Stockholm School of Economics
Department of Economics
659 Degree project in economics
Spring 2016

THE POWER OF PEERS IN ACADEMIC PERFORMANCE

– IMPLICATIONS OF ABILITY SORTING IN SWEDISH UPPER SECONDARY SCHOOLS

Marcus Strinäs (23160) and Edvard von Sydow (23080)

The formation of human capital is an integral part of economic growth. If we can optimize classroom organization, long-term economic performance is likely to be waiting on the other side. Consequently, how peers affect one another in a classroom and school setting could be a question of considerable economic importance. In this thesis, we conduct an event study in which we analyse how student outcomes evolved as segregation by ability increased when the seven largest municipalities in Sweden changed admission system to their public upper secondary schools. By using publicly available aggregate data from the Swedish National Agency for Education, we find evidence that graduation rates as well as fail rates increased after the changes. Although appearing as a contradiction, these results indicate that the least skilled students who, absent the reform, would have quit school early, now complete their education – with higher fail rates at the aggregate level as a consequence. These results have proven robust in Sweden's three largest municipalities and can with some degree of certainty be related to the change in student allocation and hence to peer effects. We find tentative support that sorting by ability could be beneficial for the lowest performing students.

Keywords: Peer effects in education, School admission reform, Ability sorting, Human capital formation, School choice

JEL: I21, I26, I28, J24

Supervisor:	Johanna Wallenius
Date submitted:	15 May 2016
Date examined:	31 May 2016
Discussants:	Tilde Halvorsen and Anton Sundberg
Examiner:	Karl Wärneryd

Acknowledgements

First and foremost, we want to thank our supervisor Johanna Wallenius for her sharp questions and encouragement along the way. You gave us an indispensible outside perspective on for us the most familiar subject of Swedish upper secondary schools.

For their untiring support and help in the empirical execution of this thesis we want to extend our most heartfelt thanks to Dany Kessel and Elisabet Olme. It has been an honour as two undergraduates to be included in your research team.

To Mattias Hallberg, without you, we would have spent weeks instead of days collecting data. For that we are forever thankful.

We also want to thank Elin Mohlin, Jonas Cederlöf and Niklas Blomqvist, who we believe are sure to contribute in unravelling the mystery that is the mishmash of Swedish educational policies.

Friends and family.

Table of Contents

1 Introduction

To what extent is a student's classroom achievement conditioned by the quality of her classmates? This is an important question to address for any policymaker who seeks to optimize human capital formation in a society. There is full consensus that education is a crucial component in the prosperity of a nation, but the question of how to maximize the returns to education remains disputed.

One such dividing issue is whether or not policymakers should try to steer the student composition in schools. Is there a problem that high-skilled students go to school separately from low-skilled students – or is everybody better off that way? At the heart of this issue lies the question of *peer effects* in education, how students affect one another in the school and classroom setting.

From an economic perspective we can think of a number of ways peer effects might come into play in education. Arguing from the point of view of a behavioural economist we might think that being included in an ambitious and talented environment would be good for low-skilled students, and that student composition in schools therefore should be mixed in order to maximize the returns to education. Conversely, arguing from the point of view of a neoclassical economist, we might think that there would be economies of scale in educating students of similar ability, and that student composition therefore should be homogenous.

What is the empirical support behind such theories of peer effects? In this thesis, we try to isolate the importance of peers on academic performance by conducting an event study in which we examine how a number of outcome variables evolved as student segregation increased in the seven largest municipalities of Sweden. Specifically, we look at *Grade Point Average* (GPA), the *fail rates* in *Mathematics A, English A* and *Swedish A* as well as the overall *graduation rates* from the public upper secondary schools in our sample. The event that we exploit as an exogenous source of student segregation is the year the municipalities changed admission system to their public upper secondary schools. This reform led to a reallocation of students across schools, where the most talented students were pooled together in certain schools and the least talented students in others.

