IFN Working Paper No. 1472, 2023

Independent-School Competition and Sweden's Performance in TIMSS

Gabriel Heller-Sahlgren

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Independent-school competition and Sweden's performance in TIMSS

Gabriel Heller-Sahlgren[*][†][‡][§]

This paper analyses the effects of independent-school competition on Sweden's performance in TIMSS, an international low-stakes test in mathematics and science among students in year 8. Exploiting variation in independent-school enrolment shares across counties over time, it finds that increasing competition has improved TIMSS scores, an impact that appears only after 2003 and is driven by for-profit schools. The results suggest that competition both slowed down Sweden's performance decline between 1995 and 2011 as well as contributed to its improving scores between 2011 and 2019. A simulation based on the estimates indicates that Sweden's average score in TIMSS 2019 would have been 20 points, or 0.24 standard deviations, lower without the expansion of the independent-school sector.

Keywords: independent-school competition; student performance; TIMSS

JEL codes: I20, L33

[*] Research Institute of Industrial Economics, Stockholm, Sweden. Email: gabriel.heller.sahlgren@ifn.se
[†] London School of Economics, Houghton Street, London, WC2A 2AE, United Kingdom.
[‡] The author thanks Marco Bertoni, Henrik Jordahl, Björn Öckert, Mikael Lindahl, and Olmo Silva for useful comments and suggestions. Research funding from the former IFN Research Programme 'The Economics of the Service Sector' and the Confederation of Swedish Enterprise is gratefully acknowledged.
[§] The author acknowledges a potential conflict of interest due to partial research funding from an interested party. In addition, he has benefitted from financial support for other research from the Hans and Barbara Bergstrom Foundation, a non-profit research foundation that holds a minority stake in a for-profit Swedish independent-school provider, and has acted as consultant on the implementation of research-based practices for several Swedish municipal and independent education providers.

1. Introduction

In the on-going debate regarding the effects of independent-school competition on student achievement, Sweden's 1992 voucher reform is an important case study. The reform allowed independent operators – both non-profit and for-profit – to set up new schools, thus injecting competition in the education system. As a result of the reform, the share of students attending independent schools increased considerably. In 1992, about 1 percent of all students in primary- and lower-secondary school attended independent schools, a figure that increased to 15 percent in 2019 (NAE 2021). Moreover, the great majority of the increase in independent-school enrolment shares can be attributed to the expansion of the for-profit sector, making Sweden rather unique in a comparative perspective (OECD 2019).

The effects of the voucher reform have been fiercely debated ever since it was implemented, but especially since Sweden's scores in international tests, such as PISA and TIMSS, began to fall in the late 1990s and early 2000s. As international test scores declined, and independent-school enrolment shares increased, the latter became a key explanation for the former in the international public debate (e.g. Fisman 2014). Indeed, the OECD (2019, p. 75) advised the Swedish government that 'school choice and competition likely weakened school performance over time'. This hypothesis continues to attract attention, even though Sweden's international test scores have improved considerably since around 2010, especially when adjusting for the mechanical effects of changing demographics due to immigration (Heller-Sahlgren 2022).

While research finds positive effects of Swedish independent-school competition on domestic performance outcomes (e.g. Böhlmark and Lindahl 2015), questions often arise whether these effects reflect bona-fide changes in achievement or merely grade inflation (e.g. Wennström 2020). This is because all marking and grading in the Swedish context are performed internally by teachers in a non-moderated criterion-referenced system, thus also making it questionable whether the effects of competition would be replicated in externally-marked international tests. Of course, effects in this respect may also differ more generally, since such tests partly reflect other types of knowledge and skills compared with domestic performance measures.

This paper analyses how independent-school competition has affected Sweden's performance in TIMSS, an international low-stakes test in mathematics and science among students in year 8. This is possible because I have obtained information on the geographical location of the participating schools in all survey rounds in which Sweden has participated – 1995, 2003, 2007, 2011, 2015, and 2019 – enabling me to link the schools to variables that vary between regions over time. To analyse the effects of competition, the paper investigates whether within-county changes in the average share of students attending independent schools predict within-county changes in TIMSS scores, once adjusting for measures of student background and other relevant variables. In this way, it is possible to study whether performance trends differ across counties depending on the evolution of independent-school competition over time.

The results display that increasing independent-school competition has had positive effects on student performance in TIMSS: a 10 percentage-point increase in the average independent-school enrolment share at the county level improves TIMSS scores by about 20–25 points, equivalent to roughly 0.24–0.30 standard deviations. These effects can only be detected from 2007 onwards, which is reasonable given that it took time for the independent sector to expand more than marginally. The findings also show that competition both decreased Sweden's performance decline between 1995 and 2011 and contributed to its improvements between 2011 and 2019. The effects are primarily driven by the expansion of for-profit schools, which is plausible given that they explain the great majority of the independent sector's growth overall. Also, independent-school competition raises the probability that students reach the intermediate, high, and advanced international performance benchmarks. The point estimate is also positive when analysing the probability that students reach the low (basic) benchmark, but it does not reach statistical significance.

Interestingly, the findings are essentially identical when adjusting for the direct association between independent-school attendance and TIMSS performance, and when excluding all participating students in independent schools, suggesting that municipal and independent schools benefit similarly from competition. Using the estimates to simulate Sweden's counterfactual performance trend without increasing competition suggests that its average scores in TIMSS 2019 would have been 20 points, or 0.24 standard deviations, lower without the expansion of the independent-school sector.

The effects of competition withstand several sensitivity checks and are supported by a placebo test in treatment: changes in the share of the variation in independent pre-school enrolment shares that cannot be predicted by independent-school enrolment shares are unrelated to changes in TIMSS scores. This is expected if the findings are driven by school competition, rather than other time-varying differences between counties that correlate with changes in preferences for private providers more generally.

Overall, the paper thus supports prior research in finding positive effects of independent-school competition on lower-secondary student performance in Sweden. Importantly, little suggests that estimates in prior research are biased upwards due to unreliable domestic performance metrics. On the contrary, the effects found in this paper are larger than the impact found in research analysing domestic outcomes. A comparison suggests this is primarily because I analyse low-stakes international test scores, which both neutralise all factors behind different marking and grading standards and partly reflect different types of knowledge and skills compared with domestic outcomes, and to some extent also because I estimate effects at the county level rather than the municipal level.

The paper proceeds as follows. Section 2 discusses the background and prior research on the Swedish voucher reform; Section 3 describes the data and methodology; Section 4 presents the results and compares them with findings in prior research; Section 5 uses the empirical estimates to simulate Sweden's counterfactual performance trend without increasing competition; and Section 6 concludes.

2. Background and prior research

The effects of independent-school competition on international test scores should theoretically depend on parental demand for the knowledge and skills captured by such scores. If parents perceive the marginal utility of the knowledge and skills to be high, we may expect independent schools to invest in quality and compete along those lines, thus providing an impetus for overall improvements (e.g. Hoxby 2003). On the other hand, since quality at a general level is difficult to observe, and parents may not value it sufficiently, independent schools may have few incentives to compete on this basis (see MacLeod and Urquiola 2012, 2019). It is thus not straightforward to predict the effects of independent-school competition on international test scores.

The Swedish 1992 voucher reform is an interesting case study for evaluating such effects. Since the reform, independent providers – both for-profit and non-profit – are allowed to own and operate schools as well as receive funding from the municipality in which students reside, calculated from the per-student expenditure in the municipalities' own schools.[1] This also means that there are economic consequences for the latter, since their budgets decrease when students opt for independent schools. At the same time, since 1997, independent schools are not allowed to charge any top-up fees, meaning that they are entirely dependent on public funding. At the primary and lower-secondary level, which are the focus of this paper, such schools are also not allowed to select students based on ability, although they are allowed to admit students on a first-come, first-served basis (Holmlund et al. 2014).

As a result of the voucher reform, the enrolment shares of students attending primary- and lower-secondary independent schools increased from about 1 percent in 1992 to 15 percent in 2019 (NAE 2021). While most of the initial (small) wave of independent schools were set up by non-profit organisations, this changed considerably after just a few years. In fact, the great majority of the increase in independent-school enrolment shares can be explained by the growth of for-profit schools (Holmlund m.fl. 2019). Consequently, increasing independent-school competition following the voucher reform has been primarily driven by for-profit actors, an important difference compared with other similar reforms (OECD 2019).

At the same time as independent-school competition increased, Sweden initially saw a considerable decline in international test scores. As displayed in Figures 1 and 2, performance in mathematics and science in TIMSS plummeted between 1995 and 2011. However, since then, student performance has increased, a change that is primarily driven by students with (mostly) a Swedish background. In both subjects, although still lower than in 1995, performance among these students was higher in 2019 than it was in 2003 by a statistically significant margin. In this group, fully 50–60% of the performance decline between 1995 and 2011 had been erased by 2019. This also means that a substantial part of the remaining difference between Sweden's average scores in 2019 and 1995 – 27 percent in mathematics and 50 percent in science – can be attributed to the mechanical effects of changing demographics due to immigration.[2]

[1] The exact funding formula varies across municipalities and has also changed over time. In 1992, the voucher covered a minimum of 85 percent of the average per-student municipal expenditure. It was decreased to 75 percent in 1995, and then increased to 100 percent in 1997 (Holmlund et al. 2014). However, when calculating the voucher after 1997, municipalities were still able to deduct a share of the average per-student expenditure, intended to be equivalent to the costs associated with having the legal responsibility for compulsory education. Since 2010, municipalities are no longer allowed to make such deductions (Proposition 2008/09:171).

[2] This trend is replicated between 2000 and 2018 in PISA, another international test that is carried out at the end of lower-secondary school in Sweden. Indeed, there was a considerable performance decline

