WORKING PAPER 2018:26

Strengthening teachers in disadvantaged schools:

Evidence from an intervention in Sweden's poorest city districts

Caroline Hall
Martin Lundin
Kristina Sibbmark

IFAU
INSTITUTET FÖR ARBETSMARKNADS- OCH UTBILDNINGSPOLITISK UTVÄRDERING

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

Strengthening teachers in disadvantaged schools:

Evidence from an intervention in Sweden's poorest city districts[a]

by

Caroline Hall[b], Martin Lundin[c] and Kristina Sibbmark[d]

December 14, 2018

Abstract

Children growing up in disadvantaged neighborhoods tend to perform significantly worse in school compared to children growing up under more favorable circumstances. We examine the impact of a three-year program ("Coaching for Teaching") targeted at ten poorly performing lower secondary schools in Sweden's most disadvantaged city districts. The aim of the intervention was to strengthen the teachers' professional development, e.g. through coaching and further training, and thereby enhance student performance. We use a difference-in-differences design and rich register data to estimate effects on several educational outcomes. Our results show a large and statistically significant positive impact on student performance on standardized tests in English language. Estimates for test results in math are also positive and large, but not statistically significant; the same applies to GPA and admission to upper secondary school. For test scores in Swedish language there is no indication of improvement. An analysis of a survey of pupils supports the idea that the teaching as well as the classroom climate improved due to the intervention. Taken together, the program seems to have generated rather promising results in the short run.

Keywords: Education, disadvantaged schools, lower secondary school, social background, teachers, professional development, student performance, government policy
JEL-codes: I20, I21, I24, I28, J24

[a] We are grateful to Anna Sjögren, Olof Åslund, and Björn Öckert as well as seminar participants at IFAU and the IFAU-UCLS conference "How to improve school results? Policy design, implementation and evaluation" for valuable comments.
[b] IFAU and Uppsala Center for Labor Studies (UCLS); email: *caroline.hall@ifau.uu.se*.
[c] IFAU and Uppsala Center for Labor Studies (UCLS); email: *martin.lundin@ifau.uu.se*.
[d] IFAU; email: *kristina.sibbmark@ifau.uu.se*.

Table of contents

1 Introduction

Improving schooling outcomes among disadvantaged children and thereby reducing inequalities in life chances is a key challenge for public education systems worldwide. Children growing up in disadvantaged neighborhoods – characterized by poverty, unemployment, bad health and crime – tend to perform significantly worse in school than children growing up under more favorable circumstances (see e.g. Sass et al. 2012; Curto and Fryer 2013). A major difficulty faced by schools in these neighborhoods is attracting and retaining effective teachers (Sass et al. 2012; Chiang, Clark and McConnell 2017). Several studies find that teachers tend to prefer to work in schools that serve students from more well-off families, and that teacher turnover is higher in schools with more disadvantaged students (e.g., Lankford, Loeb and Wyckoff 2002; Hanushek, Kain and Rivkin 2004; Scafidi, Sjoquist, and Stinebrickner 2007). At the same time, previous empirical research identifies teacher quality as a crucial factor for enhancing student performance (e.g. Rockoff 2004; Rivkin, Hanushek and Kain 2005; Aaronson, Barrow and Sander 2007).

A number of papers have examined the effectiveness of programs directed towards supplying disadvantaged schools with more effective teachers, e.g. through alternative qualification routes, with somewhat mixed findings on student performance.[1] In this paper we study an intervention that, instead of aiming to attract high-quality teachers to disadvantaged schools, aimed at enhancing student performance through strengthening professional development among the existing pool of teachers.[2] The program, which started in the school year of 2012/13 and lasted three years, was targeted at ten poorly performing lower secondary schools in Sweden's most disadvantaged city districts (often referred to as "urban development districts"). The targeted districts all have low employment rates, high reliance on social assistance, and a relatively high share of pupils that do not manage to qualify for entering high school. The schools in these

[1] For instance, Chiang, Clark and McConnell (2017) find that math teachers from the alternative certification program "Teach for America" are more effective than other math teachers in high-poverty secondary schools in the US. Glazerman, Mayer and Decker (2006) find that also elementary school students with TFA teachers outperform students of other teachers in math, but not in reading, while Clark et al. (2015) find no significant difference on average in neither math nor reading.

[2] Efforts to increase teacher effectiveness have also been conducted through offering teachers financial incentives based on student achievement. While positive effects have been found for such incentive schemes in developing countries (Glewwe, Ilias and Kremer 2010; Muralidharan and Sundararaman 2011), most evaluations in the US do not find evidence of enhanced student performance (Springer et al. 2010; Goodman and Turner 2010; Fryer 2013), an exception is Fryer et al. (2012).

districts have a high share of foreign-born pupils, including many newly arrived refugee children (Assadi et al. 2015).

The intervention, which was called "Coaching for Teaching" (CFT) (*Handledning för lärande*), was initiated by the Swedish government as an effort to identify successful strategies for raising academic results in disadvantaged city districts. The main component of the program was coaching, both in smaller groups and, if the teacher desired, individually. All teachers participated in group meetings, with a trained coach, every three weeks throughout the program period. The individual coaching was usually given at eight different occasions. A second important part of the program consisted of courses and seminars aiming to enhance teachers' skills. For instance, almost all teachers participated in a quite extensive course on knowledge and language enhancing teaching strategies, which had the explicit purpose of improving the teaching of immigrant children with poor ability in the Swedish language. The intervention also contained a few additional components that in practice were somewhat less important (the intervention is described in further detail in Section 3).

Out of the around 25 schools that were eligible for the intervention, based on being located in the targeted city districts, 10 were selected for participation. Although randomly assigning schools to treatment was not feasible, the selection process tried to ensure that treated and control schools would be balanced in terms of pre-treatment trends in student performance. We use a difference-in-differences design and rich register data to estimate the effect of the CFT intervention on students' results on standardized test in math, Swedish and English language as well as on GPA and admission to upper secondary school.

Our results show a large and statistically significant positive impact on student performance on the standardized tests in English language. The estimates for test results in math are also positive and large, but not statistically significant at conventional levels; the same applies to GPA and admission to upper secondary school. There is no indication of improvement of test scores in Swedish language. We also analyze a questionnaire with pupils, taken both before and towards the end of the intervention. These results indicate positive effects on students' perceptions of how their teachers teach and the classroom climate. Taken together, the program seems to have generated rather promising results in the short run.

It is not possible to disentangle the precise mechanisms behind our findings. However, the results suggesting a positive impact on how students perceive the teaching indicate that the program indeed increased teacher skills and/or dedication. A qualitative analysis by Assadi et al. (2015), based on in-depth interviews with teachers participating in the CFT program, suggests that the teachers' opinion about the usefulness of the group coaching varies a lot: some teachers are positive, but most are neutral or critical to this part of the program. One-to-one coaching is regarded as more valuable. Moreover, the teachers highly valued the course on knowledge and language enhancing teaching strategies. Hence, the results in Assadi et al. (2015) indicate that individual coaching and further training of teachers may be valuable parts of the program. Overall, our findings are in line with the idea that efforts to strengthen teachers' skills, confidence and motivation could be a way to enhance the performance of students in disadvantaged schools.

Our paper is related to the broader literature on how to improve schooling outcomes in disadvantaged city districts, and among children with disadvantaged background more generally (see e.g. Curto, Fryer and Howard 2011 and Cullen et al. 2013 for reviews of various programs). While promising outcomes have been documented for several interventions – e.g. smaller classes (Krueger 1999; Krueger and Whitmore 2001; Chetty et al. 2011), smaller schools (Bloom et al. 2010), individualized math instruction (Cook et al. 2015), and mandatory summer school (Jacob and Lefgren 2004) – the success of many programs is likely to be contingent on the presence of effective teachers. In fact, the results in Angrist et al. (2013) indicate that selective hiring of teachers is one of the crucial components behind the success of several urban charter schools that serve many minority and low-income students in the US.[3] Hence, low-performing schools that are looking to implement previously promising interventions, most likely also need to ascertain an effective pool of teachers in order to be equally successful.

The paper is organized as follows: In Section 2 we describe the institutional setting and in Section 3 the data and the empirical strategy. Thereafter, in Section 4, we present our results. We begin with the baseline results on educational performance, followed by robustness checks. Then we present our analysis of teacher turnover and the pupil

[3] Fryer (2014) shows that implementing a bundle of successful charter school practices, among them more effective teachers, into low-performing public schools also improved math achievement in these schools.

questionnaire with the intention to better understand the main effects. Finally, our findings are summarized and discussed in Section 5.

2 Institutional setting

2.1 The Swedish education system

Sweden has nine years of compulsory schooling with a comprehensive curriculum. Children start school in the fall of the year they turn seven. Grade repetition during compulsory schooling is rare and, hence, most finish 9th grade the year they turn 16. Traditionally, compulsory schooling has been organized in three stages: grades 1–3, 4–6 (lower vs. upper primary school) and 7–9 (lower secondary school). These days the organization is more flexible and the stages are no longer as distinct.

After compulsory school, almost all pupils continue to upper secondary education, which consists of several different educational programs. Admission is based on students' 9th grade GPA. Pupils that have not attained eligibility for a regular upper secondary school program have the possibility of instead participating in an introductory program. Through this route they can qualify for a regular program.

Officially, Sweden has a system of rather far-reaching school choice. In principle, families can choose any public or "independent" (but publicly funded) compulsory school for their children.[4] Nevertheless, it is still most common that children attend the nearest public school.[5] An important reason for this is that admission to public schools, for grades 1–9, is based on proximity to the school. Independent schools may, on top of proximity, also base admission on a first-come-first-served basis, but they cannot select pupils based on personal characteristics, such as ability or family background. To receive public funding, they are also not allowed to charge a tuition fee. About 15 percent of the pupils in compulsory school attended an independent school in the school year of 2017/18 (National Agency for Education 2018).

