1	47.6
Average school size	162.5	46.5	216.5	74.9
Number of schools	17.3	21.1	14.4	18.2
Number of children age 4-11 ($\times$1000)	3.0	4.7	3.2	5.2
Number of inhabitants ($\times$1000)	31.9	59.9	34.2	62.5
% with low educated Dutch parents	35.5	13.8	12.6	7.2
% with benefits (welfare, unemployment)	3.3	1.7	2.8	1.3
% with low educated migrant parents	5.2	6.3	6.3	6.6
Number of municipalities	345		345	

Note: The statistics in this table are based on information from the 80% of municipalities (and the schools in these municipalities) not involved in a consolidation of municipalities. The standardized test score is the score on the compulsory parts on the nationwide exit exam. The score is standardized at the student level for the population of test takers.

variation in the supply of schools. It is unclear whether this variation is exogenous. Moreover, a merger between municipalities may also lead to other changes related to local governance. We therefore only consider the municipalities that were not part of a merger between 1992 and 2004. About 20% of the municipalities in 2004 are the result of a merger; the analysis is therefore based on the remaining 80% of the municipalities. The top panel of Table 1 shows that the mean of the standardized test scores in this reduced sample is slightly below that of the total population. It also shows that the standard deviation at the school level is around half of the standard deviation at the student level.

Table 1 reports summary statistics separately for the years 1992 and 2003. The bottom panel of the table shows the substantial changes that took place in the average school size (up by 33%), the average number of schools (down by 17%), and the average minimum school size (up by 63%). The numbers of inhabitants and students increased only by about 7%.

The top part of the table also shows that the number of schools that participated in the test increased by almost 18%, and since school size increased, the proportion of students who took the test increased substantially. This is not problematic for our analysis as long as the change in test-taking students is unrelated to the predicted changes in the minimum required school size. The results in Table A1 in the appendix show that this is indeed the case.

4 Empirical approach

Before we present the empirical results, we first discuss the empirical specifications that we estimate, and the assumptions needed in order to give a causal interpretation to our estimates.

The consolidation reform was implemented through a change in the minimum school size requirements. This is also the variable which varies substantially between municipalities. We start therefore by estimating the following relationship:

$$\bar{y}_{smt} = \delta \cdot \ln(z_{mt}) + x'_{mt}\beta + w'_{smt}\pi + \lambda_m + \mu_t + \varepsilon_{smt} \tag{3}$$

where $\bar{y}_{smt}$ is the mean test score of students enrolled in school s in municipality m in year t. z_{mt} is the minimum required school size based on the number of inhabitants and student density in 1992:

$$z_{mt} = \begin{cases} \text{min\_size}(pop_{m1992}), & t = 1992 \\ \text{min\_size}(d_{m1992}), & t = 2003 \end{cases}$$

Note that z_{m2003} is the *predicted* minimum required school size based on student density at baseline (in 1992). It therefore does not pick up changes in the number of inhabitants or student density over time. Because our specification includes municipality fixed effects, we essentially exploit predicted changes in the minimum required school size due to the change in rules. We use the logarithm of the minimum required school size because the effect of a given change is likely to be very different when the initial minimum required size is 50 students from when it is 200 students. We can interpret δ therefore as the effect of a 100% change in the minimum required school size on student test scores. We include municipality fixed effects λ_m, municipality time-varying controls x_{mt}, and school characteristics w_{smt}. The vector x_{mt} consists of the logarithm of the number of inhabitants, the share of students with low educated Dutch parents, the share of students with low educated migrant parents, and the share of people on unemployment or welfare benefits (all measured at the level of municipality m in year t). Inclusion of these variables reduces the residual variance and proxies for changes in the socioeconomic conditions in a municipality over time. The vector w_{smt} consists of the share of students with low educated Dutch parents and the share of students with low educated migrant parents in the school. The idiosyncratic error term ε_{smt} is clustered at the municipality level. The year fixed effects μ_t control for changes over time which are common across municipalities, such as education policies which are implemented nationwide.

The variation in the change in the minimum required school size comes from differences between municipalities in the ratio of student density and the number of inhabitants in 1992. If these municipalities also differ in other (unobserved) characteristics, then this is captured by the municipality fixed effects. Our identifying assumption is thus that the change in a municipality's minimum required school size and the change in the average residual achievement of students in that municipality are uncorrelated.[11] To weaken this identifying assumption we allow the year fixed effects to vary across municipalities of the following population sizes (in 1000s): (0-25), [25- 50), [50- 100) and [100 or more). In addition we allow the trend in test scores to differ between urban and rural municipalities.