Figure 1. Sweden's mathematics scores in TIMSS

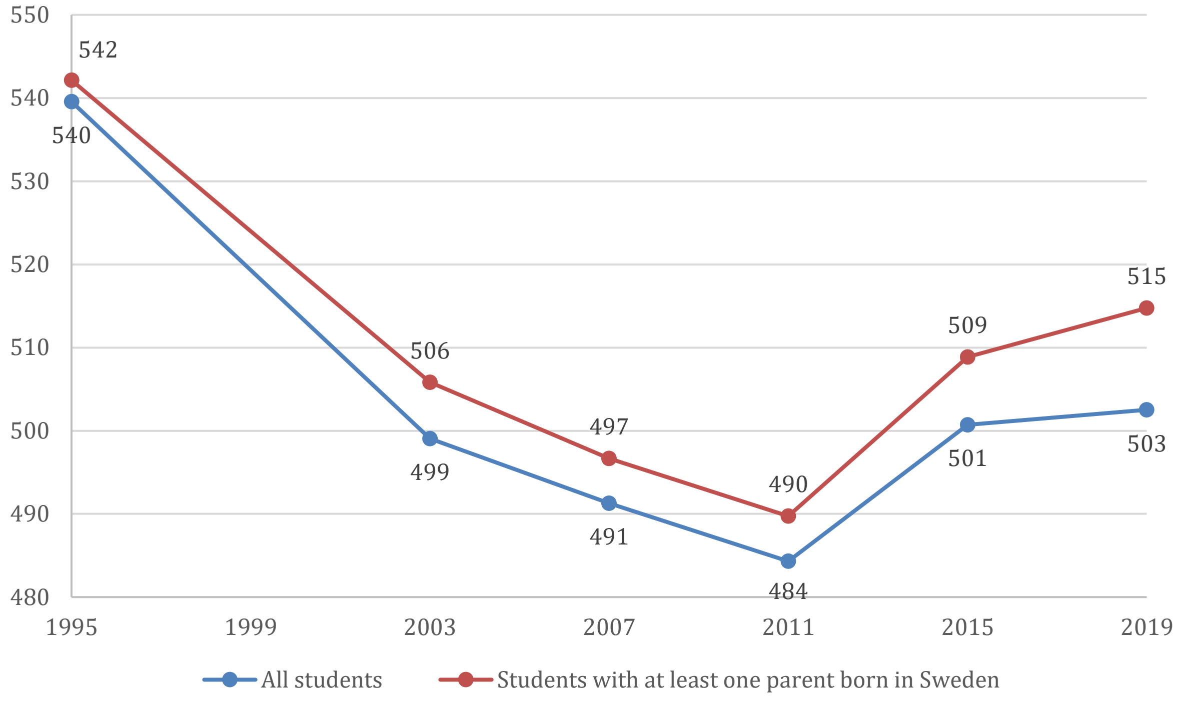
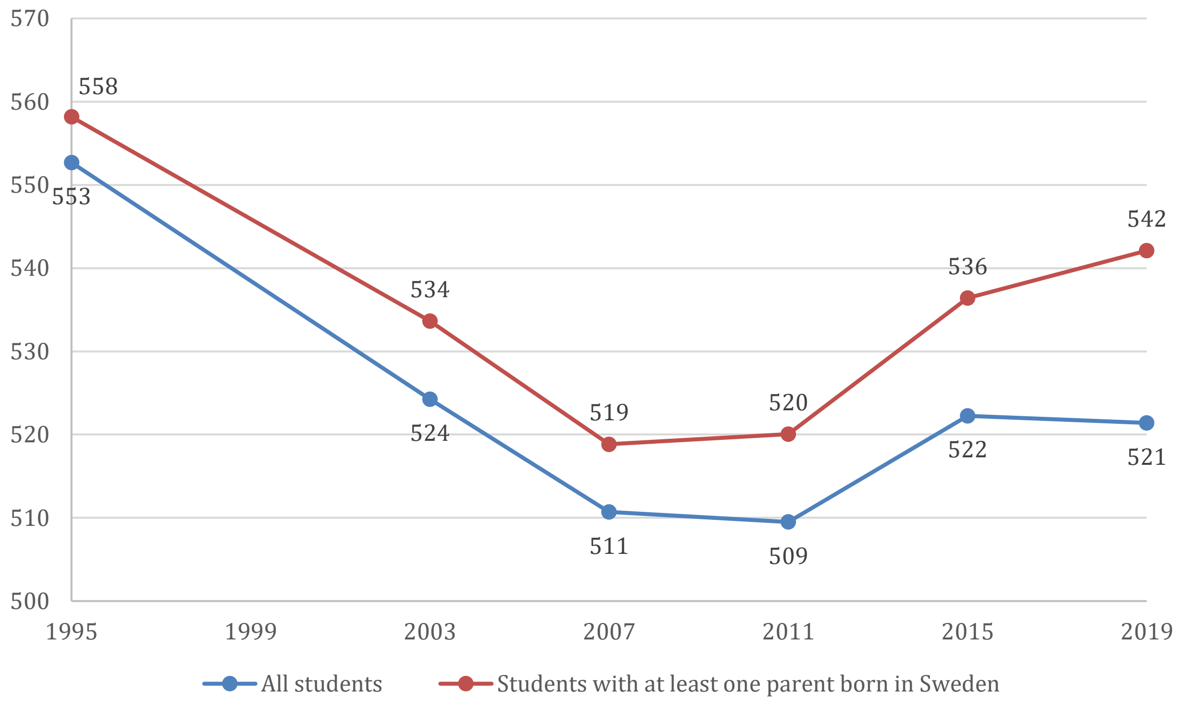


Figure 2. Sweden's science scores in TIMSS



between 2000 and 2012, but scores improved between 2012 and 2018. Again, this improvement was primarily driven by students with at least one parent born in Sweden, among whom the entire fall between 2000 and 2012 was erased in PISA 2018 (Heller-Sahlgren 2022).

While the performance decline coincides with increasing independent-school competition, the same is true for the performance improvement after 2011. Yet the extent to which Sweden's performance changes are driven by independent-school competition has not been analysed sufficiently. Prior research finds positive effects of such competition on domestic achievement measures in lower-secondary school (Ahlin 2003; Björklund et al. 2005; Böhlmark and Lindahl 2015; Sandström and Bergström 2005). In the most recent study, Böhlmark and Lindahl (2015) finds that increasing independent-school competition at the municipal level raises grades and national proficiency test scores in mathematics and English among students in year 9 by about 0.04–0.05 standard deviations between 1992 and 2009.

Still, questions are often raised about these findings, due to Sweden's decentralised grading and marking practices. Both marking of national proficiency tests and grading are today performed by teachers in a non-moderated criterion-referenced system, which decreases the reliability of performance metrics and gives rise to perverse incentives to compete by lenient marking and grading rather than by raising quality (e.g. Wennström 2020). Thus, we may not expect effects of competition to be as large when analysing international tests, which are externally marked and do not carry any incentives of manipulation due to their low-stakes nature.

Yet whether this assumption holds true depends on the marginal impact of independent-school competition, relative to other factors, on marking and grading leniency, as well as the extent to which domestic performance measures reflect knowledge and skills assessed in international tests. In fact, although difficult to study, research suggests that independent-school competition has had, at most, a marginal impact on (considerable) differences in grading standards in lower-secondary schools, with some estimates even suggesting a restraining effect (Böhlmark and Lindahl 2015; Holmlund et al. 2014; Vlachos 2010).[3] Instead, the key explanation for differences in standards appears to be the non-moderated marking and grading system itself, which leads teachers to assess student performance relatively within each school, thus generating more restrictive/lenient standards in schools with higher/lower-performing

[3] Böhlmark and Lindahl (2015) find negative point estimates when analysing the impact of competition on the performance difference between subjects in which students sit national proficiency tests and those where they do not, with the idea being that it is easier to manipulate grades without any standardised assessments. The estimate is statistically significant at the 10% level when excluding background controls but becomes more imprecise when including such controls.

students (NAE 2019, 2020). If independent-school competition improves students' knowledge and skills, we should thus also expect an endogenous response in the form of more restrictive marking and grading practices, which may in turn counteract stronger inflationary incentives in competitive regions.[4] Since studies that analyse the effects of competition on differential standards over time all compare different forms of teacher-assessed outcomes, this impact (as that of competition itself) may well be underestimated in existing research. Overall, it is thus not clear that the impact of independent-school competition should be attenuated when, by studying international test scores, all factors behind differential standards are neutralised. In fact, the opposite could be the case.

More generally, Swedish domestic performance measures are different from international tests, which may also affect the impact of competition.[5] For example, grades and national proficiency test scores also aim to capture other types of (less traditional) knowledge and skills apart from those tested in the TIMSS surveys (see Frändberg and Hagman 2017; Sollerman and Pettersson 2016). Additionally, the low-stakes nature of international assessments means that they better capture certain types of non-cognitive skills, such as internal motivation and persistence, which make students exert effort on a test that does not matter for their futures (see Gneezy et al. 2019; Zamarro et al. 2019).[6] If independent-school competition affects knowledge and skills that are better captured by low-stakes international tests, we may thus expect its effects to be different than when studying grades and national proficiency test scores.

Interestingly, in a robustness test, Böhlmark and Lindahl (2015) find point estimates that are about three-to-five times as large when studying data from TIMSS 1995 and 2007, compared to when studying grades and national proficiency test scores in equivalent models. This suggests that their findings in the latter respect are not biased due to lenient marking and grading – and that effects on low-stakes international test

[4] This may help explain why cross-sectional analyses have found that municipal schools in municipalities with more independent-school competition appear to be more restrictive in their marking of national proficiency tests than municipal schools in municipalities with less competition. In fact, the point estimates are negative when studying differential standards in both municipal and independent schools, although very imprecise (Tyrefors Hinnerich and Vlachos 2013). Yet this relationship could also be explained by other factors, such as time-invariant unobserved heterogeneity across municipalities.

[5] Indeed, while there is a relationship between students' TIMSS performance and grades/national proficiency test scores, it is far from perfect: grades and national proficiency test scores explain 39–59 percent of the variation in TIMSS 2015 scores (NAE 2017).

[6] Of course, high-stakes performance metrics often capture non-cognitive skills as well (e.g. Borghans et al. 2016), but they clearly do not measure skills that make students exert effort despite the fact that their performance has no consequences whatsoever.

scores are larger.[7] However, in this analysis, they exploit between-regional variation in independent-school competition in each survey, rather than the within-regional variation over time only. This differs compared with their analyses of grades and national proficiency test scores, and increases the likelihood of omitted-variable bias. It also means that the analysis technically does not just study the effects of competition on changes in TIMSS performance over time, but also the association between levels of competition and scores in each survey round.

Nevertheless, based on prior research, and Sweden's current performance trend, it is clearly reasonable to predict positive effects of independent-school competition in TIMSS. It is also reasonable to expect the effects of competition to be at least as large as in research analysing domestic performance metrics. This paper aims to provide a comprehensive evaluation of how competition has affected Sweden's TIMSS scores between 1995 and 2019, both during the period of declining scores and the period of improving performance.

3. Data and methodology

To study the effects of independent-school competition in TIMSS among students in year 8, I obtain Swedish student-level data from the International Association for the Evaluation of Educational Achievement (IEA 2022) and link these to regional-level variables obtained from the National Agency for Education (NAE 2021) and Statistics Sweden (2021). The TIMSS surveys have been conducted every four years since 1995, and Sweden has participated with students in year 8 in six of these surveys: 1995, 2003, 2007, 2011, 2015, and 2019. In each round, a nationally representative sample of the student population in year 8 is selected through a stratified two-stage cluster sample design with schools as the first-stage sampling unit, and one or two intact classes within these schools as the secondary sampling unit (see Martin et al. 2019). In total, 25,079 Swedish students have participated in the surveys between 1995 and 2019. In addition to sitting the tests, students also complete questionnaires that provide details about their background characteristics, which I use to obtain student- and county-level

[7] They find that a 10 percentage-point increase in the independent-school enrolment share raises grades and test scores by 0.03–0.04 standard deviations, when weighting the municipality-level regressions by the number of students in each municipality. The average impact is 0.10–0.17 standard deviations in analyses of TIMSS scores from 1995 and 2007 in equivalent student-level regressions, depending on whether (and which) controls are included.

control variables. Descriptive statistics of all variables discussed in this section are outlined in Table A1.

3.1. Student outcomes

The main outcome variable analysed is the average TIMSS score in mathematics and science, although I also display results for both subjects separately. In addition, I study the probability that students reach different international benchmarks: low (at least 400 points), intermediate (at least 475 points), high (at least 550 points), and advanced (at least 625 points). In the main analyses, I study the probability that students reach the different benchmarks in both mathematics and science, but I also display results when studying the subjects separately.