2.2 Urban development districts

Like many other countries, Sweden has neighborhoods with high levels of social exclusion. To promote positive developments in such neighborhoods, the government

[4] Note that there are very few fully private schools in Sweden.
[5] Böhlmark, Holmlund and Lindahl (2015, p. 45) approximate that around 30 percent of the pupils in 2009 opted out from their assigned public school.

often directs specific support to the city districts that are considered most disadvantaged, often referred to as "urban development districts". From 2012, 15 city districts with low employment rates, high reliance on social assistance and poor schooling outcomes were subject to such support.[6] The support was intended to focus on evaluation, knowledge acquisition and exchange as well as dissemination of knowledge (Government decision A2012/174/IU). During 2013- 2014, the government devoted financial resources, amounting to SEK 100 million per year (approximately USD 11.6 million), to stimulate urban development in these areas.[7] They also launched the project "Coaching for Teaching", which we describe in detail in Section 3.

Poor schooling outcomes is one of the defining characteristics of the urban development districts. The first column of Table 1 displays average school results for pupils finishing compulsory school in these districts in 2012, while the second column displays the same averages for students graduating from other schools. We can see that students in the urban development areas have considerably lower GPA (165 vs. 210) as well as results on the standardized tests (percentile rank 38 vs. 50). They are also much less likely to have obtained the qualifications necessary to continue to upper secondary education. The table also shows that their background characteristics are remarkably different: For instance, about 36 percent are born in another county, compared to 9 percent for the rest of the schools; and 78 percent have two parents that are born abroad, while the corresponding number for the other schools is 17 percent. The students are also much more likely to have recently immigrated to Sweden. Their parents have on average substantially lower earnings, education level, and are four times more likely to receive social assistance.

Schools located in urban development districts can be described as turbulent contexts (see e.g. National Agency for Education 2015). Social problems tend to spill over on the schools' daily activities, teacher and principal turnover is high, and the schools are used to participate in various development projects.

[6] The formal criteria for being considered an urban development district was having i) an employment rate below 52 percent; (ii) a social assistance dependency rate above 4.8 percent; and (iii) less than 70 percent of students being qualified to enter upper secondary education. The districts also needed to have a population size exceeding 4,000. (Government decision A2012/174/IU)

[7] The financial support was distributed according to a model based on improved schooling, employment, and social assistance outcomes, where municipalities were given more financial resources if the targeted districts managed to improve their results (Government Bill 2012/13:1).

Table 1. Characteristics of pupils graduating from compulsory school in urban development districts and all other districts, 2012

	Schools in urban development districts	All other schools
GPA, 9th grade	164.7	210.4
Average results on the standardized test, 9th grade (percentile ranked)	37.7	50.2
Eligible for upper secondary education	0.61	0.88
Girls	0.52	0.49
Born in another country	0.36	0.09
Immigrated 0- 4 years ago	0.14	0.03
Both parents are born in another country	0.78	0.17
Parents receive social assistance	0.35	0.08
Mother has post-secondary education	0.17	0.40
Father has post-secondary education	0.18	0.31
Mother's earnings (SEK)	119,119	252,575
Father's earnings (SEK)	164,429	358,898

Note: All variables originate from Statistics Sweden and are described in more detail in Section 4. Average test result refers to students' average grade on the standardized tests in math, English and Swedish. Parents' education, earnings, and receipt of social assistance are measured the year before the student finished 9th grade. Average earnings are computed among parents who were Swedish residents that year.

3 The "Coaching for Teaching" program

The CFT intervention took place in ten public lower secondary schools (grades 7–9), located in the urban development districts, during the three school years 2012/13 – 2014/15. The government provided the financial resources and stated the program's overall orientation and aim, and the National Agency for Education (*Skolverket*) was given the task of formulating its more precise content.

As outlined in the Introduction, the program's primary focus was to enhance teachers' professional development, with the goal of improving the instruction and thereby student performance. The intervention's main component was coaching. All teachers participated in coaching in small groups every three weeks throughout the program period. The idea was that teachers, through these group discussions, would strengthen their self-confidence and ability to handle the challenges of the teaching

profession.[8] The coaches were especially qualified teachers hired and trained by the National Agency for Education. In addition, teachers that were interested could receive one-to-one coaching, usually at eight occasions.[9] The principals were also offered coaching in groups and individually, and most of them took part. The purpose of the coaching towards principals was to include them in the program, motivate them and improve their ability to lead the intervention locally (Assadi et al. 2015).

As a second crucial part of the program, the agency organized several courses and seminars with the aim of improving the teachers' skills. Nearly all teachers participated in an extensive course that had the aim of improving teaching towards pupils with poor ability in the Swedish language. This course focused on teaching strategies (including concrete pedagogical tools) to enhance knowledge and language development among immigrant pupils. The course was organized by the National Center for Swedish as a Second Language.[10] In addition, further training was offered on topics such as formative assessment, subjective didactics and motivational teaching methods. The extent to which these courses and seminars were offered to teachers differed between schools (Assadi et al. 2015).

The schools were also given the opportunity to apply for extra funding for a variety of purposes: For organizing teaching outside regular school hours (e.g., homework assistance and teaching during school breaks), for training tutors that could assist newly arrived immigrant pupils in their native language (*studiehandledare*), and for efforts to improve communication between teachers and parents. Lastly, the National Agency for Education developed some support material related to the various parts of the program; this material was made available to all participating schools (Assadi et al. 2015).

[8] The schools were responsible for organizing the groups, which each consisted of 3–12 teachers. The meetings usually lasted around 1.5 hours. The teachers decided what to discuss (one question each meeting) and everything that was said in the discussions was confidential. A trained external coach chaired the meetings. Usually, the same group constellation was used throughout the program period (Assadi et al. 2015).

[9] Around 25 percent (94 individuals) of the teachers participated in individual coaching. The individual coaching usually took place for three months (Assadi et al. 2015).

[10] The National Center for Swedish as a Second Language is part of Stockholm University. The teachers participating in the course met eight work days spread out over a year. They also had assignments between the meetings.

Table 2. Allocation of resources in the "Coaching for Teaching" intervention, 2012–2014

	SEK	Percent
Group coaching	10 304 119	22.8
Ono-to-one coaching	2 457 724	5.4
Coaching of principals	1 787 500	4.0
Further training of teachers	8 157 595	18.0
Extra funding for a variety of purposes	7 466 701	16.5
Administration and implementation of the program	15 062 588	33.3
Total	45 245 227	100.0

Note: "Coaching of principals" also includes resources spent on further training of principals. Schools were able to apply for extra funding for the following purposes: organizing teaching outside regular school hours, for training certain tutors that could help newly arrived immigrants in their native language, and for improving communication between teachers and parents.

Table 2 shows how the resources were allocated: One third of the program budget was used for costs associated with the administration and implementation of the program, one third for coaching activities and the last third was split almost evenly between further training and extra funding for a variety of purposes. Altogether, the resources used within the project correspond to around 1.4 extra teachers per school and year. The intervention schools employed on average 35 teachers each, which means that an alternative use of the resources would have been to increase the teaching staff by approximately 5 percent.[11]

A qualitative analysis of the CFT intervention, based on interviews with teachers and principals, is included in Assadi et al. (2015). The analysis shows that the program had a clear influence on the schools. Schools in disadvantaged areas often participate in various projects but in comparison to other interventions, CFT is regarded as an extensive one by both teachers and principals. Especially the various efforts to strengthening teachers' competence, through coaching and further training, affected daily work life throughout the program period. The opinions about the usefulness of the group coaching, which was the most visible part of the program, diverge. Some teachers found it to be, basically, a waste of time. Others, however, said that they had benefited from the group discussions, although most of them could not pinpoint more precisely what they had learned. The individual coaching and training activities (especially the

[11] The budget was 60 000 000 SEK (around USD 7 million). The intervention started somewhat later than planned, which is one explanation why only 75 percent of the original budget was used.

course on knowledge and language enhancing teaching methods) were regarded as much more valuable.

As noted above, the schools could also apply for extra funding for activities such as homework assistance, teaching during school breaks, training of tutors and for improving communication between teachers and parents. The results in Assadi et al. (2015) show that the schools did apply for extra funding, but that local practices did not changed that much. For instance, the schools were already offering homework assistance and teaching during school breaks. We cannot conclude that these parts of the CFT program were unimportant, but it seems like activities aiming at improving teachers' professional development (e.g. coaching and training) were the most prominent parts of the project.

3.1 The selection of schools for participation

The National Agency for Education was given the task of selecting schools for participation. As stated above, the program was directed towards schools in urban development areas with poor academic results. Only public schools were eligible. The government also explicitly stated that the schools were to be selected in a way that would enable evaluation of the program's impact on student performance. We had the opportunity to advice the agency regarding the selection of schools, although they themselves made the final decision on the selection process.