Although our conditional mean independence assumption is counterfactual and, like any identifying assumption, cannot be tested, we can devise some falsification tests. If changes in the minimum required school size correlate with changes in characteristics of the (student) population of a municipality that are systematically related to potential outcomes, then this will invalidate our assumption. To investigate this possibility we estimate equation (3), but replace the dependent variable by i) the (log) number of inhabitants, ii) the (log) number of students, iii) the share of inhabitants on benefits (welfare or unemployment), and iv) the share of students with low educated migrant parents. If the coefficient on the minimum required school size is statistically significant, then changes in the minimum required school size are confounded by changes in the underlying population. Table 2 shows that this is not the case. Changes in these

[11] Whereby the change in minimum required school size and the change in average residual student achievement are measured as deviations from a nation wide trend, since we include year fixed effects in our specification.

Table 2. Exogeneity checks

	ln(min school size)
Dependent variable:	
ln(Number of inhabitants)	0.007 (0.013)
ln(Number of students)	0.006 (0.018)
% with benefits	-0.114 (0.101)
% with low educated migrant parents	-0.003 (0.002)
Number of schools	7,650

Note: Standard errors in parentheses are clustered at the municipality level. All regressions include municipality and year fixed effects. Year fixed effects are allowed to vary across four groups of municipalities of different size, based on the number of inhabitants (in 1000s): (0-25), [25- 50), [50- 100) and [100 or more). The regressions also include an urban specific trend. ***/**/* significant at the 1/5/10 percent level.

municipality characteristics between 1992 and 2003 are not systematically related to changes in the minimum required school size.[12] [13]

Figure 2 suggests that there are (at least) two relevant channels through which a change in the minimum required school size affects student outcomes: a change in actual school size and a change in the number of schools per municipality. Minimum required school size as an instrument for, say, actual school size, may violate the exclusion restriction as it also could affect student outcomes through a change in the number of schools per municipality.

To make progress on the examination through which channels school consolidation influences student outcomes, we need a second instrumental variable. In the spirit of, amongst others, Hoxby (2000c) and Urquiola (2006), we use variation in the number of students at the municipality level as an additional instrumental variable. This instrument varies at the municipality level over time, and should therefore address concerns about within municipality sorting at a given point in time. We continue to control for the (log) number of inhabitants, and for this

[12] Another scenario under which the conditional mean independence assumption might fail is if closeness to the minimum required school size gives an incentive to schools to perform and if, in addition, the share of schools in a municipality that are close to the norm is correlated with the minimum required school size. As a robustness check we include the share of schools in a municipality that are at most 25 students (about 10 percent of average school size) away from the norm. The coefficient for this variable is not significantly different from zero. Moreover, the estimated coefficient of $\ln(z_{mt})$ remains the same.

[13] In 1992 a student was classified as having "low educated Dutch parents" if one of his parents was low educated, while after 1993 *both* parents had to be low educated. Due to this change in definition there is a substantial reduction in the share of students with low educated Dutch parents between 1992 and 2003. We can therefore not use this variable to check the conditional mean independence assumption in Table 2. We do use it as a control variable in Tables 3 and 4. Including it as a control variable does not substantially affect the estimated coefficient on the supply of schools, but it does increase the precision of our estimates. Both tables report results with and without control variables.

instrument to be valid we require shocks to the population share of students over time to have no independent effect on achievement.[14]

To second-stage equation that we now estimate is:

$$\bar{y}_{smt} = \delta_1 \cdot \ln(n_{mt}) + \delta_2 \cdot \ln(\text{school size}_{smt}) + x'_{mt}\beta + \lambda_m + \mu_t + \varepsilon_{smt} \quad (4)$$

where n_{mt} is the number of schools in municipality m in year t, and school size$_{smt}$ is the size of school s in municipality m in year t. δ_1 and δ_2 are the parameters of interest.[15]

5 Results

We present the results in four subsections. We start with the estimates of the reform which come from the reduced form where student achievement is regressed on minimum required school size. In Section 5.2 we then present the results from the instrumental variable analysis that aims to disentangle the reform effect into its effect through a change in average actual school size and its effect through a change in the number of schools per municipality. In Section 5.3 we inquire whether the reform affected segregation patterns. Finally, in Section 5.4 we assess whether our results can in part be explained by the closure of small schools that were, given their size, underperforming.