3.2. Independent-school competition

As principal measure of independent-school competition, I use the unweighted average of the municipal share of students in years 1–9 who attend independent schools in each county, which I link to the schools that have participated in TIMSS. This measure is calculated based on students' municipality of residence, which is responsible for funding and providing education.[8] Each municipality, of which there are 290 today, belongs to a county and there are 21 counties in total. In other words, the competition measure is constructed as the average of the municipal independent-school enrolment shares at the county level. The idea is to capture the average level of independent-school competition that affects students in each county.[9]

Another option is to use the independent-school enrolment shares at the county level overall to study the effects of competition. Unlike the main measure, this variable is effectively weighted by the student population in each municipality. However, this also means that it tends to be dominated by larger municipalities, which likely makes it worse for capturing the average level of competition to which students in the different

[8] An alternative is to calculate the corresponding independent-school competition measure based on the municipality in which schools are located (NAE 2021). This variable can only be calculated from 1998 onwards. However, as reported in Table A7 in the Appendix, results are almost identical when calculating the county-level competition measure using data at the school municipality level and imputing 1995 values with the values from 1998.

[9] There is no geographical information to link two schools participating in TIMSS 2011, with 24 students in total, to the correct municipality. These are thus excluded from the analysis. In unreported analyses, I allocated these two schools and their students to all different counties, one by one, and re-ran the main regressions to ensure that this did not affect the results. The results were almost identical, which is expected given that the issue only applies to two schools with 24 students in one of the test rounds.

municipalities in each county are subjected. Regardless, in robustness tests, I also use the weighted measure to ensure that the results are not dependent on the preferred measure of competition.

Studying the impact of independent-school competition at the county level differs somewhat from prior research, which principally analyses effects at the municipality level, although it also conducts robustness tests at the county level (Böhlmark and Lindahl 2015).[10] An important reason why I choose to study the issue at the county level is that this paper, unlike prior research, analyses a sample of 4,000–5,000 students per test round (except for TIMSS 1995 when about 2,000 students participated).[11] Only between 87 and 100 municipalities are represented in each test round, and only 156 municipalities have had students participating in at least two test rounds, while the number of participating students in each municipality in any given round is also small.[12] This makes it less meaningful to study how changes in independent-school competition and changes in test scores over time are related at the municipal level. This is especially true given that I want to analyse how the effect has varied over time.

On the other hand, students from all counties have participated in all test rounds, with the exception for Gotland, which had no participating students in TIMSS 2015. The number of students at the county level is also much larger than at the municipal level.[13] Despite the fact that TIMSS is designed to be representative at the national rather than county level, this means that it is more likely that results are externally valid when studying the effects at the county level compared with the municipal level. But even if external validity is low, it does not affect the internal validity of the findings – the effects of independent-school competition among students who participated in the TIMSS

[10] Böhlmark and Lindahl (2015) use the independent-school enrolment shares in each county overall in this analysis. They find slightly larger point estimates than when they analyse effects at the municipal level, but these estimates are also less precise when they include controls. This may very well be because the measure, as argued above, does not capture average competition levels in each county as well as the unweighted municipal average at the county level.

[11] The exception is Böhlmark and Lindahl's (2015) robustness test when analysing data from TIMSS 1995 and 2007. But precisely because the statistical imprecision is high due to the sample size, they exploit the between-regional variation in independent-school competition in each survey, in addition to the within-regional variation over time.

[12] The number of participating students in each municipality varies between 4 and 560, with a median of between 18 and 44 depending on the test round. The sample size in TIMSS 1995 was about half as large as in the other test rounds, and if I exclude this round from the calculations the median was between 25 and 44 students.

[13] The number of participating students by county and test round varies between 14 and 1,135. The median varies between 69 and 158 depending on the test round. The sample size in TIMSS 1995 was about half as large as in the other test rounds, and if I exclude this test round from the calculations the median was between 108 and 158 students.

surveys – which is what matters for understanding how increasing independent-school competition has affected Sweden's performance in TIMSS over time.

Furthermore, there are other advantages of analysing the effects of competition at the county level. For example, doing so means that one can take into account spill-over effects between municipalities. This may be important since independent schools are not constrained by any geographical boundaries when recruiting students, and they are thus likely to compete with schools in different municipalities. Also, research finds positive effects of larger independent-school enrolment shares on teacher wages (at the upper-secondary school level), which may have altered the teacher composition across municipality borders (see Böhlmark and Lindahl 2015; Hensvik 2012). While endogenous teacher mobility does not affect the validity of estimates, it does affect the interpretation of any effects as they may be due to teacher sorting rather than productivity gains. Such issues are better accounted for when analysing the effects of independent-school competition at the county level.

It is also important to analyse the effects of competition from for-profit and non-profit operators separately. In this respect, I only have access to data for students in years 6–9 from 2011 onwards (Statistics Sweden 2021). However, since the average independent-school enrolment shares in 1995 were small across the country – varying between 0 and approximately 3 percent depending on county – I predict the enrolment shares of for-profit and non-profit independent schools among students in years 6–9 in 1995 from the relationship between these variables and the overall independent-school enrolment shares in the years 2011, 2015, and 2019.[14] Classifying schools, I use the same definitions of for-profit and non-profit operators as Holmlund et al. (2019): limited companies and general partnerships are defined as for-profit operators, while foundations, non-profit associations, and economic associations are defined as non-profit operators.[15]

[14] To study whether this approach is reasonable, I also analysed the impact of the overall independent-school enrolment shares among students in years 6–9, using the same method to predict the enrolment shares in 1995. As displayed in Table A7 in the Appendix, the results are very similar compared to those in models using the main measure of independent-school competition. The results are also very similar if I instead assume that the enrolment shares of both for-profit and non-profit independent schools were 0 percent in 1995.

[15] In practice, there are few schools run by general partnership, having extremely small total enrolment shares in the relevant years. Also, a total of 77 students in five municipalities attended schools run by sole traders in 2015 only. To keep the definitions the same as in Holmlund et al. (2019), these students are classified as attending a non-profit school, but results are unsurprisingly identical if I instead classify them as attending a for-profit school or exclude them from the construction of the separate measures.

3.3. Control variables

The paper adjusts for several relevant control variables at the student level, which may be related both to TIMSS scores and independent-school competition. These variables are obtained from the student survey that is conducted in conjunction with each test round.[16] First, the number of books at home is included, since this variable has been shown to capture students' socioeconomic background well in Swedish TIMSS data (Wiberg and Rolfsman 2023).[17] Furthermore, I include indicators for whether students have a computer and a desk at home respectively, which are the only indicators for home resources that are available in all TIMSS rounds. Also, I adjust for students' immigrant background – captured by separate indicators for whether students and their parents were born in Sweden, age at arrival in Sweden, and how often students speak Swedish at home – gender, and age. Since the student-level controls are constructed from survey data, I allow their effects to vary in each round to account for changes in how answers reflect student background over time.

Apart from the student-level controls, in the full model I also include the county average of the number of books at home among participating students, and the county share of participating students with some foreign background, the effects of which are also allowed to vary in each survey round for the reason noted above.[18] In addition, from administrative data, I separately add the (log) number of students at the county level to account for scale economies (Statistics Sweden 2021).[19] The latter indicator is available among students in year 9 in all test rounds, which is the best available proxy for the number of students in year 8.[20]

[16] Generally, non-response rates are low for the questions used to construct the control variables, but to ensure that I study the entire Swedish sample, any missing values are assigned to their own category and dummies for these categories are included. Similar techniques to deal with missing values on student-background characteristics are used widely in other similar research (e.g. Falck and Woessmann 2013).

[17] The questions regarding parental educational background suffer from considerable attrition in the TIMSS surveys (Wiberg and Rolfsman 2023) and are thus not included. However, results are very similar if I include an indicator for parents' highest educational level together with a missing dummy, in line with the discussion in the previous footnote.

[18] The county variables are constructed among students who have responded to the relevant questions. Students are defined as having some foreign background if they are not born in Sweden and have two Swedish-born parents. The results are almost identical if I use alternative definitions in this respect.

[19] This variable is calculated based on the municipality of residence. Overall, results are very similar if I also allow the effect of this indicator to vary over time, or add other variables from administrative data covering students in year 9, including the county average parental educational level, on a scale from 1 to 7, and the county share of students with an immigrant background.

[20] However, the results are almost identical if instead include the (log) number of 14-year-olds at the county level as a proxy for the number of students in year 8.

3.4. Empirical set-up

It is not straightforward to study the effects of independent-school competition in the Swedish context. This is because the 1992 voucher reform was implemented in the entire country at the same time and operators choose where to open new schools, meaning there is no proper, untreated control regions with which to compare. This is especially true when studying TIMSS data, since the first survey was conducted a few years after the voucher reform. To analyse the effects of competition, I thus follow prior research in exploiting the differential increase in independent-school enrolment shares across regions over time, in a panel set-up to adjust for time-invariant unobserved differences across these regions, while also adjusting for the time-varying controls discussed in Section 3.3.

Studying the effects of independent-school competition on TIMSS scores, implicitly assuming homogenous treatment effects across counties and over time at different levels of competition (Callaway et al. 2021), I thus estimate the following OLS model:

$$y_{sct} = \alpha + \beta_1 sp_{ct} + \beta_2 x_{sct} + \beta_3 z_{ct} + \delta_c + \mu_t + \varepsilon_{sct} \tag{1}$$

where y_{sct} is the average TIMSS score of student s in county c in year t; sp_{ct} denotes the average independent-school enrolment share in each county; x_{sct} is a vector of student-level controls while z_{ct} is a vector of county-level controls, discussed in Section 3.3; δ_c denotes county-fixed effects, which absorb time-invariant unobservable county characteristics; μ_t represents time-fixed effects, which absorb common shocks that affect all counties equally.[21] Standard errors are clustered at the county level because the variable capturing independent-school competition varies at this level.[22]

[21] Unless stated otherwise, regressions in this paper are adjusted for the fact that students' TIMSS scores are estimated from five plausible values in each subject (see Wu 2005). In practice, this means that the models are estimated separately for each plausible value. The point estimate is then equivalent to the average point estimate for the five plausible values, and the standard error is estimated using Rubin's rules for handling multiple imputations (see Jerrim et al. 2017). In models studying the average mathematics and science scores, I use the average of each of the five plausible values in both subjects.

[22] Since there are only 21 counties, there is a risk that the number of clusters is too small to estimate the models correctly. Furthermore, the number of observations varies quite a lot between the counties, which may also create problems for interpreting the results (see Cameron and Miller 2015; Roodman et al. 2019). Since I scale the student weights in the main models so that all students in the same county and test round sum to 1 – to ensure that all counties are given equal weight in the models – the latter is not a problem for my analysis. Yet to ensure that the number of clusters is not too small, I estimated the main models on the average of the five plausible values and used the wild cluster bootstrap procedure to adjust the relevant p-values for this possibility (Roodman et al. 2019). As displayed in Table A2, these p-values

The model's assumption is that $Cov(sp_{ct}, \varepsilon_{sct}|x_{sct}, z_{ct}, \delta_c, \mu_t) = 0$. Effectively, this means that changes in time-varying unobservable county characteristics that affect performance should not be correlated with changes in independent-school enrolment shares. This assumption is violated if changes in student performance generate changes in independent-school competition, reflecting reverse causality, if there is measurement error in the competition measure, and/or there exist omitted time-varying variables that affect both changes in competition and student performance. The overall direction of such potential bias is theoretically unclear. For example, private operators may establish more new schools in areas with declining school results, as demand may be stronger in such areas, which is likely to bias estimates downwards. Also, potential measurement error in the variable used to capture independent-school competition would also bias estimates downwards. On the other hand, private operators may prefer to start schools in areas with improving school results, for example because of influxes of socio-economically advantaged students who may effectively be cheaper and easier to educate (see Edmark 2019). If this is the case, and the phenomenon is not captured by the included controls, one would expect estimates to be biased upwards.