In order to try to ensure there would exist a control group of schools with similar pre-treatment trends in student performance, we created ten pairs of schools (located in the same municipality) with similar developments in terms of GPA and eligibility for upper secondary education during the years preceding the intervention (as similar as possible, given the limited number of schools in the relevant areas). Only one of the schools within each pair was then given the opportunity to participate. Although randomization within pairs would have been ideal for the purpose of evaluation, the National Agency for Education did not believe this to be a feasible strategy in this case. The agency instead chose the school they found most suitable for participation within each pair. Reasons for why schools sometimes were considered less suitable included, e.g., lack of interest on the part of the principal, concerns that the school would be closed down in the near future, and awareness that the school recently had been the target of other types

of interventions.[12] A couple of control schools in the initial pairs of schools in fact turned out to be located outside the boundaries of the targeted districts. These schools will be excluded from the analyses; we instead use all non-treated schools within the targeted districts as the control group (hence, regardless of whether or not the schools were included in the initial pairing of schools).[13]

The non-random selection of schools, as well as the limited number of treated and possible control schools, makes it possible that the two groups will differ in terms of average student characteristics (although it is not obvious whether we should expect the intervention schools to be positively or negatively selected). Table 3 compares the background characteristics of students enrolled in intervention schools (col. 1) and all remaining schools (col. 2) in the targeted districts. The table shows that the two groups are very similar on average, although the share of foreign-born students is slightly higher in the intervention schools as are the parents' education level and earnings. However, these differences are quite marginal in comparison to the striking differences relative to the rest of the schools in the country (col. 3).

[12] One of the ten schools that initially were selected for participation closed down shortly afterwards, and a new school was instead offered participation. The new school was able to join the project already during the first semester and therefore received almost the same amount of treatment as the other intervention schools. The school that was shut down very early on is not included in the analysis.

[13] One reason for restricting the sample to schools within the targeted districts is that the urban development districts also have been the target of other initiatives that may affect educational outcomes; see further discussion in Section 4.

Table 3. Descriptive statistics for pupils who finished grade 9 at intervention, control and other schools 2009−2014

	Intervention schools	Control schools[a]	Other schools
Age	16.1	16.0	16.0
Girl	0.487	0.485	0.487
Born in another country	0.372	0.343	0.087
Immigrated 0- 2 years ago	0.092	0.092	0.014
Immigrated 3- 4 years ago	0.064	0.057	0.013
Immigrated 5- 6 years ago	0.054	0.048	0.013
Both parents are born in another country	0.758	0.783	0.160
Mother's level of education[b]			
Compulsory school or lower	0.363	0.380	0.102
Upper secondary education	0.333	0.350	0.472
Post-secondary education	0.186	0.149	0.392
Information is missing	0.118	0.121	0.034
Father's level of education[b]			
Compulsory school or lower	0.265	0.304	0.136
Upper secondary education	0.358	0.362	0.490
Post-secondary education	0.198	0.165	0.304
Information is missing	0.179	0.168	0.071
Mother's earnings (SEK)[b]	121,996	114,255	250,303
Father's earnings (SEK)[b]	166,728	154,243	355,175
Information about mother is missing	0.043	0.040	0.017
Information about father is missing	0.103	0.092	0.039
Number of individuals	4,810	4,274	623,697
Number of schools[c]	10	10	1,680

Note: All variables originate from Statistics Sweden, and are described in Section 4. [a]The control group consists of all public schools in the targeted city districts that did not participate in the intervention. [b]Parents' education and earnings are measured the year before the student finished 9th grade. Average earnings are computed among parents who were Swedish residents that year. [c]Number of "other schools" is measured in 2014.

4 Data and research design

Our analysis is based on individual level register data, which we have obtained from Statistics Sweden. To measure school performance, we use information on students' grades on the national standardized tests in Swedish, English and math, which are given during the spring semester of grade 9.[14] To account for changes in grading standards

[14] Pupils who do not have Swedish as their native language may take the course Swedish as a Second Language instead of regular Swedish, but they will still take the same standardized test as all other pupils.

over time, the students are percentile ranked within cohort based on their performance on each test.[15] We also use information on the student's (percentile ranked) GPA, and whether he/she was admitted to a regular program in upper secondary school directly after finishing 9th grade as outcome variables. The GPA has the advantage of capturing effects for a broader range of subjects than those covered by the standardized tests. The obvious drawback is that it is a less objective measure of student performance, and we cannot dismiss the possibility that the intervention would affect local grading standards (e.g. by making the teachers more generous in their grading). Nevertheless, it is the GPA that determines admission to further education in the Swedish education system.

Our database also contains information on which school the individual graduated from (the school's name as well as location), which makes it possible to determine if he/she attended a school that took part of the intervention. We also have access to background variables: age, sex, each parent's education level and earnings[16], and immigrant background. To account for immigrant background in the analysis, we control for whether or not the pupil is born in another country, whether or not both of his/her parents are born in another country, and whether he/she recently immigrated to Sweden (0-2 years ago, 3-4 years ago, 5-6 years ago, vs. earlier).[17] The full list of background variables, as well as descriptive statistics, is displayed in Table 3.

In addition, we have access to individual data for the teachers employed at each school, which we use to estimate effects on teacher turnover. Moreover, some analyzes are based on data from a questionnaire with pupils. We describe the student questionnaire in Section 5.4.2.

To estimate the effect of the CFT intervention we compare how student performance develops for the schools that took part in the program, to the development for the untreated schools in the targeted districts, using a difference-in-differences design. Since the intervention was targeted at public schools, only public schools are included in the analysis. The rational for limiting the control group to untreated schools within the same city districts is twofold: i) The credibility of the analysis hinges on the comparability of the two groups of schools. City districts that are considered urban

[15] The students are ranked among all students in the country who took the test the same year.
[16] Parents' education level and earnings are measured the year before the student finished 9th grade.
[17] Immigration date is measured as the date the person was granted residence permit. Some individuals have more than one immigration date registered. In such cases, we base our measure of immigration year on the earliest date registered.

development districts all have low employment rates, high reliance on social assistance, poor schooling outcomes, and, as we saw in Table 3, the two groups of schools have very similar student characteristics. ii) The urban development districts have also been the target of other initiatives that may affect educational outcomes (see Section 2.2). Thus, expanding the control group to also include schools in other districts would make it harder to determine whether any observed differences in development were due to the CFT intervention or these other initiatives.

However, choosing a control group of schools from the same districts, and that operates under very similar conditions, also comes with a potential disadvantage: It is possible that the municipalities (the responsible administrative body of the public schools) would choose to compensate the schools that were not selected for participation, e.g. with additional resources, which may raise performance in the control group. Such a response, on the part of the municipalities, could lead us to underestimate the impact of the CFT intervention on student performance. The same problem could occur if it is common that teachers who have received additional training through the intervention start working at untreated schools within the same city districts.

Our analysis is based on data for the time period 2009- 2014. The start date is chosen to keep the panel of schools as balanced as possible (some of the schools in the sample are relatively new).[18] The rational for ending the evaluation period already in 2014 – only two years after the start of the intervention – is that the government thereafter introduced an additional policy change that affected some of the schools in the sample.[19] Hence, after 2014 it is difficult to separate the effect of the CFT program from the effect of other policy changes. The obvious drawback with such a short follow-up period is that the cohort that was most affected by the intervention (those finishing 9th grade in 2015), and potentially benefited the most, will not be included in the analysis.

We estimate the following regression model:

$$y_{ist} = \alpha_0 + \beta_1(CFT_s * D_{2013}) + \beta_2(CFT_s * D_{2014}) + \sum_{t=2009}^{2014} \theta_t D_t + \boldsymbol{\delta X}_i + \gamma_s + e_{ist} \qquad (1)$$

[18] All schools but one can be observed from the beginning of this time period. For the newest school we lack data for the first two years. After the school year 2012/2013 one of the schools in the control group was closed down. In Section 5.2, we show that our results are not affected by whether this school is included in the analysis or not.

[19] In March 2014 the government decided to earmark central government grants for hiring additional teachers at some disadvantaged schools; see Government Regulation (2014:145). Pupils that graduated in the spring of 2014 (or earlier) should be unaffected by this policy change.

where *i* indexes individual, *t* the year the individual finished 9th grade, and *s* the school he/she attended. y_{ist} is the individual's (percentile ranked) grade on the standardized test in math, Swedish or English; GPA; or an indicator for admittance to upper secondary school. CFT_s is an indicator that takes the value one if the school the student attended participated in the intervention; otherwise it is zero. D_t are dummies for graduation year, $\boldsymbol{X}_i$ is a vector of individual and parental background characteristics (those listed in Table 2), and γ_s represent school fixed effects. e_{ist} is the error term.