We find evidence that, post the admission reform, students tended to fail the courses to a higher extent than before, but more importantly, though, seemed to increase their overall graduation rate. From a human capital formation perspective, the latter is of immense importance. In Sweden 2015, those who lacked post-secondary education had nearly half the employment rate in comparison to the total population (SCB 2016). This underlines the macroeconomic importance of students actually graduating. Whether or not these effects are the causal outcome of peer effects, we discuss thoroughly in the thesis. In two instances in the big cities, regarding the fail rate in English A as well as the overall graduation rate, our robustness tests indicate however that this seems to be the case.

This thesis consists of ten sections. Sections 1–4 present the objective with the thesis and provide a theoretical overview of peer effects, as well as an institutional background to the Swedish school system and its changes since the 1990s. Sections 5–6 present our data and empirical strategy. In section 7–8 we describe our findings and subject them to a series of robustness tests. These findings are discussed in section 9. In section 10, this thesis concludes by summarizing our key lessons.

2 Research Purpose and Definition

Our aim with this thesis is to identify and quantify peer effects by utilizing an exogenous shock in the admission system of public upper secondary schools. From this we hope to be able to draw conclusions with regard to policy implications. We thereby want to state the following hypothesis:

H_1: Fail rates, GPA, and graduation rates changed as a consequence of changing peer effects after admission reform in Swedish upper secondary schools.

We define peer effects openly. They are the effects on performance from socialization with and the presence of peers in one's surroundings. Peer effects are in our view inherently interpersonal. Within this definition we can fit a wide range of circumstances. Peer effects may be observed in a factory along a production line. They could be read from the academic performance between siblings as in Joensen and Nielsen (2015). And they can, as will be the focus in this thesis, be thought of in the framing of a school and classroom.

The peer effects that we will try to identify will be, due to limitations of the available data, at a school and programme level. In this sense, we study the effects of the overall peer environment at a school. Is it possible to quantify the academic effects students have on one another, in the classroom and the corridor outside of it?

Following Tincani (2015) we divide peer effects into direct, indirect and correlated peer effects. Of primary interest in this thesis are the direct and indirect peer effects. Direct peer effects are the peer-to-peer interaction that affects performance on a classroom level. Indirect effects also appear in this setting, but look instead at how peer composition affects other factors, for example teacher efficiency. Correlated peer effects are effects of unobserved classroom characteristics that depend on the composition of peers as well.

3 Literature Review

3.1 Peer effects

The literature on peer effects is diverse and covers an array of academic disciplines. The interest garnered in the subject within economics stems from the potential economic gains that could result from understanding the underlying dynamics of peer effects. If one views it as a process of developing human capital, strong arguments can be made that increased knowledge will entail considerable macroeconomic gains (Hoxby 2000). Existing theories stress the importance of matching people and the implications it has on industry production, development and growth (Kremer 1993). For example, as argued by Bénabou (1996), there can be costs related to more diverse and heterogeneous groups. Bénabou's finding is however, that in the long run, overall heterogeneity decreases as a result of mixing people and costs go down, eventually offsetting the initial negative impact.

As techniques to identify and isolate supposed peer effects have developed and been refined, we find that there seems to be a convergence on the opinion supporting their existence. Manski (1993) highlighted the need to look for *endogenous* effects of the interaction between the members of a group. He separates these from *exogenous* and *correlated* effects. Correlated effects are a result from exposure to the same environment, for instance attending the same class. Exogenous effects evolve from characteristics shared by group members, usually demographics. The model Manski presents is known as the linear-in-means model, arguing that peer effects work as a function of the mean performance of one's peers while also addressing the circular problem of mutual simultaneous influence between peers.

A previously used model of peer effects failed to highlight the eventual efficiency gains from peer effects and limited peer effects to be purely distributional by nature. For one group to perform better by reallocating a good peer, another has to perform worse and vice versa (Hoxby 2000).

For an extensive list on the many reports of peer effects, Sacerdote (2014) briefly summarizes a rather wide selection. Subjects range from Norwegian fathers' likelihood to take paternity leave as an effect of peers doing so, to how probable students are to attend university as an effect of the share of low-ability students among peers.