Naturally, given the nature of the data analysed, it is impossible to conduct checks on pre-reform trends in support of a causal interpretation of the results, although prior research analysing grades and national proficiency test scores suggests they do support such an interpretation (Böhlmark and Lindahl 2015).[23] To investigate the likelihood that the findings reflect the causal impact of independent-school competition, I thus instead devise a placebo test in treatment using the share of children attending independent pre-schools. The idea is to study the extent to which the results capture more general regional preferences for private providers in fields that are related to primary- and secondary schooling. If this is the case, the main analysis risks capturing the effects of other factors than independent-school competition on TIMSS scores. Certainly, a correlation between independent pre-school enrolment shares and independent-school enrolment shares is expected, but changes in the variation in the former that cannot be explained by the latter should not be related to changes in TIMSS scores.

are almost identical to the ones produced in the equivalent models using clustering, suggesting that the findings are not biased due to a small number of clusters.

[23] Note that these checks are conducted on grades given in the pre-reform cohort-referenced system, which means that differential grade inflation is likely less of a problem for interpreting the results.

In the first step, I thus predict the municipal independent pre-school enrolment share, for each survey year, using the corresponding share of students attending independent primary- and lower-secondary schools. In the second step, I extract the variation in the independent pre-school enrolment share that cannot be explained by the independent-school enrolment share and calculate the unweighted county average.[24] In the regressions, I then study whether the residual – the difference between the actual independent pre-school enrolment shares and the variation in this variable that can be predicted by independent-school enrolment shares – is related to TIMSS scores, once adjusting for county- and time-fixed effects and all control variables.

Since I study a sample of students, I utilise the weights that adjust for students' sampling probability. These ensure that the sample is representative at the national level. But since I analyse the effects of independent-school competition at the county level – and each county is thought to represent a separate market – I adjust the weights so that all students in the same county and test round sum to 1. This means that I give all counties equal weight in the models, which is equivalent to Böhlmark and Lindahl's (2015) main approach. The models thus analyse how independent-school competition affects an average student in each county.

However, I also display results when using the unadjusted student weights, thus giving larger counties more weight in the regressions. The models then analyse how independent-school competition affects an average student in the country rather than in each county. These estimates are thus also used to simulate Sweden's counterfactual performance trend without increasing independent-school enrolment shares. Finally, I also estimate models without any weights at all to ensure that the findings are not driven by a small number of students who happen to account for a considerable share of the variation in competition across counties, while being given a large weight in the regressions merely because they also happen to be underrepresented in the sample.

In the main models, I include data from TIMSS 1995, 2003, 2011, and 2019. This ensures that I can analyse the effects of independent-school competition over time in a balanced panel with eight years between each survey. This also allows me to estimate the effects of competition over longer time periods – exploiting the fact that there has been consistent growth in independent-school enrolment shares (and the variation in

[24] The share of children attending independent pre-schools is only available from 1998, and I thus use data for 1998 instead of data for 1995 in the first year of the panel.

these shares across counties) over time – which is important when estimating general-equilibrium effects. In addition, analysing data over longer time periods decreases the estimates' sensitivity to measurement error (see Böhlmark and Lindahl 2015). However, in other models, I include data from all survey rounds, or restrict the sample in other ways to analyse the extent to which the effect of independent-school competition varies over different periods.

4. Results

This section reports the results, which are displayed in figures in the main text, showing the effect of a 10 percentage-point increase in the average independent-school enrolment share on students' TIMSS scores. The error bars in the figures display a 90% confidence interval. Details regarding the results, including levels of statistical significance, are provided in the tables in the Appendix.

4.1. Main results

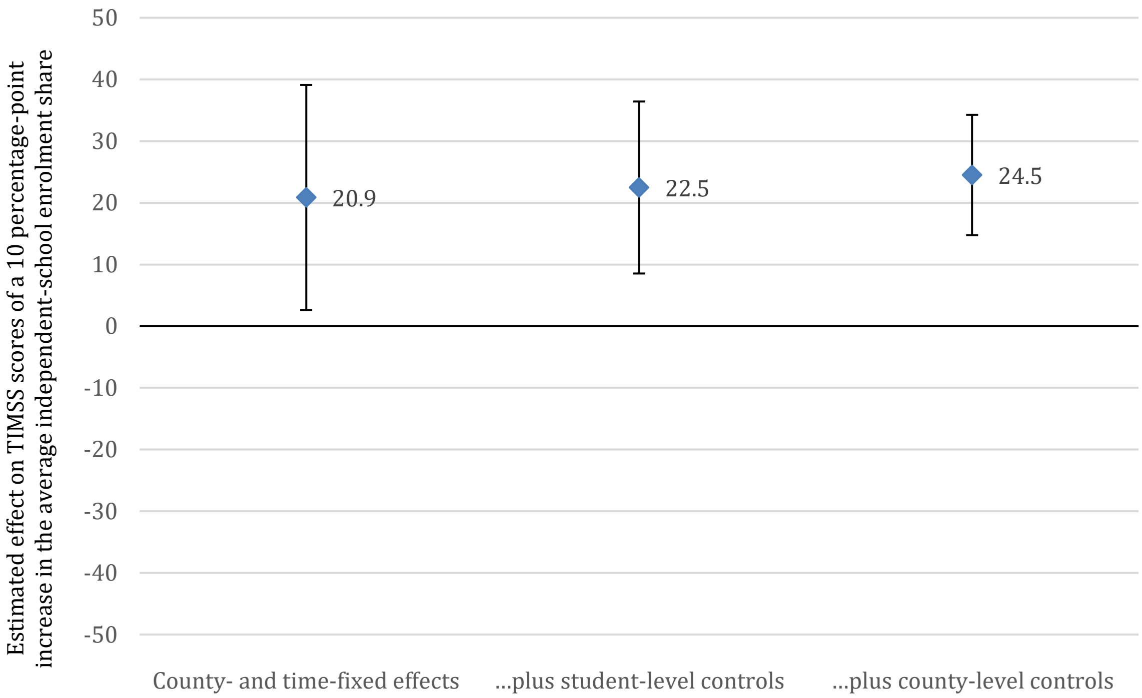
Figure 1 shows that a 10 percentage-point increase in the average independent-school enrolment share raises TIMSS scores by 21–25 points, depending on which controls are included. The results are very similar in all three models, but they become more precise when including controls at the student- and county levels. This is expected if the findings reflect a causal relationship, since it suggests the control variables are related to student performance in TIMSS, but essentially unrelated to independent-school competition, once adjusting for county- and time-fixed effects.[25] As displayed in Table A7 in the Appendix, the results are also very similar when including data from all test rounds in the analysis, and when studying the effects on mathematics and science performance separately.

Interestingly, at the same time, Table A2 also shows that there is no significant relationship between independent-school competition and student performance in TIMSS when excluding county-fixed effects, regardless of whether I adjust for student- and county-level control variables. This suggests that estimates from models that do not

[25] When studying all controls in separate models, I found that almost all of them are indeed separately related to TIMSS scores when adjusting for county-fixed effects. In contrast, no controls were related to independent-school competition when adjusting for county-fixed effects. However, when excluding county-fixed effects, some of the controls were related to independent-school competition, thus highlighting the importance of adjusting for unobserved time-invariant heterogeneity.

adjust for time-invariant unobservable characteristics that vary across counties are likely to be biased downwards.

Figure 1. Independent-school competition and TIMSS scores
(TIMSS 1995, 2003, 2011 och 2019)



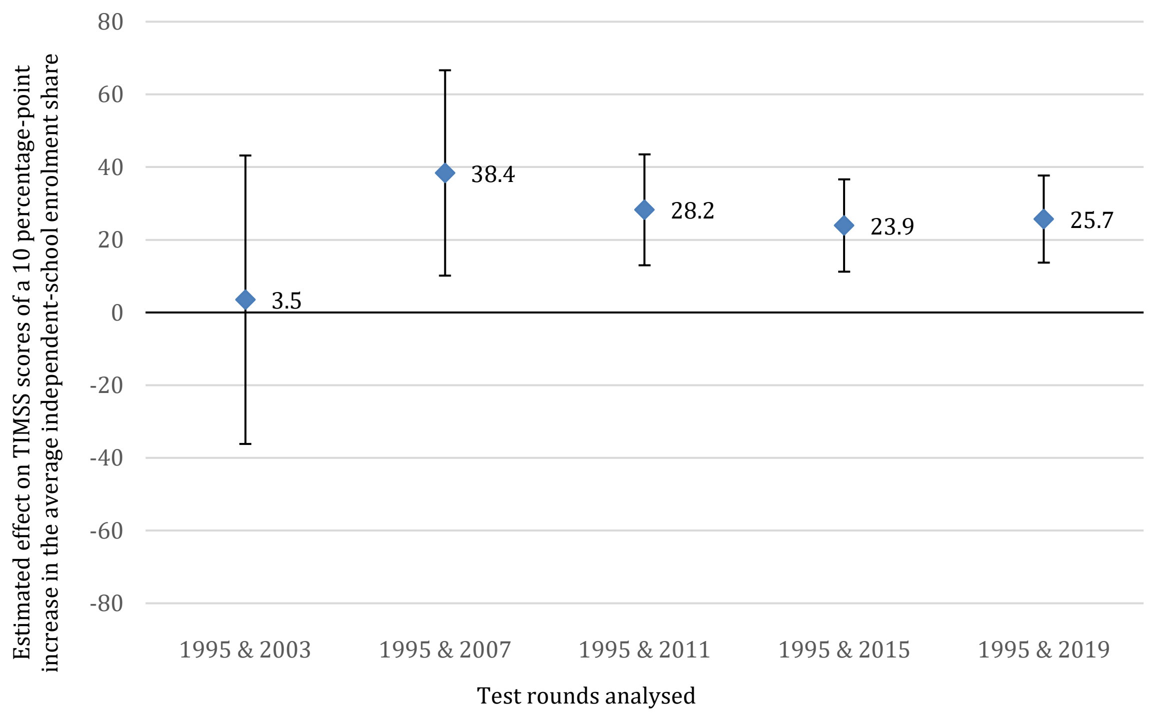
Overall, the results thus suggest that increasing independent-school competition has improved TIMSS scores over time, compared to a counterfactual situation in which competition would have remained at the levels in 1995. Counties with larger independent-school enrolment share growth have had a more positive performance trend in TIMSS, when adjusting for county- and time-fixed effects as well as relevant control variables at the student- and county levels.

4.2. When does the impact appear?

Previous research indicates that the impact of independent-school competition on student grades and national proficiency scores appeared from around 2005 onwards (Böhlmark and Lindahl 2015). In order to study whether this also applies to the impact of independent-school competition on TIMSS scores, I analyse the effect of changes in competition levels between 1995 and each following survey round in separate models.

Figure 2 shows that the positive impact of independent-school competition is only detectable after 2003.[26] The point estimate in the model that only includes data from 1995 and 2003 is close to zero and very imprecise. The lack of precision likely arises because the variation in the change in independent-school competition was quite small in this period. Consequently, I cannot rule out that the impact between 1995 and 2003 is of the same magnitude as the effect displayed in Figure 1. Regardless, only four years later, the effect is positive and statistically significant – and it becomes more precise in the following test rounds. This is to be expected since the variation in the change in independent-school competition increases over time.