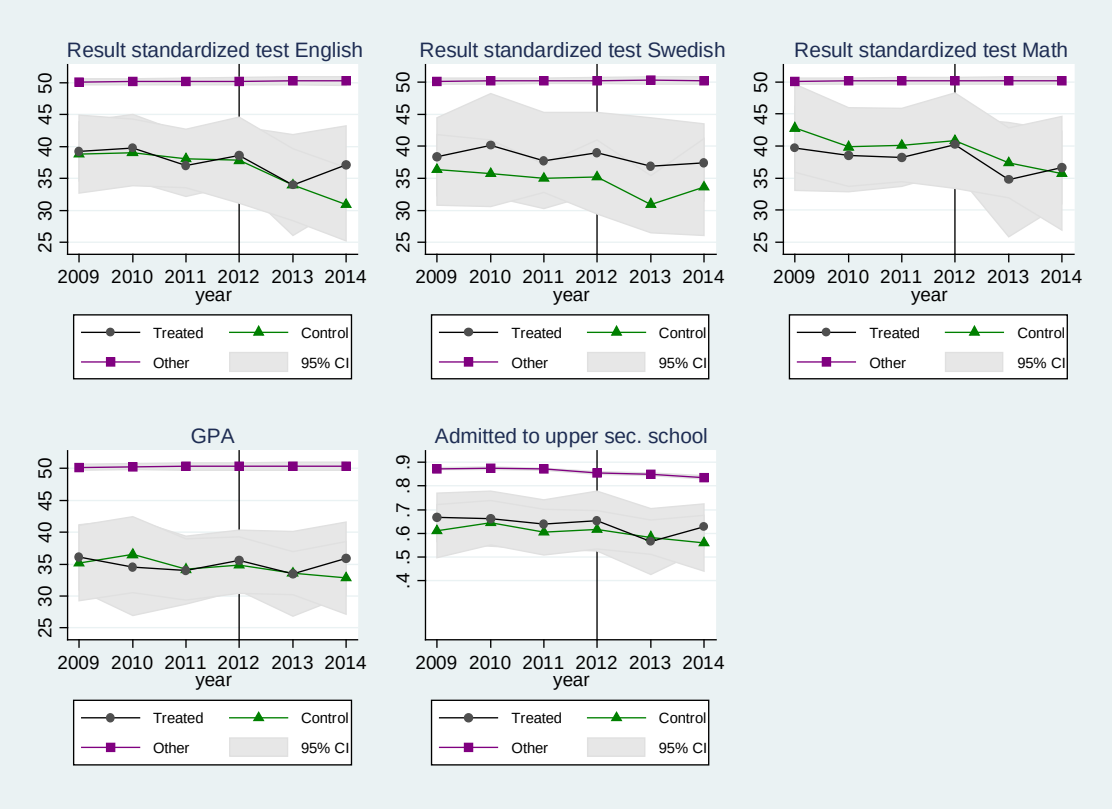
The parameters of interest, β_1 and β_2, are the difference-in-differences estimates of the effect of the intervention for the two cohorts of pupils that were treated – those who finished grade 9 in 2013 and those who finished in 2014. The first of these cohorts began 9th grade when the intervention was launched in the fall of 2012. However, the activities at the schools generally did not start until later during that semester, which means that they were just in place for about one semester before these students took the standardized tests (Assadi et al. 2015). Since the activities were directed towards the teachers, and not directly towards the students, this is probably a too short time period for there to be any noticeable effects on student performance.[20] It is more likely that the intervention would affect the performance of the second cohort of students, those who finished in 2014. The activities at the schools were in place for about three semesters before this cohort of students took the standardized tests. Hence, if the CTF intervention was effective in raising student performance, we expect $\beta_2 > \beta_1$.

By incorporating school fixed effects, our model accounts for all unobserved differences across schools that remained constant over time. But to interpret β_1 and β_2 causally we need to assume that trends in student performance would not have differed between intervention and control schools in the absence of the CFT intervention. Figure 1, which plots average student outcomes for intervention and control schools (as well as all other schools) during our study period, indicates that this may very well be a realistic assumption in this case. For none of the outcomes we study are there any clear indications of divergent trends for treated and control schools during the period before

[20] The teachers must first learn lessons that can be used in their daily work in the classroom. They must thereafter change their method of teaching to make use of these lessons, which hopefully will enhance learning among students.

the intervention started (2009- 2012).[21] In Section 5.2 we test this more formally by performing a set of placebo-tests.

Figure 1.Results on standardized tests, GPA, and admission to upper secondary school for intervention schools, control schools and all other schools, 2009−2014



Note: Results (grades) on the standardized tests and GPA are percentile ranked within cohort. The standard errors allow for clustering by school. The vertical line marks the year before the CFT intervention started.

When it comes to impacts of the program, Figure 1 suggests that performance in English language may have improved among students graduating in 2014 (although the difference between the two groups of schools is not statistically significant in this simple year to year comparison). But we see no signs of improvements in math or Swedish. There is also a tendency for students to obtain higher GPAs following the intervention, but no clear indication of an increase in admission to upper secondary education

[21] That intervention and control schools have similar trends in the period before the intervention is perhaps not surprising given that the selection process tried to ensure that treated and control schools would be balanced in terms pre-treatment trends in student performance. However, since the number of schools in the targeted districts was rather small, it would still have been possible to end up with rather unbalanced groups. Moreover, the control group used in the analysis diverges somewhat from the initial set of control schools.

5 Results

This section presents the results from the empirical analyses. We begin by showing results for the full sample of students (Section 5.1). This presentation is followed by some robustness checks (Section 5.2), and thereafter we show results for different subgroups (Section 5.3). Last, we exploit pupil questionnaire data as well as register data on teachers to try to shed light on the mechanisms behind our findings (Section 5.4).

5.1 Main results

Table 4 displays results for all five outcome variables for the full sample of students. We show results both with and without individual background controls included in the model. To allow for correlation in error terms of individuals attending the same school, standard errors should be clustered at the school level. However, since we rely on rather few clusters (only 20 schools), cluster-robust standard errors may be underestimated. To make correct inference with few clusters, we instead use wild bootstrap as suggested by Cameron et al. (2008) and Cameron and Miller (2015). Table 4 (as well as all following tables of results) displays the p-values resulting from wild bootstrap.

Table 4. Effects of the CFT intervention on results on standardized tests, GPA, and admission to upper secondary school (p-values in parentheses)

A. Test results	English		Swedish		Math	
	(1)	(2)	(3)	(4)	(5)	(6)
Effect 2013	0.453	1.512	3.288	3.548	1.054	1.597
	(0.873)	(0.590)	(0.280)	(0.237)	(0.818)	(0.712)
Effect 2014	6.640*	6.393**	2.218	1.170	5.615	5.538
	(0.052)	(0.040)	(0.610)	(0.784)	(0.326)	(0.326)
Controls	no	yes	no	yes	no	yes
Observations	7,140	7,139	7,296	7,295	7,087	7,086
R-squared	0.073	0.214	0.072	0.226	0.078	0.169
Mean of outcome variable	37.182	37.184	36.522	36.526	38.785	38.790

B. GPA and admission to upper secondary school	GPA		Admitted	
	(7)	(8)	(9)	(10)
Effect 2013	1.084	2.011	-0.036	-0.017
	(0.672)	(0.428)	(0.453)	(0.721)
Effect 2014	5.109*	4.246	0.067	0.059
	(0.058)	(0.118)	(0.312)	(0.350)
Controls	no	yes	no	yes
Observations	9,084	9,082	9,084	9,082
R-squared	0.056	0.273	0.083	0.290
Mean of outcome variable	34.753	34.759	0.621	0.621

Notes: Results (grades) on standardized tests and GPA are percentile ranked within cohort. All regressions control for year and school fixed effects. Col. (2), (4), (6), (8) and (10) additionally control for sex, age, born in another country, years since immigration (4 categories), parents born in another country, mother's and father's education level (4 categories), and mother's and father's (log) earnings (see Table 3 for details). P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

As expected, given the short time of exposure, there is no indication that the CFT intervention affected educational outcomes among students finishing 9th grade in 2013. Among those who graduated in 2014, we see (in line with the pattern in Figure 1) improved results on the standardized test in English among students attending intervention schools. This effect is statistically significant at the 5 percent level, and the estimate is robust to the inclusion of background controls in the model. In terms of magnitude, the 6.4 percentile points increase (col. 2) corresponds to 17 percent of the sample mean and close 50 percent of the gap in test results between students in urban development districts and other schools in the country (see Figure 1). This must be considered a substantial improvement. The point estimates are positive and large also for some of the other outcomes variables, but they are all imprecisely estimated and for the most part not significantly different from zero at conventional levels of significance.

5.2 Robustness checks

As discussed above, a causal interpretation of our estimates relies on the assumption that trends in student performance would not have differed between intervention and control schools in the absence of the intervention. The patterns in Figure 1 gave us no reason to doubt the validity of this assumption. However, in order to more formally assess this, we estimate a number of placebo-regressions: We estimate the same model, but (artificially) set the intervention year to one, two and three years before the actual start year.[22] Any significant placebo-estimates for these years would cast doubts on the validity of the identifying assumption as it would indicate divergent trends already before the intervention was launched. Reassuringly, we find no statistically significant placebo-estimates, and the estimates are, for the most part, much smaller in magnitude compared to those displayed in Table 4 for 2014.

Table 5. Placebo-tests (p-values in parentheses)

	(1) English	(2) Swedish	(3) Math	(3) GPA	(4) Admitted
Placebo estimate 2012	1.790	0.687	2.746	2.004	0.029
	(0.455)	(0.860)	(0.298)	(0.181)	(0.364)
Observations	4,874	5,000	4,863	6,224	6,224
R-squared	0.195	0.218	0.167	0.282	(0.308)
Placebo estimate 2011	-1.233	-1.150	0.736	0.863	0.001
	(0.467)	(0.798)	(0.745)	(0.670)	(0.984)
Observations	3,657	3,738	3,671	4,732	4,372
R-squared	0.198	0.217	0.175	0.293	0.318
Placebo estimate 2010	1.143	4.372	2.466	-0.884	-0.006
	(0.480)	(0.233)	(0.496)	(0.685)	(0.888)
Observations	2,539	2,629	2,591	3,240	3,240
R-squared	0.206	0.239	0.188	0.309	0.336

Notes: Results (grades) on standardized tests and GPA are percentile ranked within cohort. All regressions control for year and school fixed effects as well as sex, age, born in another country, years since immigration (4 categories), parents born in another country, mother's and father's education level (4 categories), and mother's and father's (log) earnings (see Table 3 for details). P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

The fact that our sample consists of only 20 schools means that the development of an individual school potentially can have a large influence on the results. This raises

[22] In each regression the data is censored after the placebo intervention year. That is, the placebo-regression for year 2012 is based on data for 2009-2012; the placebo-regression for year 2011 is based on data for 2009-2011; and the placebo-regression for 2010 is based on data for 2009-2010.

concerns that the estimated effect on English test scores would be largely driven by improvements in a particular school. To investigate the sensitivity of our results in this regard, we re-estimate all regressions presented in Table 4 (with background controls included) excluding one school at a time. The estimate for English is stable across these 20 regressions; see Table A 1 in the Appendix. The patterns for the other outcomes are, for the most part, also unaffected.