5.1 Reduced form results

Table 3 presents estimates for eight different specifications of the reduced form equation (equation (3)). The basic specification in column (1) indicates that a 10% increase in the minimum

[14]We control for changes in the composition of the student population by including changes in the share of minority students in a municipality over time and changes in the share of students with low educated Dutch parents as control variables.

[15]Note that equation (4) does not represent an education production function in which all relevant inputs appear on the right-hand side. In practical applications some relevant inputs will be unobserved and subsumed in the error term. Correlation between the unobserved variables and observed inputs will then lead to biased estimates of the parameters of the observed inputs in the production function (e.g. Krueger, 1999). Our aim is to obtain consistent estimates of specific parameters of the education production function, the parameters connected to the number of schools and school size. School and municipality characteristics and fixed effects are included in equation (4) because the independence assumption is more credible conditional on these variables, and to improve precision. Our approach concurs with many recent contributions in the economics of education literature (cf. Meghir and Rivkin, 2011).

required school size increases test scores on average by 0.55% of a standard deviation. The point estimates increase somewhat when we add different sets of control variables, and are always positive and statistically significant. According to the most elaborate specification in column (5) a 10% increase in the minimum required school size results in an increase in school average test scores of 0.72% of a standard deviation. When we exclude observations from the largest twenty municipalities, the effect of a 10% increase in the minimum required school size on test scores is 0.75%. When we include data for the transition years, the effect also remains the same (compare columns (5) and (7)). Finally, when we consider student-level data instead of school-level data, the estimate becomes somewhat smaller, but is still 0.57% of a standard deviation for a 10% increase in the minimum required school size, and is significant at the 5%-level.

5.2 Disentangling effects of school size and supply

When the minimum required school size increases, actual average school size increases while the number of schools decreases. The estimates that we presented in the previous subsection are the result of both changes. In this subsection we attempt to disentangle the separate contributions of the increase in average school size and the reduced number of schools. We do this using the instrumental variable approach (equation (4)) outlined in Section 4.

Table 4 reports results for the same specifications as in Table 3. Table 4 consists of three panels: Panels A and B report the first-stage results for the two endogenous variables, while panel C reports the 2SLS results. The results in panels A and B show strong and consistent relationships between school size and the number of schools on the one hand, and the minimum required school size and the number of children between 4 and 11-years-old on the other hand. The number of schools is negatively related to the minimum required school size and positively to the number of school-aged children. School size is positively related to both instrumental variables.

The first column in panel C shows a positive and statistically significant effect of (log) school size on test scores, while the effect of the (log) number of schools is negative, but not

Table 3. The effect of the school consolidation reform on test scores

	School level							Pupil level
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(min. school size)	0.055**	0.060**	0.083***	0.071***	0.072***	0.075***	0.076***	0.057**
	(0.025)	(0.024)	(0.025)	(0.025)	(0.025)	(0.026)	(0.025)	(0.026)
Controls								
- Municipality		✓	✓	✓	✓	✓	✓	✓
- Size trends[a]			✓	✓	✓	✓	✓	✓
- Urban trend[b]				✓	✓	✓	✓	✓
- School					✓	✓	✓	✓
Sample								
- No Big 20[c]						✓		
- All years[d]							✓	

Note: Note: Dependent variable is school average standardized test score. Standard errors in parentheses are clustered at the municipality level. All regressions include municipality and year fixed effects. Municipality controls: ln(municipality size), municipality share of students with low educated dutch parents, municipality share of students with low educated migrant parents, share with unemployment/welfare benefits.[a]Trend in test scores allowed to differ between municipalities with number of inhabitants (1000s) of respectively (0-25), [25-50), [50-100) and [100+).[b]Trend in test scores allowed to differ between urban and rural areas. School controls: share of students with low educated Dutch parents, share of students with low educated migrant parents. [c]Excludes 20 largest cities. [d]Uses all available years, pre: 1992, post: 1999–2003. ***/**/* significant at the 1/5/10 percent level.