Figure 2. The variation in the effect of independent-school competition over time



Overall, these findings are thus in line with previous research: the positive impact of independent-school competition becomes detectable only from around 2005. This is reasonable given that the increase in independent-school enrolment shares, and the variation between counties in this respect, were rather modest in the first ten years after the voucher reform. Yet as competition increased, so apparently did its impact on student performance.

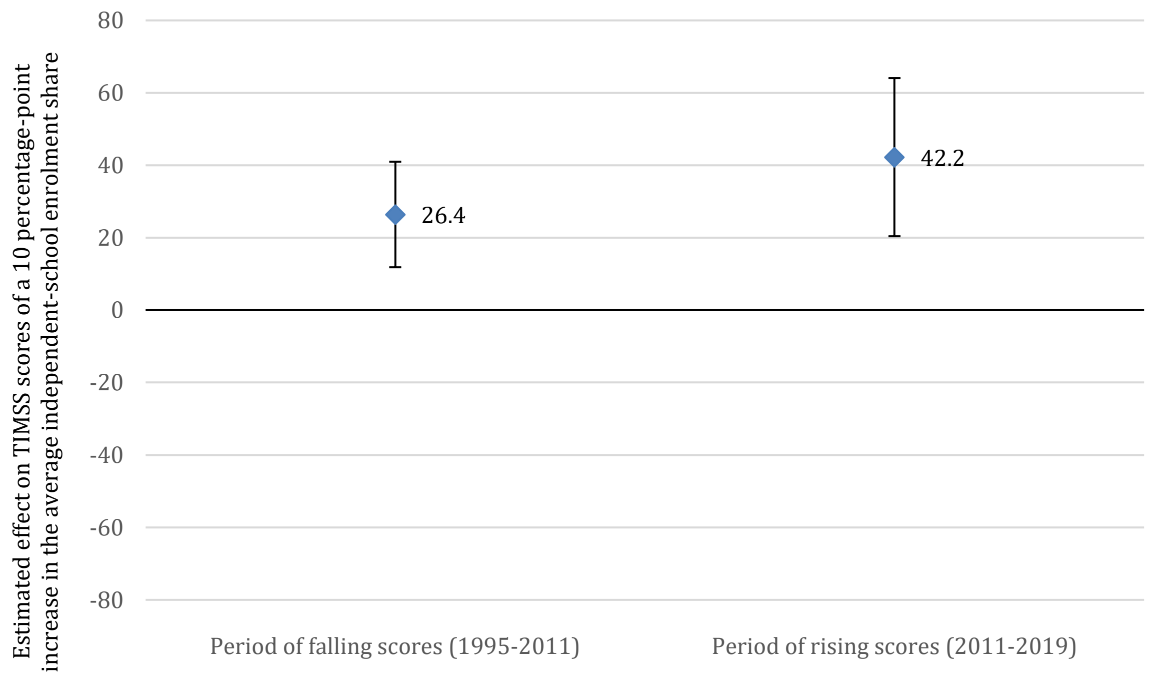
[26] All models in Figures 2–5 include controls at the student and county levels.

4.3. Effects during the periods of falling and rising performance

The above analyses suggests that independent-school competition slowed down Sweden's fall in TIMSS between 1995 and 2011, at least from 2007 onwards. But has it also contributed to the improvements that have been seen since 2011? And, if so, what is the relative effect size in the respective periods? To answer these questions, I analyse the impact of independent school-competition during the performance decline and rise, including all test rounds in each period.

Figure 3 shows that the positive effects are apparent in both the period of falling scores and the period of rising performance. The point estimate is larger, but it is also less precise, between 2011 and 2019, which is to be expected since the increase and the variation in changes in independent-school competition were larger in the period 1995–2011. Consequently, the difference is not statistically significant. Regardless, the positive effect is apparent when analysing the periods of declining and improving scores separately. In other words, independent-school competition appears to have both slowed down Sweden's performance decline between 1995 and 2011 and contributed to its improving scores since 2011.

Figure 3. The effect of independent-school competition during the periods of falling and rising performance



In Table A4, I also show separate results between 2011 and each following test round during the period of rising performance, equivalent to the estimates during the period of falling performance in Section 4.2. The effect size is the largest between 2011 and 2015, which is when Sweden's scores improved on average, but the precision in the estimate is lower than between 2011 and 2019. This also means that I cannot rule out that the effects are of the same magnitude in both periods.

4.4. Competition from for-profit and non-profit independent schools

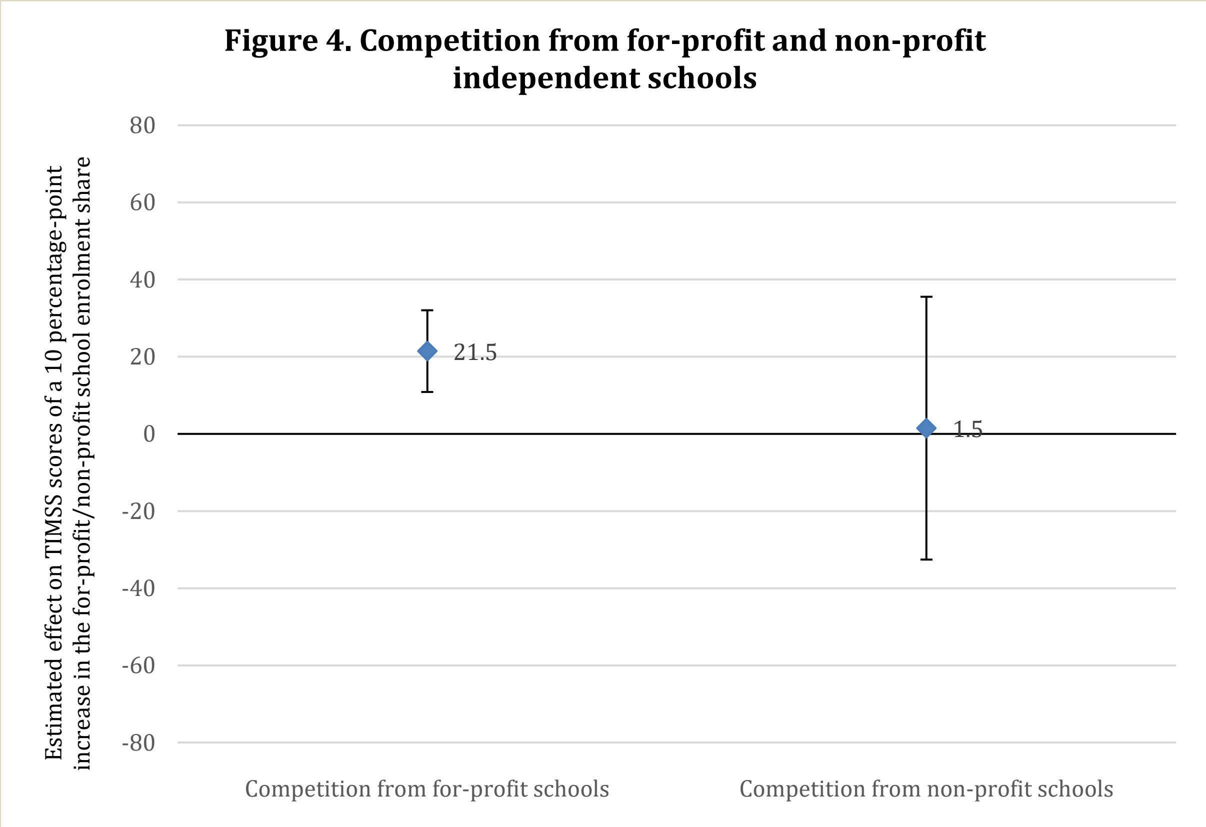
The largest share of the increase in independent-school enrolment shares can be explained by the proliferation of for-profit schools; the growth in the share of students attending non-profit independent schools has been more modest (Holmlund et al. 2019). It is thus plausible that the effects of independent-school competition are primarily driven by the growth of for-profit schools specifically. To study this issue, I separately analyse how competition from for-profit and non-profit independent schools has affected student performance in TIMSS.[27]

Figure 4 shows that the positive impact of independent-school competition can be entirely explained by increasing shares of students attending for-profit schools. While the difference in effect compared with competition from non-profit schools is not statistically significant, the latter is unrelated to student performance when analysing both measures in the same model. Table A5 shows that this is also the case when studying the effects of competition from for-profit and non-profit schools in separate models, although the non-profit independent-school enrolment share coefficient increases in size (and in one case turns negative). This is because changes in for-profit and non-profit independent-school enrolment shares are positively correlated overall. But competition from non-profit independent schools does not appear to have a separate impact once taking this correlation into account.

Certainly, since the increase in non-profit independent school enrolment shares has been much smaller than the increase in for-profit independent school enrolment shares, it is difficult to draw strong conclusions regarding whether these results reflect differences in the effects of for-profit and non-profit schools more generally. Yet it does

[27] As described in Section 3.2, due to data availability, I can only include the years 1995, 2011, 2015, and 2019 in these analyses. As Table A5 shows, results are very similar if I analyse data over the entire period between 1995 and 2019.

appear as if the impact of independent-school competition in TIMSS is driven primarily by the growth of for-profit schools specifically.



4.5. Effects of competition on the probability of reaching different benchmarks

Another important issue is the extent to which independent-school competition has affected the probability that students reach different international performance benchmarks. I thus analyse the probability that students reach the low (basic) level (at least 400 points), intermediate level (at least 475 points), high level (at least 550 points), and advanced level (at least 625 points).

Figure 5 shows that independent-school competition has a positive impact on the probability that students reach the intermediate, high, and advanced benchmarks. An increase in the average share of students attending independent schools by 10 percentage points increases the probability that students reach the intermediate and high levels in both subjects by 12 and 11 percentage points respectively, and the probability that students reach the advanced level by 4 percentage points. These are important effects given that the share reaching the intermediate level in both subjects was 60 percent, the share reaching the high level was 24 percent, and the share reaching the advanced level amounted to 4 percentage points in 2019. The point estimate when analysing the probability that students reach the low level in both subjects is positive, but it does not reach statistical significance. However, the share reaching the low benchmark is much larger than the other categories (86 percent in 2019), and the variation that can be explained is smaller.