As mentioned in Section 4, one of the schools in the control group was closed down in 2013. In Table A 1 we can see that all estimates stay very similar if this school is excluded from the analysis; see Panel B, col. 7.

5.3 Heterogenous effects?

Some of the intervention's activities explicitly focused on enhancing the teaching of immigrant children with poor ability in the Swedish language. It is therefore of special interest to examine how the intervention impacted students that relatively recently immigrated to Sweden. Table 6 shows results from separate regressions for students that immigrated 0- 4 years ago[23] and all other students.

[23] Immigration date is defined as the date the person was granted residence permit.

Table 6. Effects of the CFT intervention for newly arrived immigrants and others (p-values in parentheses)

A. Test results	English		Swedish		Math	
	New immigrant	Others	New immigrant	Others	New immigrant	Others
Effect 2013	1.741	1.127	-0.567	3.446	-0.567	1.747
	(0.761)	(0.708)	(0.960)	(0.231)	(0.960)	(0.697)
Effect 2014	6.670	6.032*	-2.954	1.338	-2.954	6.808
	(0.265)	(0.078)	(0.680)	(0.792)	(0.680)	(0.268)
Controls	yes	yes	yes	yes	yes	yes
Observations	522	6,617	684	6,746	684	6,402
R-squared	0.209	0.189	0.193	0.204	0.193	0.172
Mean of outcome var.	18.501	38.658	18.168	38.020	31.148	39.606

B. GPA and admission upper sec. school	GPA		Admitted	
	New immigrant	Others	New immigrant	Others
Effect 2013	-1.122	2.173	-0.009	-0.025
	(0.758)	(0.418)	(0.881)	(0.649)
Effect 2014	-4.725	6.162**	0.015	0.068
	(0.302)	(0.039)	(0.852)	(0.308)
Controls	yes	yes	yes	yes
Observations	1,388	7,694	1,388	7,694
R-squared	0.231	0.199	0.216	0.163
Mean of outcome var.	12.610	38.755	0.170	0.702

Notes: Results (grades) on standardized tests and GPA are percentile ranked within cohort. "Newly arrived" is defined as children that immigrated 0-4 years ago. All regressions control for year and school fixed effects as well as sex, age, born in another country, parents born in another country, mother's and father's education level (4 categories), and mother's and father's (log) earnings (see Table 3 for details). P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

The estimated effect on test results in English, for the 2014 graduation cohort, is similar in size for newly arrived immigrant children and others (although it is less precisely estimated for the former group); for all other outcomes the estimates are smaller in size for the newly arrived and sometimes negative. Thus, we find nothing to suggest that the CFT intervention primarily benefitted pupils with poor ability in the Swedish language; if anything, the pattern is the reverse. Among pupils that have resided in Sweden longer than four years, we see a positive and statistically significant effect of the intervention on the GPA, indicating that the intervention potentially benefitted students' performance in more subjects than those covered by the standardized tests.

To further investigate whether the intervention seems to primarily have benefitted relatively weak or strong pupils, we estimate effects on the probability of achieving a

GPA above the 25th, 50th, and 75th percentile of the (sample) grade distribution.[24] The results from these regressions are presented in Table 7. The results are in line with those in Table 6 as they suggest that the intervention primarily benefited students in the middle and upper part of the grade distribution. We find a positive and statistically significant effect on the probability of receiving a GPA above the median as well as above the 75th percentile, while the estimate for the 25th percentile is much smaller and statistically insignificant. However, it is important to point out that pupils in the middle-upper part of the grade distribution for these schools may still be relatively weak in relation to the grade distribution for all students in the country.

Table 7. Effects on the probability of receiving a GPA above the 25th, 50th, and 75th percentile in the GPA distribution (p-values in parentheses)

	GPA> 25th percentile	GPA> median	GPA> 75th percentile
Effect 2013	0.011	0.048	0.037
	(0.814)	(0.206)	(0.178)
Effect 2014	0.022	0.079*	0.059*
	(0.467)	(0.076)	(0.062)
Controls	yes	yes	yes
Observations	9,082	9,082	9,082
R-squared	0.293	0.216	0.149
Mean of outcome variable	0.752	0.500	0.248

Notes: GPA is percentile ranked within cohort. All regressions control for year and school fixed effects as well as sex, age, born in another country, years since immigration (4 categories), parents born in another country, mother's and father's education level (4 categories), and mother's and father's (log) earnings (see Table 2 for details). P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

Last, we have analyzed whether the impact of the intervention differs for boys and girls; see Table A 2 in the appendix. The positive effect on students' results in English is clearer among girls; other than that, we find no indications of gender differences in impact.

5.4 Mechanisms

5.4.1 Teacher turnover

As discussed in the introduction, a major challenge for schools located in disadvantaged city districts lies in attracting and retaining effective teachers. A possible explanation for the positive impact on student performance is that the schools, through the program, became more attractive workplaces and thereby, to a higher extent, managed to keep

[24] We focus on GPA in this analysis as this is our only continuous outcome variables.

their teaching staff. To examine if reduced teacher turnover could be a potential mechanism behind our findings, we use the Teacher register (*Lärarregistret*) and construct a dataset consisting of all teachers employed at intervention and comparison schools from the fall semester of 2008 to the fall semester of 2012.[25] We then estimate the same type of difference-in-differences model as before (with school fixed effects), but where the outcome is the probability of remaining employed at the same school the following school year. Table 8 displays the results from this analysis.

The first group of teachers that potentially could be affected by the intervention are those employed during the school year 2011/2012, as the selection of schools for participation was made during the spring of 2012. However, awareness that the school would be subject to the CFT program the following school year does not seem to have affected teachers' decision to remain at the schools. The estimate (-0.011) is very small in size and statistically insignificant. Those employed during the school year 2012/2013 were directly affected by the intervention, but neither for this year do we find evidence of an effect on teacher turnover. The estimate is negative (hence indicating a higher rather than a lower turnover rate), but statistically insignificant. In sum, we find nothing to suggest that changes in teacher mobility would be the explanation behind the improved school results.

Table 8. Effect of the CTF intervention on the probability of remaining employed at the same school the following school year (p-values in parentheses)

	Remaining employed t+1
Effect for teachers employed during the fall of 2011	-0.011 (0.863)
Effect for teachers employed during the fall of 2012	-0.079 (0.189)
Observations	3,429
R-squared	0.038
Mean of outcome variable	0.745

Notes: The regression controls for year and school fixed effects. P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

[25] All individuals who were employed as teachers for the relevant age groups are included. That is, we exclude school administrators, student counselors, staff for extracurricular activities, and teachers that are registered as primary school or preschool teachers. Both full and part-time teachers are included.

5.4.2 Students' perceptions of the learning environment

To learn more about how the CFT program affected the instruction and classroom climate, we make use of a survey of pupils in intervention schools as well as in a group of ten similar non-treated schools. The control group (which is not identical to the control group used in the previous analyses) was constructed to match the intervention schools in terms of location (municipality) and pre-intervention trends in GPA and eligibility for upper secondary education.[26] The survey was taken both before the intervention started (during the fall of 2012) and again two years later. Although all ten pairs of schools answered the first round of the survey, only seven pairs can be included in the analysis: One of the schools that initially was selected for participation closed down just before the intervention started, but after the first round of the survey. The Swedish National Agency for Education then decided to include one of the comparison schools in the project. Furthermore, one of the remaining comparison schools refused to participate in the second round of the survey.

We approached all pupils in grade 7–9, and we received a response rate of 75 percent in the first survey round and 82 percent in the second. The average response rates were similar for intervention and comparison schools. To examine if the program affected students' responses, we again use a difference-in-differences strategy. Admittedly, there is some uncertainty as to whether the estimates can be given a causal interpretation in this case. One obvious deficit is that it is not possible to examine pre-intervention trends in the relevant outcomes as the questionnaire was only taken once before the intervention started. Somewhat reassuringly, the two groups of schools are comparable in terms of average student characteristics in the beginning of the intervention (see Table A 3).[27]

The main idea of the program was to strengthen the teachers' professional skills and confidence through coaching, collegial learning and further training. As a consequence, the instruction as well as the classroom climate was expected to improve, which should enhance learning and performance among the students. In line with this chain of events, we anticipate positive effects of the intervention on students' opinions about their teachers' ability to teach and the learning environment in the classroom. The survey

[26] The treatment and control group here consist of the initial pairs of schools described in Section 3.1.
[27] As discussed in Section 3.1 and 4, it may also be problematic that a couple of the schools in this control group were located outside the boundaries of the targeted districts.

therefore included a number of questions on students' perception of how their teachers teach and the classroom climate.[28]

We use factor analysis of six survey items to empirically examine whether it is possible to identify a common underlying dimension of opinions about *teacher ability* in the responses (see Table A 4). For instance, we asked the students to what extent they think their teachers are good at teaching and creating an interest among the students. All six items load positively and high on a single dimension. These findings indicate that an index based on the items is likely to be a valid measure of students' perceptions of teacher ability. We therefore use the factor scores from the analysis to compute an index standardized to have a mean of 0 and a standard deviation of 1. A higher score implies a more positive assessment of teacher ability. We also use the overall question: "How many of your teachers are good at teaching?" as an additional alternative outcome.

We employ a similar procedure to construct an index of opinions about the *learning environment* in the classroom. Three survey items capture the extent to which pupils find that the classroom climate is favorable for learning. Once again, the items load positively and high on a single dimension (see Table A 5). We construct an index (mean 0, standard deviation 1) where a higher score implies a more positive view of the classroom climate.