Table 4. Scale versus supply effects – 2SLS Estimates

	School level							Pupil level
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. First-stage: ln(# schools) as dependent variable								
log(min. school size)	-0.174***	-0.167***	-0.172***	-0.177***	-0.177***	-0.159***	-0.188***	-0.197***
	(0.019)	(0.020)	(0.019)	(0.020)	(0.020)	(0.018)	(0.019)	(0.023)
log(# children aged 4–11)	0.401***	0.311***	0.193**	0.179**	0.179**	0.197**	0.191***	0.173**
	(0.073)	(0.085)	(0.080)	(0.078)	(0.078)	(0.078)	(0.069)	(0.087)
Angrist-Pischke (weak id) F	113.38	78.21	89.39	90.39	89.65	92.94	109.81	90.42
B. First-stage: ln(school size) as dependent variable								
log(min. school size)	0.171***	0.164***	0.167***	0.173***	0.175***	0.165***	0.195***	0.179***
	(0.024)	(0.025)	(0.025)	(0.026)	(0.026)	(0.026)	(0.023)	(0.026)
log(# children aged 4–11)	0.510***	0.584***	0.691***	0.706***	0.703***	0.718***	0.737***	0.809***
	(0.091)	(0.103)	(0.103)	(0.103)	(0.099)	(0.100)	(0.084)	(0.115)
Angrist-Pischke (weak id) F	65.10	54.34	58.32	58.75	63.62	64.22	104.88	60.00
C. Second-stage: School average standardized test scores as dependent variable								
log(# schools)	-0.082	-0.086	-0.190	-0.162	-0.160	-0.198	-0.229	-0.153
	(0.112)	(0.144)	(0.164)	(0.157)	(0.158)	(0.169)	(0.143)	(0.143)
log(school size)	0.236**	0.287**	0.312**	0.256**	0.259**	0.284**	0.174*	0.167
	(0.114)	(0.138)	(0.137)	(0.123)	(0.118)	(0.125)	(0.099)	(0.104)
Kleibergen-Paap (weak id) rk F	45.08	47.25	57.61	58.40	65.36	57.73	139.35	51.64
N	7,650	7,650	7,650	7,650	7,650	5,810	24,136	182,509
Controls								
- Municipality		✓	✓	✓	✓	✓	✓	✓
- Size trends[a]			✓	✓	✓	✓	✓	✓
- Urban trend[b]				✓	✓	✓	✓	✓
- School					✓	✓	✓	✓
Sample								
- No Big 20[c]						✓		
- All years[d]							✓	

*Note:*Standard errors in parentheses are clustered at the municipality level. All regressions include municipality and year fixed effects. Municipality controls: ln(municipality size), municipality share of students with low educated dutch parents, municipality share of students with low educated migrant parents, share with unemployment/welfare benefits.[a]Trend in test scores allowed to differ between municipalities with number of inhabitants (1000s) of respectively (0-25), [25-50), [50-100) and [100+).[b]Trend in test scores allowed to differ between urban and rural areas. School controls: share of students with low educated Dutch parents, share of students with low educated migrant parents. [c]Excludes 20 largest cities. [d]Uses all available years, pre: 1992, post: 1999–2003. ***/**/* significant at the 1/5/10 percent level.

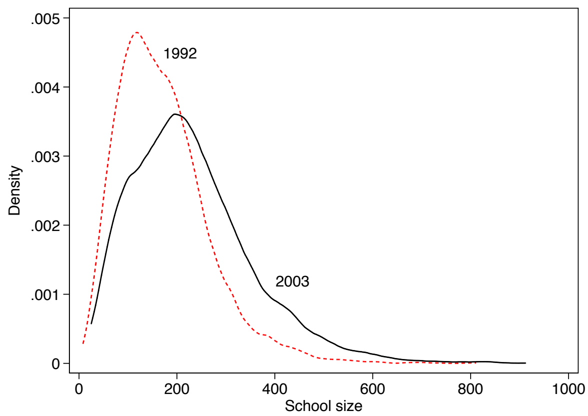


Figure 3. Kernel density of school size in 1992 and 2003

significantly different from zero. A 10% increase in school size improves test scores by 2.4% of a standard deviation. When we include control variables (columns (2) to (5)), the estimates of the effect of school size remain very similar and always statistically significant. The estimate of the effect of the number of schools remains insignificant, but becomes more negative when control variables are included. Excluding the largest twenty municipalities has almost no impact on the estimates (column (6)). When we also include the post-reform cohorts that were only partially exposed to the reform, the estimate of the effect of school size becomes somewhat smaller, while the estimate of the effect of the number of schools become a bit more negative (column (7)). Finally, when we analyze the student-level data instead of the school-level data, the estimates keep the same sign but the estimated effect of school size becomes somewhat smaller and is no longer significantly different from zero at conventional levels (p=0.108).