Figure 5. The effect of independent-school competition on the probability of reaching different benchmarks

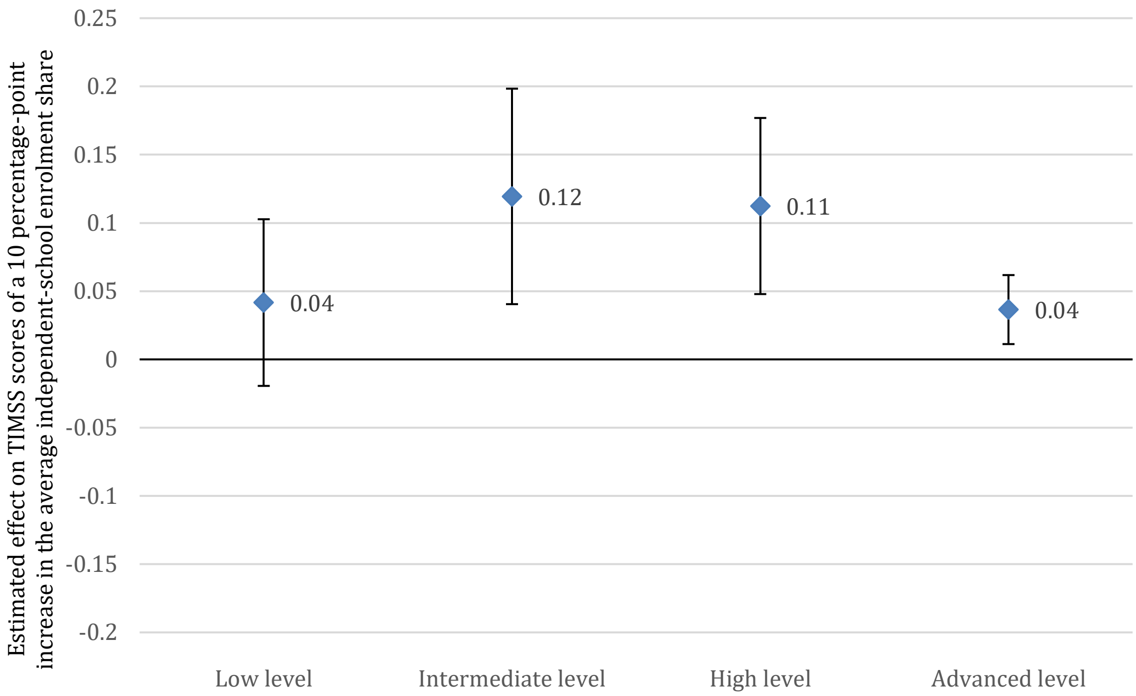


Table A7 shows that the results are generally similar when analysing the subjects separately. However, the effects on the probability that students reach the high and advanced benchmarks are considerably larger in science (although still significant in mathematics), while the impact on the probability that students reach the intermediate benchmark is considerably larger in mathematics (and not significant in science). There is no significant impact on the probability that students reach the low threshold in either subject. While no differences between subjects are statistically significant, this indicates that independent-school competition affects lower- and higher-performing students somewhat differently in mathematics and science.[28]

Overall, the effect of independent-school competition is thus also detectable when analysing the probability that students reach different performance benchmarks. Competition appears to benefit both lower- and higher-performing students, although the impact at the lowest benchmark is not significant.

4.6. Further analyses

In addition to the results reported above, I also conducted several further analyses, the results of which are displayed in Table A7. The estimate is very similar when including all test rounds in the analysis, and when using the unadjusted student weights, thus giving larger counties larger weights in the regressions. This also holds true when not using any weights at all. The results are also very similar when excluding Stockholm, Skåne, and Västra Götaland counties to ensure that the results are not driven entirely by large metropolitan areas, even though about 50 percent of the sample are dropped.[29] To ensure that the results are not driven by outliers, I also carried out the analysis using median regression. The impact is again very similar.[30]

Also, estimates are essentially identical when calculating the competition measure based on the municipality in which schools are located rather than students' municipality of residence, and when using the independent-school enrolment share in each county overall, rather than the unweighted municipal average, to study the effects of competition.[31] This is also the case when adjusting for the average share of students attending a municipal school outside students' home municipality, a measure intending to capture competition between municipal education providers.[32]

---

[28] This may reflect the fact that the share of students who do not reach the intermediate benchmark is larger in mathematics (36 percent in 2019) than in science (29 percent in 2019), while the shares of students who reach the high and advanced thresholds are larger in science (41 and 13 percent respectively in 2019) than in mathematics (28 and 5 percent respectively in 2019).

[29] This also holds true when using the unadjusted student weights or no weights at all. In unreported analyses, I also excluded each of the 21 counties, one by one, and results were always very similar.

[30] In this model, I study the average of the five plausible values (and cannot include weights). As the model did not converge when including interactions between the student-level controls and year-fixed effects, I exclude these interactions and display the equivalent OLS estimates for comparison purposes.

[31] However, when using the weighted average, while the coefficient is essentially identical as in the main models, it is less precise. This suggests that it is a noisier measure than the unweighted average, perhaps because the former does not capture average competition levels in each county as well as the latter.

[32] This indicator has no relationship with student performance in TIMSS. The point estimate is marginally negative, but very close to zero and far from statistically significant.

Furthermore, Table A7 shows that the effect is almost identical if I include a dummy indicating whether participating students attend an independent school, which allows me to adjust for the direct association between independent-school attendance and performance in TIMSS.[33] Also, the results are almost identical if I instead exclude all participating students who attend independent schools. In other words, independent-school competition appears to impact performance in municipal and independent schools approximately equally.[34]

Finally, Table A7 also shows that there is no relationship between student performance and the share of the variation in independent pre-school enrolment shares that cannot be predicted by independent-school enrolment shares. As discussed in Section 3.4, this placebo test supports the argument that the analysis captures the causal effects of increasing independent-school competition on TIMSS scores in Sweden.[35]

4.7. Discussion

Overall, the paper displays a positive effect of independent-school competition on Swedish TIMSS scores over time. Importantly, my findings suggest that estimates in prior research are not biased upwards due to unreliable domestic performance metrics. On the contrary, the effects found in this paper are much larger than the impact found in prior research analysing domestic outcomes. Indeed, Böhlmark and Lindahl (2015) find that a 10 percentage-point increase in independent-school enrolment shares raises grades and national proficiency test scores by about 0.04–0.05 standard deviations in the period 1992–2009, when weighting municipalities equally in the regressions. Analysing the effect at the county level, they find an effect size of about 0.05–0.06

[33] The independent-school coefficient in this regression is positive (15.84) and significant (standard error = 6.65). There is no information regarding school ownership for eight schools, including the two schools for which there is no geographical information, which are thus excluded.

[34] While adjusting for the independent-school dummy holds constant an endogenous variable, and excluding students in independent schools introduces sample-selection issues, both modifications are likely to bias the estimates of competition downwards. This is because independent schools historically have had more advantaged students on average, although differences have decreased over time (e.g. Holmlund et al. 2019), suggesting positive selection into these schools.

[35] In addition to analyses reported in Section 4.6, I also checked whether the results changed when adding county-specific linear trends to the full main model, thus exploiting deviations from linear trends, rather than changes in levels, as the source of variation. Doing so is clearly problematic in my setting, where it is only possible to adjust for post-reform trends, as such trends often control for the treatment effect itself (e.g. Goodman-Bacon 2021; Meer and West 2016). Nevertheless, the results are in fact very similar when including county-specific linear trends in the full main model, with a slightly larger coefficient, although precision also decreases. This is unsurprising given the reduction in the variation exploited.

standard deviations. This is much smaller compared with the equivalent finding of about 0.24–0.30 standard deviations (about 20–25 TIMSS points) at the county level in this paper.[36]

However, Böhlmark and Lindahl (2015) also find an effect of 0.10–0.17 standard deviations (about 7–12 TIMSS points) when studying low-stakes TIMSS scores in 1995 and 2007, depending on whether (and which) controls are included, which is about three-to-five times as large compared to the equivalent population-weighted estimates when studying domestic performance metrics (0.03–0.04 standard deviations). As noted in Section 2, this may well be reasonable since all factors behind differential standards in domestic marking and grading are neutralised when studying international test scores, rather than just those stemming from competitive incentives, and since competition may affect knowledge and skills that are better captured by international low-stakes tests differently.

Yet, as also noted in Section 2, when studying data from TIMSS 1995 and 2007, Böhlmark and Lindahl (2015) also utilise the between-regional variation in independent-school competition in each survey, rather than the within-regional variation over time only. And this is an important difference compared with this paper. Indeed, when modifying my analysis to be more similar to their set-up, thus exploiting the between-municipal variation in independent-school competition within counties and including the unadjusted student weights, the effect size in a model with student-level controls is 5 TIMSS points and thus very similar to their main model.[37] Meanwhile, when excluding county-fixed effects in an equivalent analysis at the county level, thus also exploiting between-county variation, I obtain an almost identical estimate of about 6 TIMSS points in a model with student-level controls.[38]

---

[36] The standard deviation of the average mathematics and science score in TIMSS 2019 was 82.38 points.

[37] The coefficient is 54.13 (standard error clustered at the municipal level = 19.20). If I also include municipal-level control variables equivalent to the county-level controls used in this paper, the estimate is close to zero and insignificant. This appears broadly similar to Böhlmark and Lindahl's (2015) equivalent findings, as discussed in the note to their Table 9.

[38] The coefficient is 58.22 (standard error clustered at the county level = 13.99). This is the same model as the one including student-level controls in the second panel in Table A2 – which produces a very similar yet insignificant estimate – but includes unadjusted student weights to replicate Böhlmark and Lindahl's approach when analysing TIMSS data. However, when including county-level controls the estimate is close to zero and insignificant also with the unadjusted weights, which, as noted in the previous footnote, also is similar to Böhlmark and Lindahl's (2015) findings. Combined with the results in this paper, it also suggests that models exploiting the variation in independent-school competition between regions are more sensitive to the inclusion of regional-level controls than models that only exploit the variation within regions over time.

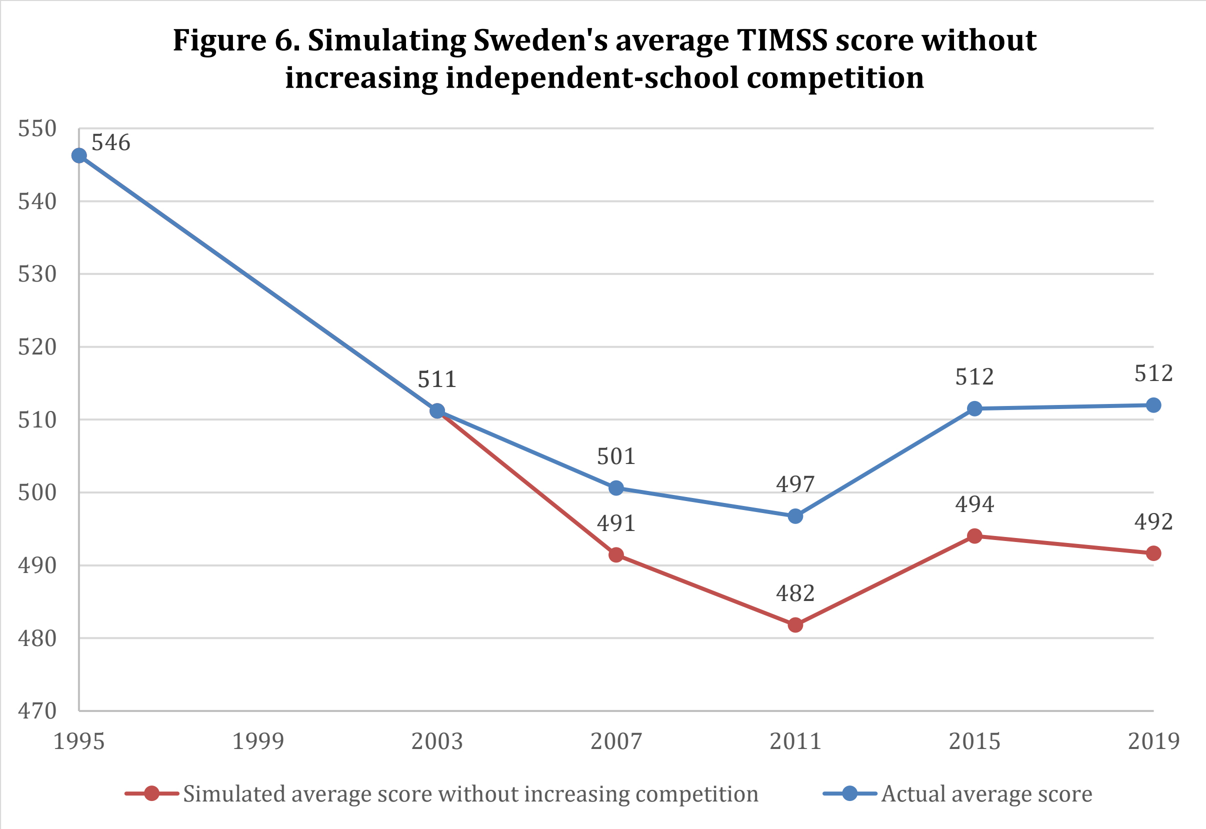
Overall, this suggests that the larger effect size in this paper compared to prior research can likely be explained by the combination of the fact that I (1) study low-stakes international test scores, which both neutralise all factors behind differential marking and grading standards as well as partly reflect different types of knowledge and skills compared with domestic outcomes, (2) analyse the effects of competition at the county rather than municipal level, and (3) only utilise the variation in independent-school competition that varies within counties over time.