[28] Note that students' perceptions of the teachers' ability, expressed in a survey, is not necessarily the same as the teachers' actual ability. However, it seems more likely that the intervention really affected learning if we can see an impact of the program on students' perceptions of the teachers and the learning environment in the classroom.

Table 9 shows the results when we use students' opinions of teacher ability as the dependent variable. In columns 1–3, the full index of six items is the outcome, whereas the single survey question is the outcome in columns 4–6. We estimate models without control variables (col. 1 and 4) and models controlling for background characteristics (col. 2 and 5). In the most restrictive specification, we also include school fixed effects (col. 3 and 6).

Table 9. Effects of the CFT intervention on students' perceptions of teacher ability (p-values within parentheses)

	Index of six survey items			One survey item		
	(1)	(2)	(3)	(4)	(5)	(6)
Effect of CFT	0.199**	0.170**	0.139	0.172***	0.210***	0.174***
	(0.036)	(0.026)	(0.109)	(0.010)	(0.000)	(0.002)
Controls	no	yes	yes	no	yes	yes
School fixed effects	no	no	yes	no	no	yes
Observations	4,204	3,599	3,599	5,069	4,256	4,256
R-squared, adj.	0.005	0.098	0.124	0.003	0.054	0.087
Mean of outcome var.	0.000	0.000	0.000	3.769	3.769	3.769

Note: In models 1–3, an index of perceived teacher ability based on six survey items from the student survey is used as the dependent variable. The index is computed using the factor loadings from a factor analysis of the questions. The index is set to have a mean of 0 and a standard deviation of 1, and a higher score implies a more positive assessment of teacher ability (see Table A 3 for details). In models 4–6, the following survey question is used as the dependent variable: "How many of your teachers are good at teaching?" The response scale is 1–5, where 1 = "none of them" and 5 = "all of them". Models (2), (3), (5) and (6) include the following control variables: sex, age, Swedish as native language (yes/no), have attended Swedish school for less than four years (yes/no), attends a class for newly arrived immigrants (yes/no), father's educational level (5 levels), mother's educational level (5 levels). P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

We find positive estimates in all specifications, indicating that the intervention improved students' perceptions about their teachers' ability to teach. The estimates suggest an increase by around 0.15 of a standard deviation when the index is used as the outcome, and an increase of about 0.2 on a scale 1–5 when the single survey item is used as outcome. The estimate becomes marginally insignificant (p-value = 0.109) in the most restrictive specification when we use the index as the outcome (col. 3). However, we can only include pupils who have answered all six survey questions, and where we have information on all control variables, in this analysis. There are quite a lot of internal missing values, which decreases the sample size and therefore statistical precision.[29]

Table 10 displays the results when we study student perceptions of the classroom climate. The positive estimates suggest that the situation in the classroom became more favorable for learning due to the intervention. The estimate is marginally insignificant (p-value = .106) in the model without control variables, but statistically significant when covariates and/or school fixed effects are added to the model.[30] In sum, the

[29] As a robustness check, we have estimated models with each survey question as the outcome separately (see Table A6). The results show positive estimates between 0.09 and 0.21. In 11 of 18 specifications, the estimate is statistically significant at least at the 10 percent level.

[30] Table A6 shows that the results are mainly driven by students' opinions regarding the extent to which the classroom can be described as a place where "it's peace and quiet, and it's possible to concentrate on school work".

analysis of survey responses suggests that the CFT intervention may have improved the teaching as well as the classroom climate at the schools, in line with the stated intentions. These developments may thus account for the improved school results at the intervention schools.

Table 10. Effects of the CFT intervention on student's perceptions of the learning environment in the classroom (p-values within parentheses)

	(1)	(2)	(3)
Effect of CFT	0.078	0.133***	0.093*
	(0.106)	(0.006)	(0.056)
Controls	no	yes	yes
School fixed effects	no	no	yes
Observations	4,739	4,066	4,066
R-squared, adj.	0.002	0.011	0.058
Mean of outcome variable	0.000	0.000	0.000

Notes: The dependent variable is an index computed from the factor loadings from a factor analysis of three questions in the student survey. This index is set to have a mean of 0 and a standard deviation of 1, and a higher score implies a more positive assessment of the learning environment in the classroom (see Table A 4 for details). Models (2) and (3) include the following control variables: sex, age, Swedish as native language (yes/no), have attended Swedish school for less than four years (yes/no), attends a class for newly arrived immigrants (yes/no), father's educational level (5 levels), mother's educational level (5 levels). Models (3) also include school fixed effects. P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

As outlined in Section 3, the CFT program included some additional components. Most notably, schools could apply for extra funding to enhance teaching outside of regular school hours (e.g. homework assistance) and for tutors that could assist immigrant students in their native language. Table 11 shows estimates from regressions where we use responses to six survey questions related to these aspects as outcomes.[31]

[31] Note that the number of observations varies among the columns. This is no surprise: First, "don't know"-answers are excluded from the analysis in columns 2, 3, and 6. Second, the questions in columns 3 and 6 were only answered by students receiving homework assistance/help with school work in native language. Third, extreme values (answers over 15 hours per week) are excluded in column 3. Fourth, there are some internal missing values on these questions.

Table 11. Effects of the CFT intervention on students' perceptions of homework assistance, teaching outside regular school hours and help with school work in native language (p-values within parentheses)

	Receives homework assistance?	Minutes per week of homework assistance?	Pleased with homework assistance?	Receives teaching outside regular school hours?	Receives help with school work in native language?	Pleased with help with school work in native language?
	(1)	(2)	(3)	(4)	(5)	(6)
Effect of CFT	-0.102 (0.170)	-15.944 (0.266)	0.005 (0.934)	0.015 (0.780)	0.071** (0.040)	0.138 (0.232)
Controls	yes	yes	yes	yes	yes	yes
School FE	yes	yes	yes	yes	yes	yes
Observations	4,205	3,513	1,692	4,132	3,057	613
R-squared, adj.	0.126	0.215	0.037	0.048	0.080	0.019
Mean of outcome var.	0.417	47.819	2.939	0.135	0.236	3.312

Notes: The outcomes in columns (1), (4) and (5) are dichotomous (0/1=no/yes). The outcomes in columns (3) and (6) are measured on an attitude scale 1–4, where 4 = "very pleased". All models include school fixed effects and the following control variables: sex, age, Swedish as native language (yes/no), have attended Swedish school for less than four years (yes/no), attends a class for newly arrived immigrants (yes/no), father's educational level (5 levels), mother's educational level (5 levels). Findings are robust if these variables are excluded from the models, with the exception that the estimate in column (1) becomes slightly larger (-0.140) and statistically significant (p-value=0.048) if no controls are added. P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

We find no indication that the intervention affected homework assistance and other forms of teaching outside of regular school hours (col. 1-4). But we find a rather large positive estimate when we study the presence of tutors that could assist students in their native language. The probability that students born abroad answered that they receive help in their native language increased by 7 percentage points (col. 5). This seems like a quite sizeable effect (corresponding to around 30 percent of the mean). However, the fact that positive impacts on student performance are primarily found among students who have resided in Sweden longer than four years (see Table 6), suggests that improved access to this type of assistance is unlikely to account for the overall improvement in school results at the intervention schools.[32]

[32] The survey also included several questions on teaching methods. For example, how often students have homework and written tests, how often they participate in group work, and how often they work individually. We find no indication that the CFT intervention affected teaching in such dimensions.

6 Conclusion

Children growing up in neighborhoods with high levels of social exclusion in the form of unemployment, poverty and reliance on social assistance tend to perform significantly worse in school compared to children growing up under more favorable conditions. In this paper we have examined a public program, "Coaching for Teaching", initiated by the Swedish government to improve school results in ten disadvantaged lower secondary schools. The program consisted of several components, but primarily focused on teachers' professional development. We have used rich register data and a difference-in-differences design to estimate the impact of the program on several educational outcomes.

We find a sizeable and statistically significant positive effect on student performance on standardized tests in English language. The estimates for students' test scores in math, GPA, and admission to upper secondary school are also positive and large, but not statistically significant at conventional levels. The low number of schools participating in the intervention may explain why the effects are not precisely estimated. Hence, we cannot rule out that there were positive impacts also in these regards. For test scores in Swedish language there is no indication of improvement.

We have also investigated whether the program primarily benefitted relatively weak or strong pupils. We find some indication of stronger effects for children who have lived in Sweden for more than four years. For this group we also see a positive and statistically significant impact on the GPA. The effects are also more accentuated for students in the middle and upper part of the grade distribution in the sample. However, it is essential to emphasize that these pupils may still be relatively weak in relation to the grade distribution for all students in the country.

Furthermore, we have analyzed data from a student questionnaire to better understand the implications of the program. The results indicate that students' perceptions of the learning environment in classroom and their teachers' ability to teach improved. Hence, the program may, in line with its intentions, have improved the learning environment at the schools. Taken together, we interpret the empirical results as quite promising in the short run.

Like many other programs designed to fight social exclusion, the CFT program consisted of several components. This implies that it is hard to determine which parts of

the program that were more and less successful. Coaching of teachers in small groups constituted the main part of the program. However, qualitative results in Assadi et al. (2015) warn against adopting the specific coaching model used within the project. According to that study, one-to-one coaching and further training were more promising parts of the program. Thus, we think that it is appropriate to interpret our results as indications that strategies to strengthen teachers' professional development – without pinpointing the exact design of such a policy – can be a fruitful way to enhance student performance in disadvantaged areas. It is also reasonable to assume that such strategies will be more successful if teachers themselves perceive them as useful.