Figure 3 shows kernel density plots of school size for the years 1992 and 2003. Average school size increased from 162 students per school in 1992 to 216 students in 2003. This increase in average school size can explain our findings if there are returns to scale. As a first step to investigate the potential for returns to scale, we report the relationship between school size and actual school practices. In the year in which the reform was announced (1992), a survey of 177 primary school principals asked, among other things, about the organization of teaching in their schools and about the schools' contacts with parents.[16] In Table 5, we report results

[16]This survey is part of a larger project that collected data from primary school students, their parents and teachers; the Landelijke Evaluatie Onderwijsvoorrangsbeleid. We only use information from the representative

Table 5. Associations between school size and school characteristics

	ln(school size)	
Dependent variable:		
Share of time the principal spends on teaching	-0.139***	(0.023)
School has at least one full time director	0.233***	(0.044)
Share of classes that contain students from multiple grades	-0.575***	(0.045)
School has a remedial teacher	0.181**	(0.052)
School is involved in extracurricular parent-student activities	-0.020	(0.053)
School has agreement with parents about:		
- parents attending parent-teacher meetings	-0.002	(0.062)
- discussing the school report of the students	0.002	(0.023)
- time spend on the different subjects	0.087	(0.061)
- minimum goals that students should achieve	-0.035	(0.066)

Note: Results are based a survey among principals of 177 schools in 1992. Robust standard errors in parentheses. ***/** significant at the 5/10 percent level.

from regressions of these organizational features and school-parent contacts on the size of the school. Each row comes from a separate regression. The results show that larger school size is associated with: i) less teaching by the principal, ii) a higher probability of having at least one full-time director, iii) fewer classes with students from multiple grades, and iv) a higher probability of having a remedial teacher. At the same time a larger school is not associated with less involvement of the parents with the school (as indicated in the second half of the table). These findings are consistent with the view that just before the reform was announced, increases in school size could increase the efficiency of the teaching process without harming parental involvement.

The results in this subsection suggest that a substantial share of the positive effect of consolidation reform that we reported in Section 5.1, can be attributed to the increase in average school size. We find no evidence that the reduction in school choice and competition has been harmful for student achievement.

5.3 Did changes in the number of schools affect school segregation?

An alternative explanation for the positive effect of the school consolidation reform on test scores is that it reduced school segregation, which in turn may have had a positive effect on

sample of schools.

student outcomes. With fewer schools there is less scope for similar students to separate themselves from other students. If low-ability students benefit more from being placed with high-ability students than high-ability students do, less segregation will raise the average of students' outcomes. To assess whether this mechanism can explain our findings, we analyze in this subsection the impact of school consolidation on school segregation.

For each primary school we know the number of students in each of the following three categories: i) students with low-educated migrant parents, ii) students with low-educated Dutch parents and iii) all students who do not fall into the first two categories. Given this division of students by socioeconomic status we can calculate the Blau index of heterogeneity $H = 1 - \sum_{r=1}^{R} S_r^2$ where R is the number of groups and S_r is the share of group r in the population. On the basis of the division into groups, we can calculate the heterogeneity index for each school and for the municipality in which the school is located. By taking the ratio of the two we obtain a measure of relative heterogeneity (cf Urquiola, 2005). As a robustness check we also calculate the (relative) heterogeneity index on the basis of two groups: i) students with low educated migrant parents and ii) all other students.[17]

Table 6 shows the estimates of the effect of minimum required school size on the two measures of relative heterogeneity. These results are based on the specification in column (4) of Table 3 that includes all control variables except the variables measuring the shares of students with low educated Dutch (migrant) parents. These variables are omitted because they are used to construct the heterogeneity indexes that are the dependent variables in this analysis.