3.1.1 Peer Effects in Education

The search for peer effects in education has yielded some results in quantitative studies. Hoxby (2000) found peer effects while looking at the race and gender composition in Texas schools. Sacerdote (2001) observed an effect on the grade point averages (GPA) between randomly allocated roommates at Dartmouth College. Using the same data as Sacerdote (2001), Bhattacharya (2009) found that though segregation by race or prior academic performance did not affect GPA mean or median, it polarized the aggregate results, increasing their spread.

In an experiment at the United States Air Force Academy, Carrell, Sacerdote, and West (2013) attempted policymaking to improve the lowest ability students' academic performance by assigning them into designed peer groups. Due to increased homophilic socialization, the design proved detrimental. The result highlights the challenges regarding the application of theories in practice.

Joensen and Nielsen (2015) studied educational peer effects in a household setting using a pilot scheme of an advanced mathematics course in Danish high schools. They found that the choices of an older sibling affects those of a younger, indicating sizable heterogeneity though. The effects are strongest between siblings close in age, especially brothers. In their setting, peer effects are however limited to work in one direction, from the older to the younger sibling.

Sacerdote (2014) argues for the relaxation of the assumptions in the linear-in-means model. Tincani (2015) instead assumes non-linearity and adds student rank concerns as a driving force behind the effort students will spend. She claims achievement is valued in relation to that of peers and argues that the classroom environment is more accurately described through the entire distribution of ability, rather than the mean ability. What drives student effort in this setting is the presence of peers of similar ability for students to outrank without spending too much effort. Tincani shows that especially low-performing students perform better when such conditions hold.

3.1.2 Peer Effects Outside of Education

There is a growing strand of research on the subject of peer effects within production (Horton and Zeckhouser, 2016). One set of research links peer effects to theories on peer pressure. Kandel and Lazear (1992) developed a theory of peer pressures in partnerships. Peer pressure could be used to combat free riders within partnerships. They categorize peer pressure into two components: one internal, the feeling of guilt, and one external, which is shame, both induced by not doing your best.

By using high-frequency data, Mas and Moretti (2009) found evidence that worker effort is related to the productivity of workers that can observe each other. They explain this effect through mechanisms similar to Kandel and Lazear's shame and guilt. Falk and Ichiro (2006) found evidence for the existence of peer effects through a controlled field experiment, where those who had a peer in view, produced more output.

3.2 Related Effects of Changing Admission Policies

The specific set of admission reforms that we intend to investigate has previously been studied in search of other outcomes than those connected to peer effects. Karbownik (2014) looked at the effects on teacher mobility in relation to when Stockholm decided to switch its upper secondary school admission system for publicly owned schools from one primarily based on residential proximity to one based on elementary school grades in the fall 2000. He found, through a difference-in-differences analysis, that after the reform there was no effect on teacher mobility within the first year, using registry data. There was, however, a change in mobility after that time span. Karbownik linked this change in mobility to change in student quality – a proxy based on the grade level of incoming students – coming to the conclusion 'that a 10-percentile-point increase in student quality decreases the probability of a separation by up to 9 percentage points.' This result is controlled for in terms of potential other underlying causes, such as share of immigrant students, parental income, and parental education. The result is more or less the same across different kinds of teachers in terms of subjects taught, work experience, and gender among other factors.

Söderström and Uusitalo (2010) studied the Stockholm case as well with the help of registry data, focusing on its effect on segregation. They found that the reform, originally aimed at reducing the residential segregation, did result in other kinds of segregation. By comparing the sample variation to a constructed sample in which the students had been randomly allocated to a school, they constructed a Duncan dissimilarity index indicating a deviation from the expected segregation. A difference-in-differences analysis in relation to other municipalities in Stockholm County too lent support for this finding. Segregation by ability increased as grades became the deciding factor, it did however also increase segregation in terms of family background, for example parental education and income, as well as immigrant status. The former increased in a manner, which can be explained by the effects of sorting by ability. The latter, however, cannot.