We have only been able to examine short run effects. However, an important idea behind the program was to generate a more lasting impact compared to temporary increases in resources. For instance, while the impact of additional teachers is likely to disappear when resources are removed, increased professional competence among existing teachers might produce positive effects for subsequent cohorts even when the program is over. This idea builds on at least two important assumptions: First, that the teachers will remain employed at disadvantaged schools, or that improved teaching strategies can be transferred to new recruits. Second, the effects cannot simply be driven by enthusiasm over the fact that the schools were selected to participate in a project. We leave to future research to explore these issues.

Our study contributes to research on how to improve schooling outcomes in disadvantaged city districts, and among pupils with disadvantaged background more generally (e.g. Angrist 2013; Chiang et al. 2017; Cullen et al. 2013; Curto, Fryer and Howard 2011). Previous research has shown that teacher quality is a crucial factor for enhancing student performance (e.g. Rockoff 2004; Rivkin, Hanushek and Kain 2005; Aaronson, Barrow and Sander 2007), but also that it is difficult for schools in disadvantaged neighborhoods to attract and retain effective teachers (Sass et al. 2012; Chiang, Clark and McConnell 2017). We have added to the existent literature by presenting evidence suggesting that programs aiming to strengthen teachers' professional development might be a promising way forward for these schools.

References

Aaronson, Daniel, Lisa Barrow, and William Sander (2007), "Teachers and Student Achievement in the Chicago Public High Schools", *Journal of Labor Economics* 25(1): 95–135.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters (2013), "Explaining Charter School Effectiveness", *American Economic Journal: Applied Economics* 2013, 5(4): 1–27.

Assadi, Anahita, Caroline Hall, Martin Lundin, and Kristina Sibbmark (2015), "Erfarenheter och effekter av satsningar på lärare i skolor med låga elevresultat", Report 2015:23, IFAU, Uppsala.

Bloom, Howard S., Saskia Levy Thompson, Rebecca Unterman, Corinne Herlihy, and Collin F. Payne (2010), "Transforming the high school experience: How New York City's new small schools are boosting student achievement and graduation rates", Report, Manpower Demonstration Research Corporation, New York.

Böhlmark, Anders, Helena Holmlund, and Mikael Lindahl (2015), "School choice and segregation: evidence from Sweden", Working Paper 2015:8, IFAU, Uppsala.

Cameron, Colin A., Jonah. B. Gelbach, and Douglas. L. Miller (2008), "Bootstrap-based improvements for inference with clustered errors", *Review of Economics and Statistics* 90(3):414–427.

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

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan (2011), "How does your kindergarten classroom affect your earnings? Evidence from Project STAR", *Quarterly Journal of Economics* 126(4):1593–1660.

Chiang, Hanley S., Melissa A. Clark, and Sheena McConnell (2017), "Supplying Disadvantaged Schools with Effective Teachers: Experimental Evidence on Secondary Math Teachers from Teach for America", *Journal of Policy Analysis and Management* 36(1): 97–125.

Cook P, Dodge K, Farkas G, Roland G. Fryer J, Guryan J, Ludwig J, Mayer S, Pollack H, Steinberg L. (2015), "Not Too Late: Improving Academic Outcomes for Disadvantaged Youth", WP-15-01, Institute for Policy Research, Northwestern University.

Cullen, Julie B., Steven D. Levitt, Erin Robertson, and Sally Sadoff (2013), "What Can Be Done To Improve Struggling High Schools?", *Journal of Economic Perspectives* 27(2): 133–152.

Curto, Vilsa E., Roland G. Fryer Jr., and Meghan L. Howard (2011), "It May Not Take a Village: Increasing Achievement among the Poor", in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances* (eds. Greg J. Duncan and Richard J. Murnane), Russel Sage Foundation.

Curto, Vilsa E. and Roland G. Fryer Jr. (2014), "The Potential of Urban Boarding Schools for the Poor: Evidence from SEED", *Journal of Labor Economics* 32(1): 65–93.

Fryer, Roland G., (2013), "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools", *Journal of Labor Economics* 31(2): 373–407.

Fryer, Roland G., (2014), "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments", *Quarterly Journal of Economics* 129(3): 1355–1407.

Fryer, Roland G., Steven D. Levitt, John List, and Sally Sadoff (2012), "Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment", NBER Working Paper 18237.

Glazerman, Steven, Daniel Mayer, and Paul Decker (2006), "Alternative Routes to Teaching: The Impacts of Teach for America on Student Achievement and Other Outcomes", *Journal of Policy Analysis and Management* 25(1): 75–96.

Glewwe, Paul, Nauman Ilias, and Michael Kremer (2010), "Teacher incentives", *American Economic Journal* 2(3):205–27.

Goodman, Sarena, and Lesley Turner (2010), "Teacher incentive pay and educational outcomes: Evidence from the NYC bonus program", Working Paper No. 10–07, Program on Education Policy and Governance, Cambridge, MA.

Government Bill 2012/13:1, *Budgetpropositionen för 2013*

Government decision A2012/174/IU, *Kommuner som ska omfattas av urbant utvecklingsarbete från 2012.*

Government Regulation (2014:145), Förordning om extra statsbidrag till skolhuvudmän som inrättar karriärsteg för lärare i förskoleklasser och grundskolor i utanförskapsområden.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin (2004), "Why Public Schools Lose Teachers", *Journal of Human Resources* 39(2): 326–354.

Jacob, Brian and Lars Lefgren (2004), "Remedial education and student achievement: A regression-discontinuity analysis", *Review of Economics and Statistics* 86(1): 226–44.

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

Krueger, Alan B., and Diane M. Whitmore (2001), "The effect of attending a small class in the early grades on college–test taking and middle school test results: Evidence from Project STAR", *Economic Journal* 111, No. 468: 1–28.

Lankford, Hamilton, Susanna Loeb, and James Wyckoff (2002), "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis", *Educational Evaluation and Policy Analysis* 24(1): 37–62.

Muralidharan, Karthik, and Venkatesh Sundararaman (2011), "Teacher performance pay: Experimental Evidence from India", *Journal of Political Economy* 119(1): 39–77.

National Agency for Education (2015), "Uppdrag att stödja grundskolor i utanförskaps-områden. Redovisning av regeringsuppdrag", Skolverket, Stockholm.

National Agency for Education (2018), "Skolor och elever i grundskolan läsåret 2016/17", The Internet homepage of the National Agency for Education (*Skolverket*), *https://www.skolverket.se/statistik-och-utvardering/statistik-i-tabeller/grundskola/skolor-och-elever/skolor-och-elever-i-grundskolan-lasaret-2017-18-1.267701* (accessed 2018-05-15).

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain (2005), “Teachers, Schools, and Academic Achievement”, *Econometrica* 73(2): 417–458.

Rockoff, Jonah E. (2004), “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data”, *American Economic Review* 94(2): 247–252

Sass, Tim R., Jane Hannaway, Zeyu Xu, David N. Figlio, and Li Feng (2012), “Value Added of Teachers in High-Poverty Schools and Lower Poverty Schools”, *Journal of Urban Economics* 72: 104–122.

Scafidi, Benjamin, David L. Sjoquist, and Todd R. Stinebrickner (2007), “Race, Poverty, and Teacher Mobility”, *Economics of Education Review* 26:145–159.

Springer, Matthew G., Dale Ballou, Laura S. Hamilton, Vi-Nhuan Le, J. R. Lockwood, Daniel F. McCaffrey, Matthew Pepper, and Brian M. Stecher (2010), “Teacher pay for performance: Experimental evidence from the project on incentives in teaching”, Nashville: National Center on Performance Incentives.

Appendix

Table A 1. Effect of the CFT intervention 2014, excluding one school at a time from the analyses (p-values in parentheses)

A. Excluding intervention schools, one by one	(1) *Without school 1*	(2) *Without school 2*	(3) *Without school 3*	(4) *Without school 4*	(5) *Without school 5*	(6) *Without school 6*	(7) *Without school 7*	(8) *Without school 8*	(9) *Without school 9*	(10) *Without school 10*
Test results English	6.720**	5.466*	6.356*	6.186*	5.204*	6.723**	5.970*	7.130**	6.920**	7.022**
	(0.040)	(0.085)	(0.058)	(0.056)	(0.08)	(0.038)	(0.070)	(0.027)	(0.024)	(0.025)
	[6,871]	[6,294]	[6,509]	[6,874]	[6,664]	[6,849]	[6,877]	[6,823]	[6,934]	[6,912]
Test results Swedish	0.954	1.868	0.944	0.613	-0.406	2.750	0.781	0.908	1.872	1.310
	(0.840)	(0.694)	(0.845)	(0.884)	(0.928)	(0.532)	(0.864)	(0.848)	(0.670)	(0.792)
	[7,015]	[6,478]	[6,640]	[7,037]	[6,808]	[6,997]	[7,012]	[6,980]	[7,087]	[7,044]
Test results Math	5.348	5.198	5.411	5.229	6.649	5.884	5.312	5.077	5.561	5.625
	(0.344)	(0.381)	(0.338)	(0.373)	(0.253)	(0.311)	(0.357)	(0.370)	(0.318)	(0.339)
	[6,816]	[6,270]	[6,439]	[6,824]	[6,594]	[6,806]	[6,834]	[6,783]	[6,889]	[6,846]
GPA	4.095	4.522	4.169	3.428	5.220*	4.145	3.891	3.807	4.768*	4.418
	(0.152)	(0.116)	(0.139)	(0.183)	(0.057)	(0.127)	(0.149)	(0.157)	(0.092)	(0.116)
	[8,645]	[8,136]	[8,377]	[8,755]	[8,422]	[8,711]	[8,722]	[8,673]	[8,793]	[8,778]
Admitted upper secondary school	0.079	0.068	0.065	0.045	0.068	0.047	0.046	0.046	0.066	0.061
	(0.203)	(0.279)	(0.324)	(0.442)	(0.278)	(0.424)	(0.427)	(0.432)	(0.288)	(0.325)
	[8,645]	[8,136]	[8,377]	[8,755]	[8,422]	[8,711]	[8,722]	[8,673]	[8,793]	[8,778]