5. Simulating Sweden's counterfactual performance trend without competition

In this section, using the empirical estimates from the paper, I simulate Sweden's counterfactual performance trend in TIMSS without any expansion of the independent-school sector. To do so, I extrapolate the impact of independent-school competition from a model that includes all test rounds between 1995 and 2019 and use the unadjusted student weights. This is to ensure that the estimate picks up the effect of competition on an average student in the country across all years, which is the interesting parameter when simulating Sweden's counterfactual trend overall. All control variables at the student and county levels are included. As displayed in Table A7, the estimate from this model is very similar to the main findings.

Since this paper, like prior research, does not find any positive effects between 1995 and 2003, I assume that Sweden's average trend in this period would have been the same without any increase in competition. In the simulation, I first multiply the coefficient for the average share of students attending independent schools with the change in this share at the national level between 1995 and each survey year from 2007 onwards. I then subtract the product from Sweden's actual average TIMSS score for each year.

The results are displayed in Figure 6. According to the simulation, Sweden would have scored 15 points lower in 2011, when the country's performance was at its lowest. The improvement that occurred between 2011 and 2015 is in turn predicted to have been 12 points instead of 15 points, to be followed by a marginal decline between 2015 and 2019. The average score in 2019 is thus predicted to have been 10 points better than in 2011, instead of 15 points better. Finally, the difference between the actual and predicted average TIMSS score in 2019 amounts to 20 points (or 0.24 standard deviations).



Of course, it is important to interpret the results from the simulation with caution. It is impossible to determine conclusively how Sweden's score trend in TIMSS would have looked like without increasing independent-school competition since 1995. Yet the analysis gives a rough estimation of how Swedish average performance had changed over time without any increase in independent-school competition, based on the empirical estimates in this paper.

6. Conclusions

This paper has analysed the impact of increasing independent-school competition on Sweden's scores in TIMSS, an externally marked international test measuring performance in mathematics and science among students in year 8. I found that independent-school competition has increased Swedish students' performance compared with a counterfactual situation without competition. The results imply that an increase of 10 percentage points in the average municipal independent-school enrolment share at the county level improves TIMSS scores by about 20–25 points on average, equivalent to roughly 0.24–0.30 standard deviations, an effect that only emerges from 2007 onwards. It also appears that competition both decreased the performance decline in the period 1995–2011 and contributed to the performance rise in the period 2011–2019. A simulation based on the estimates in the paper suggests that Sweden would have scored on average 20 points (or 0.24 standard deviations) lower in TIMSS 2019 had the average share of students attending independent schools remained at the level of 1995.

Interestingly, the positive impact appears to be driven by competition from for-profit independent schools, which is not surprising given that the increase in competition overall can be mostly attributed to the establishment of such schools. Also, the effect is almost identical when adjusting for the direct association between independent-school attendance and TIMSS scores, and when excluding all participating independent-school students in the analysis, suggesting that competition affects students in municipal and independent schools similarly. Apart from having a positive impact on students' average scores, competition also raises the probability that students reach the intermediate, high, and advanced international benchmarks, indicating that competition has positive effects among both lower- and higher-performing students.

Given the nature of the data analysed, it is not possible to conduct checks on pre-reform trends in support of a causal interpretation of these findings. Yet in a placebo test in treatment, I found no evidence that changes in the share of the variation in independent pre-school enrolment shares that cannot be predicted by independent-school enrolment shares were positively related to changes in TIMSS scores. This is expected if the findings are driven by school competition, rather than other time-varying regional differences that correlate with differential changes in preferences for private providers more generally.

Overall, the paper thus supports prior research in finding positive effects of independent-school competition on lower-secondary student performance in Sweden. Importantly, little suggests that estimates in prior research are biased upwards due to unreliable domestic performance metrics. On the contrary, the effects in this paper are larger than those found in prior research analysing domestic outcomes. A comparison suggests this is primarily because I analyse low-stakes international test scores, which both neutralise all factors behind differential marking and grading standards as well as partly reflect different types of knowledge and skills compared with domestic outcomes, and because I estimate effects at the county level rather than municipal level.

Certainly, these findings do not imply that the effects of the voucher reform have been positive in all respects. For example, other research suggests it may also have increased school segregation at the lower-secondary level (Böhlmark et al. 2016). Furthermore, the paper is silent on the effects of competition on student performance at the upper-secondary level, where the share of students attending independent schools has increased about twice as much. More generally, the education system and market are very different at the upper-secondary level, making it impossible to extrapolate this paper's effects in this respect. In recently commenced research, I thus intend to analyse how independent-school competition at the upper-secondary level has affected Sweden's performance trend in TIMSS Advanced, an international test in advanced mathematics and physics among students in the final year of upper-secondary school. This will allow me to provide further evidence of the effects of independent-school competition at the upper-secondary level as well.

References

Ahlin, Åsa. 2003. "Does School Competition Matter? Effects of a Large-Scale School Choice Reform on Student Performance." Working Paper, No. 2003:2, Department of Economics, Uppsala University.

Böhlmark, Anders, Helena Holmlund, and Mikael Lindahl. 2016. "Parental choice, neighbourhood segregation or cream skimming? An analysis of school segregation after a generalized choice reform." *Journal of Population Economics* 29:1155–1190.

Böhlmark, Anders and Mikael Lindahl. 2015. "Independent schools and long-run educational outcomes: Evidence from Sweden's large-scale voucher reform." *Economica* 82(327):508–551.

Borghans, Lex, Bart H. H. Golsteyn, and James J. Heckman. 2016. "What grades and achievement tests measure." *PNAS* 113(47): 13354–13359.

Björklund, Anders, Melissa A. Clark, Per-Anders Edin, Peter Fredriksson, and Alan B. Krueger. 2005. *The Market Comes to Education in Sweden: An Evaluation of Sweden's Surprising School Reforms*. New York: Russell Sage Foundation.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with a Continuous Treatment." Working Paper, arXiv:2107.02637v2.

Cameron, A. Colin and Douglas Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2):317–372.

Edmark, Karin. 2015. "Location choices of Swedish independent schools." *Annals of Regional Science* 63:219–239.

Falck, Oliver and Ludger Woessmann. 2013. "School competition and students' entrepreneurial intentions: International evidence using historical Catholic roots of private schooling." *Small Business Economics* 40(2):459-478.

Frändberg, Birgitta and Mats Hagman. 2017. ”Med fokus på naturorienterade ämnen – Analys av samstämmigheten mellan svenska styrdokument i NO och de internationella studierna TIMSS 2015 och PISA 2015.” Report, National Agency for Education, Stockholm.

Fisman, Ray. 2014. “Sweden's School Choice Disaster.” Article, *Slate,* 15 July, https://slate.com/news-and-politics/2014/07/sweden-school-choice-the-countrys-disastrous-experiment-with-milton-friedman-and-vouchers.html.

Gneezy, Uri, John A. List, Jeffrey A. Livingstone, Xiangdong Qin, Sally Sadoff, and Yang Xu. 2019. “Measuring Success in Education: The Role of Effort on the Test Itself.” *American Economic Review: Insights* 1(3):291–308.

Goodman-Bacon, Andrew. 2019. “Difference-in-differences with variation in treatment timing.” *Journal of Econometrics* 225(2):254–277.

Heller-Sahlgren, Gabriel. 2022. ”Kunskapsskolans återkomst.” Report, Timbro, Stockholm.

Hensvik, Lena. 2012. ”Competition, Wages and Teacher Sorting: Lessons Learned from a Voucher Reform.” *Economic Journal* 122(561):799–824.

Holmlund, Helena, Josefin Häggblom, Erica Lindahl, Sara Martinson, Anna Sjögren, Ulrika Vikman, and Björn Öckert. 2014. “Decentralisering, skolval och fristående skolor – resultat och likvärdighet i svensk skola.” Report 2014:25, Institute for Evaluation of Labour Market and Education Policy (IFAU), Uppsala.

Holmlund, Helena, Anna Sjögren, and Björn Öckert. 2019. “Jämlikhet i möjligheter och utfall i den svenska skolan.” Appendix 7 in the 2019 Long-Term Survey of the Swedish Economy, SOU 2019:40.

Hoxby, Caroline M. 2003. “School choice and school competition: Evidence from the United States.” *Swedish Economic Policy Review* 10:9–65.

IEA. 2022. Data retrieved from the international database: https://www.iea.nl/data-tools/repository/timss.

Jerrim, John, Luis Alejandro Lopez-Agudo, Oscar D. Marcenaro-Gutierrez, and Nikki Shure. 2023. ” What happens when econometrics and psychometrics collide? An example using the PISA data” *Economics of Education Review* 61:51–58.

Meer, Jonathan and Jeremy West. 2016. “Effects of the Minimum Wage on Employment Dynamics.” *Journal of Human Resources* 51(2):500–522.

NAE. 2017. “TIMSS, TIMSS Advanced och betygen – Analys av sambandet mellan svenska betyg och de internationella TIMSS-studierna.” Report, National Agency for Education, Stockholm.

NAE. 2019. “Analyser av likvärdig betygssättning mellan elevgrupper och skolor – Jämförelser mellan betyg och nationella prov i årskurs 9.” Report, National Agency for Education, Stockholm.

NAE. 2020. “Likvärdiga betyg och meritvärden – Ett kunskapsunderlag om modeller för att främja betygens och meritvärdenas likvärdighet.” Report, National Agency for Education, Stockholm.

NAE. 2021. Data retrieved from the National Agency for Education's databases and annual reports.

OECD. 2019. “OECD Economic Surveys: Sweden 2019”, Report, OECD Publishing, Paris.

Proposition 2008/09:171. 2009. “Offentliga bidrag på lika villkor”, Government bill, Stockholm.

MacLeod, W. Bentley and Miguel Urquiola. 2012. ”Competition and Educational Productivity: Incentives Writ Large.” IZA Discussion Paper No. 7063.

MacLeod, W. Bentley and Miguel Urquiola. 2019. ” Is Education Consumption or Investment? Implications for School Competition” *Annual Review of Economics* 11:563–589.

Martin, Michael O., Matthias von Davier, and Ina V.S. Mullis. 2022. Methods and Procedures: TIMSS 2019 Technical Report.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. 2019. ”Fast and wild: Bootstrap inference in Stata using boottest.” *Stata Journal* 19(1): 10.1177/1536867X19830877.

Sandström, F. Mikael and Fredrik Bergström. 2005. “School vouchers in practice: competition will not hurt you.” *Journal of Public Economics* 89(2–3):351–380.