4 Institutional Background

4.1 The Swedish School System

The Swedish school system comprises of nine years of compulsory education, where students in general are between seven and sixteen years of age, followed by three years of voluntary upper secondary schooling. According to the Swedish National Agency for Education, Skolverket, about 85 per cent of the students complete compulsory school and are qualified to upper secondary education (Skolverket

2015a). Among those, almost all students continue directly to upper secondary school. In 2014 the share was slightly over 98 per cent (Skolverket 2016a).

The upper secondary school system consists of a number of different programmes, among which the fundamental difference is whether or not the programmes make students eligible for higher education. Some of the programmes are specialized in providing work related skills, such as manufacturing or retailing, whereas other programmes provide broader skill sets in preparation for higher education.

In Sweden, most schools are publicly owned and run by municipalities. However, the number of students in independent schools – schools that receive public funding but operate independently of the public school system – has increased steadily since the beginning of the 1990s. In 2014, 26 per cent of all students in the upper secondary stage of their education were affiliated to independent schools. The corresponding number for compulsory schools was 14 per cent (Skolverket 2015b).

Concerning the three dependent variables of interest in this thesis: GPA, the fail rate in so called core subjects – Swedish, English and mathematics – and the overall graduation rate, we report the averages during the latest ten years in Sweden for comparison, in Table A.1, Appendix A. In order to graduate from the Swedish upper secondary school in Lpf 94 – the upper secondary education system in place for students starting prior to 2011 – a student had to receive grades in all the 30–35 courses that normally were included in each programme. Note that this did not imply that the student had to receive *passing* grades in all courses. Instead, a student failed to graduate when the school lacked assessment grounds for the student. This happened mostly for one of three reasons: 1) classroom absence had been too high, 2) the mandatory project assignment had not been handed in on time, or 3) the student had chosen to leave the school before graduation (Skolverket 2016c).

4.2 Institutional Change in the Swedish School System

The Swedish educational system went through a series of fundamental institutional reforms during the 1990s. Holmlund et al. (2014) represents a very comprehensive study of the reforms undertaken in Swedish compulsory and upper secondary education during the time. Their report follows and analyses the effects of four distinct reforms, whereof three are of particular interest to this thesis.

First, the Swedish school system was decentralized in 1991, with municipalities being given the main responsibility in ensuring compulsory and upper secondary education for their inhabitants. Still, the state retained some degree of control through a system of management based on national goals.

Second, a bill that was introduced in 1992 proposed new curriculums for the compulsory and the upper secondary schools, named Lpo 94 and Lpf 94 respectively. They were both adopted, and Lpf 94, the one of specific interest to us, came into effect for students starting their upper secondary schooling in the academic year of 1994/95. It restructured all upper secondary school programmes, introduced a new grading system and standardized the length of all programmes to three years.

Thirdly, independent schools were allowed in Sweden as of July 1992. As stated before, independent schools receive public funding but operate as separate organizations. International equivalents can be found in the American charter schools or the British academies. At the time of their introduction in Sweden they were allowed to take out a fee from students, as long as it was deemed 'reasonable'. Five years later in 1997, however, the succeeding government removed this option in return for more extensive public funding for independent schools.

4.3 Admission in the Swedish Upper Secondary School System

As the Swedish school system is organized at the municipal level, there has been and still is a high degree of variety in how the local school systems are organized. Today, the freedom assigned to the municipalities is somewhat lesser than it was before, given more thorough state regulation recently, but the degree of freedom is still substantial in comparison to state administered school systems.

One way in which this freedom manifested itself was that different municipalities applied different admission policies to their upper secondary schools. By gathering information from local education officials, we know that most municipalities applied proximity-based admission prior to 2000. Students applied for a certain programme and their grades determined admission. If there were several schools providing the same programme, this admission system stated that students be assigned to the school closest to their residence. Other examples of municipal admission policies include randomized admission to popular schools – albeit given that the students had sufficiently high grades to be accepted to the programmes – and a third hybrid version of both proximity-based and conditionally randomized admission. The different admission policies prior to free school choice for the municipalities in our sample are summarized in Table A.2 in Appendix A.