B. Excluding control schools, one by one	(1) *Without school 11*	(2) *Without school 12*	(3) *Without school 13*	(4) *Without school 14*	(5) *Without school 15*	(6) *Without school 16*	(7) *Without school 17*	(8) *Without school 18*	(9) *Without school 19*	(10) *Without school 20*
Test results English	5.830*	7.226**	5.829	7.457**	5.309	7.588***	6.488**	5.751*	5.908*	6.387**
	(0.084)	(0.015)	(0.107)	(0.014)	(0.133)	(0.006)	(0.048)	(0.092)	(0.073)	(0.042)
	[6,738]	[6,944]	[6,601]	[6,909]	[6,528]	[6,669]	[6,925]	[6,834]	[7,010]	[6,876]
Test results Swedish	2.131	2.243	-1.856	2.019	-0.737	2.569	1.119	1.433	0.920	1.348
	(0.656)	(0.619)	(0.699)	(0.664)	(0.907)	(0.618)	(0.804)	(0.778)	(0.844)	(0.782)
	[6,943]	[7,101]	[6,747]	[7,046]	[6,688]	[6,818]	[7,033]	[6,979]	[7,165]	[6,987]
Test results Math	7.580	5.461	4.624	4.800	1.323	8.292	5.783	5.706	5.679	5.719
	(0.287)	(0.369)	(0.542)	(0.473)	(0.773)	(0.146)	(0.314)	(0.373)	(0.338)	(0.302)
	[6,710]	[6,892]	[6,550]	[6,863]	[6,530]	[6,599]	[6,853]	[6,775]	[6,948]	[6,813]
GPA	4.663	4.391*	2.618	4.772*	3.216	4.883*	4.364	4.219	4.787*	4.287
	(0.109)	(0.098)	(0.291)	(0.097)	(0.317)	(0.070)	(0.100)	(0.151)	(0.09)	(0.143)
	[8,666]	[8,881]	[8,420]	[8,766]	[8,334]	[8,499]	[8,716]	[8,692]	[8,909]	[8,663]
Admitted upper secondary school	0.073	0.061	0.029	0.063	0.047	0.091*	0.061	0.045	0.061	0.052
	(0.243)	(0.322)	(0.621)	(0.327)	(0.515)	(0.069)	(0.305)	(0.484)	(0.330)	(0.436)
	[8,666]	[8,881]	[8,420]	[8,766]	[8,334]	[8,499]	[8,716]	[8,692]	[8,909]	[8,663]

Notes: Results (grades) on standardized tests and GPA are percentile ranked within cohort. All regressions control for year and school fixed effects as well as sex, age, born in another country, years since immigration (4 categories), parents born in another country, mother's and father's education level (4 categories), and mother's and father's (log) earnings (see Table 3 for details). P-values (in parentheses) are obtained using wild bootstrap; *** p<0.01, ** p<0.05, * p<0.1. Number of observations in brackets.

Table A 2. Effects of the CFT intervention, separately for boys and girls (p-values in parentheses)

A. Test results	English		Swedish		Math	
	Girls	Boys	Girls	Boys	Girls	Boys
Effect 2013	1.390	2.182	3.610	3.237	0.537	2.239
	(0.636)	(0.509)	(0.334)	(0.272)	(0.928)	(0.658)
Effect 2014	7.163***	5.888	3.719	-0.962	6.120	4.898
	(0.001)	(0.170)	(0.408)	(0.872)	(0.309)	(0.450)
Controls	yes	yes	yes	yes	yes	yes
Observations	3,528	3,611	3,616	3,679	3,487	3,599
R-squared	0.217	0.222	0.225	0.210	0.188	0.164
Mean of outcome var.	36.576	37.774	40.521	32.592	38.010	39.536
B. GPA and admission upper sec. school	GPA		Admitted			
Effect 2013	-1.517	4.921*	-0.054	0.019		
	(0.652)	(0.068)	(0.293)	(0.722)		
Effect 2014	3.084	5.367	0.054	0.061		
	(0.173)	(0.124)	(0.322)	(0.410)		
Controls	yes	yes	yes	yes		
Observations	4,414	4,668	4,414	4,668		
R-squared	0.283	0.253	0.300	0.287		
Mean of outcome var.	38.549	31.166	0.628	0.614		

Notes: Results (grades) on standardized tests and GPA are percentile ranked within cohort. All regressions control for year and school fixed effects as well as age, born in another country, years since immigration (4 categories), parents born in another country, mother's and father's education level (4 categories), and mother's and father's (log) earnings (see Table 3 for details). P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.

Table A 3. Descriptive statistics for pupils in grade 7–9 in intervention and comparison schools, questionnaire data (2012)

	Intervention schools	Comparison schools
Age	13.8	13.8
Girl (%)	50.4	48.8
Swedish as native language (%)	31.0	33.3
Less than 4 years in Swedish school (%)	11.3	9.6
Attends a class for newly arrived immigrants (%)	4.3	1.5
Mother's level of education (%)		
No education	6.2	8.2
Compulsory school	9.9	7.4
Upper secondary education	15.3	13.7
Post-secondary education	21.2	24.9
Information is missing	47.4	46.1
Father's level of education (%)		
No education	7.9	6.7
Compulsory school or lower	7.4	7.7
Upper secondary education	14.0	15.9
Post-secondary education	21.0	19.6
Information is missing	49.7	50.1
Average number of pupils/school	302.5	225.6
Average GPA, school level (gr 9)	168.0	172.0
Number of teachers/school	35.3	31.3
Number of schools	7	7

Note: Average number of pupils per class and school, and average GPA originate from the Swedish National Agency for Education. The number of teachers was collected directly from the schools. All other variables originate from the questionnaire.

Table A 4. Factor analysis: student perceptions of teacher ability

Item	How many of your teachers…	Factor loading
1	…are good at teaching?	0.79
2	…are good at engaging and creating interest among students?	0.79
3	…can tell you what to do in order to learn more?	0.78
4	…give you challenges so that you continue to develop learning?	0.77
5	…give you support and help?	0.82
6	…are good at explaining when you don't understand?	0.78

Note: The factor analysis is conducted on 4,204 students and based on principal components. All items are measured on a scale 1 to 5, where 1 = "None of my teachers" and 5 = "All of my teachers". The retention of factors is based on the Kaiser criterion (eigenvalues > 1). Only the eigenvalue for the first dimension (3.73) exceeds 1. The first factor explains 55 percent of the variation of the variables. Bartlett's test for sphericity: $p > 0.001$. Kaiser-Meyer-Olkin measure of sample adequacy: 0.90.

Table A 5. Factor analysis: student perceptions of the learning environment in the classroom

Item	To what extent would you describe the situation in your classroom in the following way?	Factor loading
1	It's peace and quiet, and it's possible to concentrate on school work	0.83
2	The atmosphere is nice and positive	0.86
3	It's easy to participate in discussions	0.73

Note: The factor analysis is conducted on 4,739 students and based on principal components. All items are measured on a scale 1 to 5, where 1 = "Never/almost never" and 5 = "Always/almost always". The retention of factors is based on the Kaiser criterion (eigenvalues > 1). Only the eigenvalue for the first dimension (1.97) exceeds 1. The first factor explains 66 percent of the variation of the variables. Bartlett's test for sphericity: $p > 0.001$. Kaiser-Meyer-Olkin measure of sample adequacy: 0.65.

Table A 6. Robustness of difference-in-differences estimates: individual items in the student questionnaire as outcome variables (p-values within parentheses)

	(1)	(2)	(3)
A. Teacher ability			
Item 1	0.172***	0.210***	0.174***
	(0.010)	(0.000)	(0.002)
Observations	5,069	4,256	4,256
Item 2	0.171***	0.178**	0.149**
	(0.008)	(0.016)	(0.048)
Observations	4,817	4,063	4,063
Item 3	0.146*	0.134*	0.108
	(0.060)	(0.052)	(0.240)
Observations	4,964	4,181	4,181
Item 4	0.133	0.110	0.089
	(0.248)	(0.216)	(0.370)
Observations	4,845	4,111	4,111
Item 5	0.149	0.162**	0.131
	(0.156)	(0.050)	(0.182)
Observations	5,007	4,227	4,227
Item 6	0.167*	0.205***	0.164
	(0.084)	(0.038)	(0.130)
Observations	5,051	4,256	4,256
B. Learning environment			
Item 1	0.111*	0.160***	0.118**
	(0.070)	(0.000)	(0.018)
Observations	5,109	4,315	4,315
Item 2	0.022	0.057	0.018
	(0.716)	(0.308)	(0.718)
Observations	4,966	4,228	4,228
Item 3	0.044	0.073	0.050
	(0.390)	(0.204)	(0.336)
Observations	4,906	4,176	4,176
Control variables	No	Yes	Yes
School dummies	No	No	Yes

Note: All outcomes measured on a scale 1–5, where a higher value implies a more positive assessment of teacher ability/the learning environment. See Table A 4 and A5 for descriptions of the survey items. Models (2) and (3) include the following control variables: sex, age, Swedish as native language (yes/no), have attended Swedish school for less than four years (yes/no), attends a class for newly arrived immigrants (yes/no), father's educational level (5 levels), mother's educational level (5 levels). Models (3) also include school fixed effects. P-values (in parentheses) are obtained using wild bootstrap. *** p<0.01, ** p<0.05, * p<0.1.