Sollerman, Samuel and Astrid Pettersson. 2016. ”Med fokus på matematik – Analys av samstämmigheten mellan svenska styrdokument och den internationella studien TIMSS 2015.” Report, National Agency for Education, Stockholm.

Statistics Sweden. 2021. Data retrieved from Statistics Sweden.

Tyrefors Hinnerich, Björn and Jonas Vlachos. 2013. “Systematiska skillnader mellan interna och externa bedömningar av nationella prov.” Appendix 5 in “Olikheterna är för stora”, Report, Swedish Schools Inspectorate, Stockholm.

Vlachos, Jonas. 2010. “Konkurrensens konsekvenser – En analys av hur konkurrens påverkar betygssättningen vid svenska skolor.” Rapport, Swedish Competition Authority, Stockholm.

Wennström, Johan. 2020. “Marketized Education: How Regulatory Failure Undermined the Swedish School System.” *Journal of Education Policy*, 35(5):665–691.

Wiberg, Marie and Ewa Rolfsman. 2023. ”Students’ Self-reported Background SES Measures in TIMSS in Relation to Register SES Measures When Analysing Students’ Achievements in Sweden.” *Scandinavian Journal of Educational Research* 67(1):69–82.

Wu, Margaret. 2005. ”The role of plausible values in large-scale surveys.” *Studies in Educational Evaluation* 31(2–3):114–128.

Zamarro, Gema, Collin Hitt, and Ildefonso Mendez. 2019. ”When Students Don’t Care: Reexamining International Differences in Achievement and Student Effort.” *Journal of Human Capital* 13(4):519–552.

Appendix

Table A1: Descriptive statistics

	1995	2003	2007	2011	2015	2019
Average TIMSS score	546.13	511.66	501.00	496.90	511.50	511.95
	(73.12)	(66.60)	(71.07)	(70.69)	(75.11)	(82.38)
Reached low level (both subjects)	0.94	0.89	0.87	0.86	0.88	0.86
	(0.24)	(0.31)	(0.34)	(0.35)	(0.33)	(0.35)
Reached intermediate level (both subjects)	0.74	0.58	0.56	0.53	0.61	0.60
	(0.44)	(0.49)	(0.50)	(0.50)	(0.49)	(0.49)
Reached high level (both subjects)	0.37	0.19	0.17	0.14	0.23	0.24
	(0.48)	(0.39)	(0.38)	(0.35)	(0.42)	(0.43)
Reached advanced level (both subjects)	0.08	0.02	0.02	0.01	0.02	0.04
	(0.26)	(0.13)	(0.13)	(0.10)	(0.15)	(0.20)
Books at home (5 categories)	3.94	3.58	3.40	3.23	3.01	2.90
	(1.09)	(1.24)	(1.25)	(1.29)	(1.31)	(1.35)
Foreign born	0.07	0.09	0.08	0.09	0.11	0.14
	(0.26)	(0.29)	(0.27)	(0.28)	(0.32)	(0.35)
Mother foreign born	0.13	0.19	0.18	0.20	0.23	0.27
	(0.33)	(0.39)	(0.38)	(0.40)	(0.42)	(0.44)
Father foreign born	0.13	0.20	0.18	0.21	0.25	0.29
	(0.39)	(0.44)	(0.43)	(0.42)	(0.52)	(0.52)
Age at arrival in Sweden	0.38	0.34	0.26	0.21	0.33	0.33
	(1.80)	(1.27)	(1.11)	(0.82)	(1.16)	(0.95)
Speak Swedish at home (3 categories)	3.08	2.92	2.93	2.91	2.90	2.84
	(1.04)	(0.31)	(0.29)	(0.33)	(0.35)	(0.44)
Boy	0.51	0.49	0.52	0.52	0.52	0.51
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age	14.93	14.89	14.83	14.77	14.75	14.83
	(0.31)	(0.35)	((0.36)	(0.40)	(0.36)	(0.36)
Computer at home	0.97	0.98	0.98	0.99	0.98	0.98
	(0.17)	(0.16)	(0.15)	(0.12)	(0.12)	(0.16)
Study desk at home	0.98	0.98	0.98	0.98	0.99	0.97
	(0.13)	(0.15)	(0.15)	(0.13)	(0.10)	(0.16)
Independent-school enrolment share	0.01	0.04	0.06	0.09	0.10	0.12
	(0.01)	(0.03)	(0.05)	(0.06)	(0.07)	(0.07)
For-profit enrolment share (years 6–9)	0.01			0.07	0.09	0.11
	(0.01)			(0.05)	(0.07)	(0.07)
Non-profit enrolment share (years 6–9)	0.00			0.03	0.03	0.04
	(0.00)			(0.02)	(0.02)	(0.03)
Books at home (county)	3.96	3.55	3.43	3.21	3.05	2.92
	(0.23)	(0.14)	(0.19)	(0.20)	(0.24)	(0.25)
Share of students with some foreign background (county)	0.19	0.26	0.25	0.28	0.30	0.33
	(0.08)	(0.11)	(0.09)	(0.10)	(0.08)	(0.08)
(log) Number of year 9 students (county)	8.72	8.86	9.02	8.95	8.90	9.03
	(0.82)	(0.91)	(0.88)	(0.93)	(0.92)	(0.98)
Number of students in TIMSS	1,949	4,256	5,215	5,573	4,090	3,996

Note: The data display means with standard deviations in parentheses. All statistics are calculated using student-level TIMSS data, weighted by participating students' sampling probabilities. Only observations without missing values are used. The independent-school enrolment shares are given as the municipal averages at the county level.

Table A2. The impact of independent-school competition on average TIMSS scores

Main analysis TIMSS 1995, 2003, 2011, 2019			
Independent-school enrolment share	208.61* (110.96)	224.79** (84.74)	245.07*** (59.29)
County-fixed effects	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes
Student-level controls	No	Yes	Yes
County-level controls	No	No	Yes
P-value (clustered standard errors)	0.08	0.02	<0.01
P-value (wild bootstrap, 99999 replications)	0.07	0.03	<0.01
Students		15,750	
Counties		21	
Excluding county-fixed effects			
Independent-school enrolment share	75.71 (67.11)	63.31 (46.74)	3.12 (35.98)
County-fixed effects	No	No	No
Time-fixed effects	Yes	Yes	Yes
Student-level controls	No	Yes	Yes
County-level controls	No	No	Yes
Students		15,750	
Counties		21	

Note: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the county level in parentheses. The background controls constructed from survey data are interacted with the year-fixed effects. All counties are given equal weight in the analysis. In the comparison of the p-values with clustering and wild bootstrap, the average of all five plausible values is analysed, thus ignoring the uncertainty in students' test scores that arise from the multiple imputation procedure to obtain the plausible values.

Table A3. The variation in the effect of independent-school competition over time

	1995 & 2003	1995 & 2007	1995 & 2011	1995 & 2015	1995 & 2019
Independent-school enrolment share	35.10 (241.21)	383.87** (171.73)	282.43*** (92.75)	239.24*** (77.20)	257.00*** (72.86)
Students	6,205	7,164	7,498	6,000	5,945
Counties	21	21	21	20	21

Note: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the county level in parentheses. All models include student- and county-level controls. The background controls constructed from survey data are interacted with the year-fixed effects. All counties are given equal weight in the analysis.

Table A4. The effects during the periods of falling and improving scores

	1995–2011	1995, 2003 & 2011	2011–2019	2011 & 2015	2011 & 2019
Independent-school enrolment share	264.03***	254.80**	422.36***	632.67**	402.15**
	(88.58)	(91.39)	(132.74)	(264.64)	(157.72)
Students	16,969	11,754	13,635	9,617	9,545
Counties	21	21	21	20	21

Note: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the county level in parentheses. All models include student- and county-level controls. The background controls constructed from survey data are interacted with the year-fixed effects. All counties are given equal weight in the analysis.

Table A5. The impact of competition from for-profit and non-profit independent schools

	1995, 2011, 2015 & 2019	1995 & 2019
Variables included in the same model		
For-profit school enrolment share	214.57***	242.54***
	(64.38)	(79.71)
Non-profit school enrolment share	14.80	-147.45
	(207.07)	(184.32)
Variables included in separate models		
For-profit school enrolment share	216.03***	232.94***
	(56.51)	(73.22)
Non-profit school enrolment share	235.45	66.39
	(209.01)	(253.55)
Students	15,584	5,945
Counties	21	21

Note: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the county level in parentheses. All models include student- and county-level controls. The background controls constructed from survey data are interacted with the year-fixed effects. All counties are given equal weight in the analysis. The for-profit and non-profit school enrolment shares are calculated among students in years 6–9 only.

Tabell A6. Effects on the probability of reaching international performance benchmarks

TIMSS 1995, 2003, 2011, and 2019			
	Coefficient	Students	Counties
Reached low level in both subjects	0.42	15,750	21
	(0.37)		
Reached intermediate level in both subjects	1.19**	15,750	21
	(0.48)		
Reached high level in both subjects	1.12***	15,750	21
	(0.39)		
Reached advanced level in both subjects	0.37**	15,750	21
	(0.15)		

Note: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the county level in parentheses. All models include student- and county-level controls. The background controls constructed from survey data are interacted with the year-fixed effects. All counties are given equal weight in the analysis.

Tabell A7. Further analyses

	Coefficient	Students	Counties
All test rounds included	238.05***	25,055	21
	(56.41)		
Skåne, Stockholm, and Västra Götaland counties excluded	263.58***	7,616	18
	(72.10)		
Median regression	165.53***	15,750	21
	(54.52)		
OLS with median regression specification	180.76***	15,750	21
	(48.68)		
Mathematics score	234.26***	15,750	21
	(69.28)		
Science score	255.88***	15,750	21
	(61.90)		
Low level in mathematics	0.35	15,750	21
	(0.34)		
Intermediate level in mathematics	1.11**	15,750	21
	(0.46)		
High level in mathematics	1.15***	15,750	21
	(0.32)		
Advanced level in mathematics	0.41**	15,750	21
	(0.20)		
Low level in science	0.15	15,750	21
	(0.23)		
Intermediate level in science	0.70	15,750	21
	(0.48)		
High level in science	1.49***	15,750	21
	(0.41)		
Advanced level in science	0.83***	15,750	21
	(0.26)		
Competition measure calculated at school municipality level	257.42***	15,750	21
	(63.11)		
Independent-school enrolment share in the county overall	243.39***	15,750	21
	(82.92)		
Control for independent-school indicator	250.10***	15,532	21
	(61.07)		
Students in independent schools excluded	261.50***	14,108	21
	(67.48)		
Control for enrolment shares in other municipalities' schools	244.70***	15,750	21
	(61.10)		
Unadjusted student weights	194.36***	15,750	21
	(46.60)		
Unadjusted student weights & all surveys (for simulation)	192.48***	25,055	21
	(43.70)		
No weights	189.25***	15,750	21
	(50.82)		
Independent-school enrolment share, years 6–9 (1995, 2011, 2015 & 2019)	184.50***	15,584	21
	(47.52)		
Independent-school enrolment share, years 6–9 (1995 & 2019)	197.53***	5,945	21
	(57.09)		
Placebo test (independent pre-school enrolment, residual)	-71.63	15,750	21
	(76.57)		

Note: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the county level in parentheses. All models include student- and county-level controls. The background controls constructed from survey data are interacted with the year-fixed effects. All counties are given equal weight in the analysis, with the exception for the median regression and its OLS equivalent, the models that include unadjusted weights, and the model that include no weights.