The top panel of the table reports descriptive statistics. The heterogeneity index is commonly interpreted as the probability that two individuals selected at random belong to different groups. When we consider three groups, this probability is on average equal to 0.36 when measured at the school level, while the average equals 0.44 when measured at the municipality level. This indicates that schools are more homogenous than municipalities, which is summarized in

[17] The reason for this robustness test is that the definition of students with low-educated Dutch parents changed between 1992 and 2003, also see footnote 13. Since this change in the definition of the second category applied to all schools in all municipalities in the Netherlands this should be captured by the year fixed effect and therefore not affect the results. The index based on the second division is not affected by the change in the definition of one of the categories.

Table 6. Effect of minimum required school size on sorting

Sample:	All schools				Schools with test results			
# Groups in Index:	3		2		3		2	
Summary statistics:	*Mean*	*S.D.*	*Mean*	*S.D.*	*Mean*	*S.D.*	*Mean*	*S.D.*
Heterogeneity school	0.361	0.197	0.124	0.149	0.358	0.192	0.126	0.149
Heterogeneity municipality	0.442	0.139	0.167	0.141	0.425	0.137	0.162	0.138
Relative heterogeneity index	0.843	0.568	0.862	1.226	0.872	0.565	0.902	1.212
Results:								
log(min. school size)	0.024		-0.024		0.021		-0.075*	
	(0.016)		(0.028)		(0.020)		(0.040)	
Nr. observations (schools)	11,403		11,391		7,648		7,644	

Note: Dependent variable is school heterogeneity (relative to municipality heterogeneity). Standard errors in parentheses are clustered at the municipality level. All regression include municipality and year fixed effects. Control variables: ln(municipality size), share with unemployment/welfare benefits. Trend in heterogeneity indexes allowed to differ between municipalities with number of inhabitants (1000s) of respectively (0-25), [25-50), [50-100) and [100+). Trend in heterogeneity indexes allowed to differ between urban and rural areas.***/**/* significant at the 1/5/10 percent level.

the relative heterogeneity index being smaller than unity. The results in the bottom panel of the table show that changes in the minimum school size rules did not have a significant effect on the sorting of students in terms of socioeconomic status. The estimates are small and not significantly different from zero. Table 6 also shows results for the sub-sample of schools for which we observe test results. The estimate using three groups in the heterogeneity index is very similar to the estimate obtained using the full sample of schools. The estimate using two groups is larger for the reduced sample of schools and marginally significant. The estimated effect is however negative, indicating that an increase in the required school size reduces the heterogeneity of schools relative to the heterogeneity of the municipality. Since increasing the required school size reduces the number of schools and thereby the scope for sorting, the sign of this estimate is the opposite of the expected direction.[18] We therefore interpret the results in Table 6 as indicating that sorting cannot explain our findings.

5.4 Can the scale effect be explained by weaker schools being shut down?

The results so far indicate that the closure of small schools was an important channel through which the consolidation reform improved test scores. It matters, however, whether the small

[18] Comparing the estimates in columns (4) and (5) in Table 3 also shows that controlling for school composition does not affect the estimated effect.

schools that were closed were performing poorly because they were small, or whether they were small because they were performing poorly. In the first case, the school size effect reflects genuine scale effects. In the second case, the school size effect reflects the disappearance of underperforming (and therefore small) schools.

If students are indeed less likely to enroll in poorly performing schools, then we should see that (before the reform) the smallest schools in a municipality are also the worst performing schools in the municipality when keeping absolute school size constant. To test this hypothesis we regress average student achievement of schools in 1992 on relative school size while conditioning on absolute school size and municipality fixed effects. This regression exploits the fact that a school with for example 125 students is a large school in a municipality with low student density, but a small school in a municipality with high student density. If it is indeed the case that low quality schools are small because students choose to enroll in a higher quality school in the municipality, we should see that a school with 125 students belongs to the group of worst performing schools in a municipality with high student density and is among the best performing schools in a municipality with low student density.