Beginning in 2000, many municipalities altered their admission policies. Henceforth, in a variety of municipalities, students were allowed to apply to *both* programme and school with their grades. This meant that students no longer were bound to the schools in the area where they lived or had to rely on luck in order to be accepted to the most popular schools elsewhere. Instead, the higher GPA a student had, the more mobility he or she had with respect to which schools that were possible to be admitted to. Not surprisingly, this led to segregation by ability, where high-performing students – and conversely low-performing students – to a greater extent were pooled together on different schools (Söderström and Uusitalo 2010). Note that this did not systematically change the aggregate ability distribution among the students. Instead it was reallocated across different schools.

From communication with local education officials, we have learnt of primarily two reasons as to why municipalities decided to change admission system. First, the reform was implemented in order to mitigate residential segregation and open up the most popular schools to the most talented students – regardless of where they lived. This was considered desirable not only on the individual level, but also for the school system as a whole since schools now would have to compete in order to attract students. In turn, this was thought to improve quality in the local school system.

Secondly, the increasing number of independent schools at the municipal level, which by default applied grade-based admission, motivated a transition for the public schools. It became a liability for the public schools that they, in practice, only could accommodate students from their near surroundings, as opposed to independent schools which could attract students from everywhere. Moving from proximity-based to grade-based admission could hence be seen as an important step to balance the operational terms between public and independent schools.

4.4 Grade Inflation

There is also the issue of grade inflation in the Swedish schools system, which to a degree can be linked to the marketization of and increased competition in the educational market (Vlachos 2010). As reported by him, the share of students leaving upper secondary school with the best grades increased dramatically and the average GPA rose significantly in from 1992 to 2008. In our data we can find similar results. As shown in Figures 4.1 and 4.2, there is definitely an interesting development for many of our variables. Graduation rates are trending upwards, where the mean in our baseline sample set, from which we derive all of the others, show a noticeable increase. The mean fail rates are more or less halved for all three subjects as well.

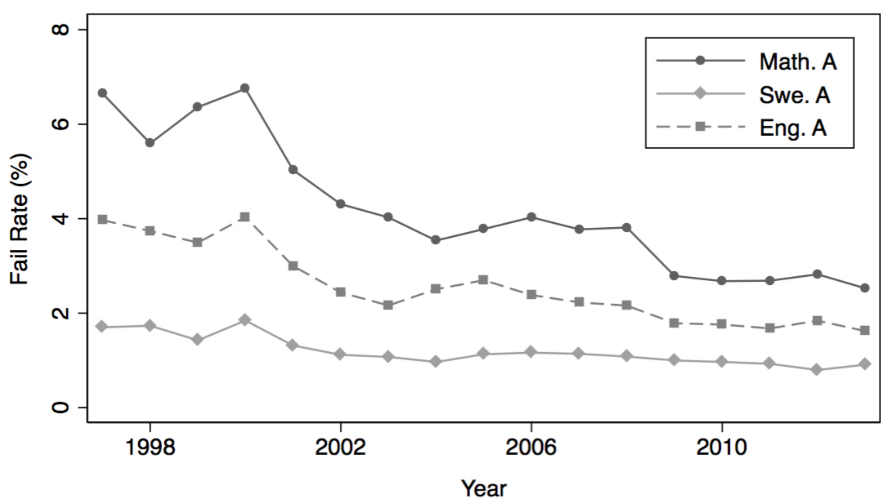


Figure 4.1: *Average fail rates over time in selected courses in our baseline sample*