We estimate the following specification by OLS

$$\bar{y}_{sm1992} = \alpha + \beta \cdot \text{Relative size}_{sm,1992} + \delta \cdot ln(\text{school size}_{sm,1992}) + \lambda_m + \varepsilon_{sm,1992} \qquad (5)$$

where relative size is defined as

$$\text{Relative size}_{sm1992} = \frac{\text{school size}_{sm,1992} - \min \text{school size}_{m,1992}}{\max \text{school size}_{m,1992} - \min \text{school size}_{m,1992}} \qquad (6)$$

and ranges from 0 to 1, with the value 0 for the smallest school in a municipality and the value 1 for the largest school in the municipality. Conditional on school size relative school size is likely to be correlated with (unobserved) municipality characteristics, for example student density; we therefore include municipality fixed effects.

Table 7 shows the results that we obtain when estimating equation (5). We see that there is a significant positive relation between absolute school size and school performance, but no sig-

Table 7. Relative school size and average test scores of schools in 1992

	All municipalities		Excl. 20 largest municipalities (in 2003)	
Relative size	-0.035	-0.037	-0.021	-0.027
	(0.077)	(0.069)	(0.086)	(0.080)
ln(school size)	0.182***	0.101**	0.159***	0.104**
	(0.048)	(0.043)	(0.056)	(0.053)
School controls		✓		✓
N	3,520	3,520	2,747	2,747

Note: Dependent variable is average standardized test score by school in 1992. Standard errors in parentheses are clustered at the municipality level. All regressions include municipality fixed effects. School controls: share with low educated dutch parents and share with low educated migrant parents. ***/**/* significant at the 1/5/10 percent level.

nificant relation between relative school size and average school performance. This shows that it is unlikely that the positive effect of the reduction in the supply of schools can be explained by an increase in school quality due to the disappearance of low quality schools. We interpret this as showing that although the small schools were performing worse than the larger schools, this was not because of lower school quality, but more likely because of returns to scale.

6 Conclusion

This paper studies the effects of a large school consolidation reform in Dutch primary education on student achievement. The reform was implemented as an increase of the minimum required school size. On average the minimum required school size in municipalities increased from 62 to 101 students. Our estimates indicate that such an increase had a small positive effect on student achievement of 3.5% of a standard deviation ($0.072 \times \ln(101/62)$).

Using an instrumental variable strategy we disentangled this reform effect into an effect due to a change in average actual school size and an effect due to a change in the number of school per municipality. The results of this analysis point to the change in average actual school size as the most important factor. This finding is supported by an analysis of survey data collected just before the reform, and which shows that increases in school size could improve the efficiency of the teaching process without harming parental involvement. We find no evidence that the reduction in school choice and competition has been harmful for student achievement. Further results show that the reform effect cannot be attributed to changes in segregation between schools and by the closure of small and weak schools.

The results in this paper show that, compared to the baseline school size range, the larger schools that resulted from school consolidation led to improved student achievement. Although cost saving, school consolidation can therefore be a beneficial policy response to shrinking student numbers. To judge the potential benefits of school consolidation reforms in other contexts, both the range of school size changes as well as the potential trade-off between school size and competition need to be considered. Informed policy would require further work about these causal impacts for different school size ranges and institutional contexts.

References

Andrews, M., Duncombe, W., and Yinger, J. (2002). Revisiting economies of size in American education: are we any closer to a consensus? *Economics of Education Review*, 21(3):245–262.

Berry, C. (2006). School consolidation and inequality. *Brookings Papers on Education Policy*, pages 49–75.

Berry, C. and West, M. (2010). Growing pains: The school consolidation movement and student outcomes. *Journal of Law, Economics, and Organization*, 26:1–29.

Beuchert, L. V., Humlum, M. K., Nielsen, H. S., and Smith, N. (2015). The short-term effects of school consolidation on student achievement: Evidence of disruption? Unpublished working paper.

Böhlmark, A. and Lindahl, M. (2013). Independent schools and long-run educational outcomes: Evidence from Sweden's large scale voucher reform. IZA Discussion Paper No. 6683, Institute for the Study of Labor (IZA).

Brasington, D. (1997). School district consolidation, student performance and housing values. *Journal of Regional Analysis and Policy*, 27:43–54.

Brasington, D. (1999). Joint provision of public goods: the consolidation of school districts. *Journal of Public Economics*, 73(3):373–393.

Brasington, D. (2003). Size and school district consolidation: Do opposites attract? *Economica*, 70(280):673–690.

Brummet, Q. (2012). The effect of school closings on student achievement. Unpublished working paper.

Bunschoten, B. (2008). Hoe ver woon ik van ... *Bevolkingstrends, 2e kwartaal 2008*, pages 19–22.

Cullen, J. B., Jacob, B., and Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica*, 74:1191–1230.

Engberg, J., Gill, B., Zamarro, G., and Zimmer, R. (2011). Closing schools in a shrinking district: Do student outcomes depend on which schools are closed? *Journal of Urban Economics*.

Gibbons, S., Machin, S., and Silva, O. (2008). Choice, competition, and pupil achievement. *Journal of the European Economic Association*, 6(4):912–947.

Hoxby, C. (2000a). Does competition among public schools benefit students and taxpayers? *American Economic Review*, 90(5):1209–1238.

Hoxby, C. (2000b). Peer effects in the classroom: Learning from gender and race variation. *NBER working paper*.

Hoxby, C. M. (2000c). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics*, 115(4):1239–1285.

Hsieh, C. and Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics*, 90(8-9):1477–1503.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114(2):497–532.

Kuziemko, I. (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review*, 25(1):63–75.

Liu, C., Zhang, L., Luo, R., Rozelle, S., and Loyalka, P. (2010). The effect of primary school mergers on academic performance of students in rural China. *International Journal of Educational Development*, 30(6):570–585.

Meghir, C. and Rivkin, S. (2011). *Handbook of the Economics of Education*, chapter Econometric methods for research in education, pages 1–87. Elsevier.

Ministry of Education (1990). Schaal en kwaliteit in het basisonderwijs. Technical report, Ministerie van Onderwijs en Wetenschappen: Projectgroep Schaalvergroting Basisonderwijs, Staatsuitgeverij: Zoetermeer.

Ritsema van Eck, J., van Dam, F., de Groot, C., and de Jong, A. (2013). *Demografische ontwikkelingen 2010-2040; Ruimtelijke effecten en regionale diversiteit*. Planbureau voor de Leefomgeving, Den Haag.

Rothstein, J. (2007). Does competition among public schools benefit students and taxpayers? Comment. *American Economic Review*, 97(5):2026–2037.

Schlömer, C. (2013). Trends in population, households and labour force. Technical report, Federal Institute for Research on Building, Urban Affairs and Spatial Development within the Federal Office for Building and Regional Planning.

Schwartz, A., Stiefel, L., and Wiswall, M. (2013). Do small schools improve performance in large, urban districts? Causal evidence from New York City. *Journal of Urban Economics*, 77:27–40.

Urquiola, M. (2005). Does school choice lead to sorting? Evidence from Tiebout variation. *American Economic Review*, 95(4):1310–1326.

Urquiola, M. (2006). Identifying class size effects in developing countries: Evidence from rural Bolivia. *Review of Economics and Statistics*, 88(1):171–176.

Table A1. Change in rules and change in share of schools participating in Cito test

	School level							Pupil level[e]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(min. school size)	0.021	0.010	0.028	-0.001	-0.001	-0.019	-0.006	0.0103
	(0.033)	(0.037)	(0.038)	(0.036)	(0.036)	(0.038)	(0.029)	-0.0322
N	7,650	7,650	7,650	7,650	7,650	5,810	24,139	182,509
Controls								
- Municipality		✓	✓	✓	✓	✓	✓	✓
- Size trends[a]			✓	✓	✓	✓	✓	✓
- Urban trend[b]				✓	✓	✓	✓	✓
- School					✓	✓	✓	✓
Sample								
- No Big 20[c]						✓		
- All years[d]							✓	

Note: Dependent variable is share of schools in municipality that participates in the Cito test. Standard errors in parentheses are clustered at the municipality level. All regressions include municipality and year fixed effects. Municipality controls: ln(municipality size), municipality share of students with low educated dutch parents, municipality share of students with low educated migrant parents, share with unemployment/welfare benefits.[a]Trend in test scores allowed to differ between municipalities with number of inhabitants (1000s) of respectively (0-25), [25-50), [50-100) and [100+).[b]Trend in test scores allowed to differ between urban and rural areas. School controls: share of students with low educated Dutch parents, share of students with low educated migrant parents. [c]Excludes 20 largest cities. [d]Uses all available years, pre: 1992, post: 1999–2003. [e]Dependent variable is number of test-taking pupils divided by the number of 11-year-olds in the municipality. ***/**/* significant at the 1/5/10 percent level..