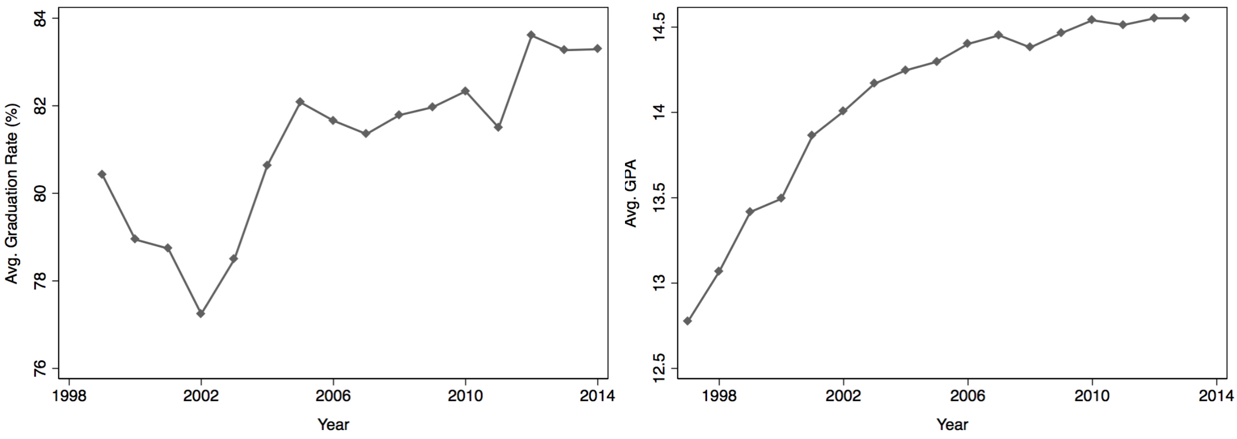


Figure 4.2: *Average GPA and graduation rate over time in our baseline sample*

5 Data

5.1 Accessed Data

The data used in this paper comes from the Swedish National Agency for Education and its online database SIRIS. In total, three sets of data have been used in creating our complete data set: one data set on upper secondary school leaving certificate statistics, another on graduation rates and the third on student characteristics. All sets are aggregated at the level of each programme taught at every publicly owned upper secondary school in Sweden. The SIRIS data is comprehensive but due to reasons of integrity, statistics are omitted to maintain anonymity when the sample group grows small.

The data set on school leaving certificate statistics includes the number of students receiving a certificate, the pass rates in the mandatory courses, and the average GPA each year within the time frame Lpf 94. We are therefore subject to a lag, as grades are not reported until a student graduates. School leaving certificates data is therefore collected for the academic years 1996/97 through 2012/13. Though reported, the data for the academic year 2012/13 is implicated by yet another reform. This divided certain schools into several different administrative objects, leading to SIRIS reporting results from each subsection, complicating long term comparability.

Looking at the data set on graduation rates, we can identify the percentage of students graduating within four and five years of starting their studies at the upper secondary school level. This data covers the academic years 1998/99 through 2013/2014, the 1998/99 data being the first available observation for statistics regarding those starting upper secondary schooling in Lpf 94.

Lastly, the data on student characteristics presents the fraction of students in each year. It also shows the composition of students based on their gender, and family background. SIRIS categorizes a student to have a foreign background if the student is, or if both its parents are, born outside of Sweden. Students are classified as coming from a highly educated family if they have one parent with post-secondary education. This data covers the academic years from 1997/98 to 2011/12 for all Swedish schools, independent and public.

As the characteristics data is reported on an annual basis, we have decided to create averages over time periods to mirror the time frames of the other statistics where possible. As a clarifying example, for the fail rates statistics of 2002/03, they are reported in the graduation certificates of that spring semester. Students graduating that semester are likely to have started their studies in 2000/01. The averages of the characteristics data in this example are consequently for 2000/01–2002/03.

5.2 Excluded Data

We have through contact with upper secondary school admission officials limited the data of interest to the twelve largest municipalities, those mentioned in Table A.2. Regarding the many municipalities overlooked, we found a cut-off point coming from the scale at which an upper secondary school optimally functions. When municipalities drop below a certain size, they start to limit each programme to be localized at only one upper secondary school to achieve scale economies. By doing so, they effectively remove the purpose of an admission reform as the one we describe. In the few instances where the same programme is available at more than one school in smaller municipalities, it represents such a small share of the total number of students that we deem the inclusion of the entire municipality likely to primarily add noise.

Seeing as five of the municipalities in our data set did not change admission system until the new curriculum (Gy 11) is introduced, it means that the effects from the reforms would be seen earliest in 2014. In turn, this implies that we cannot observe them as treated within our accessed time frame or with comparable variables. Therefore, we have chosen to run the regressions without these municipalities as they fail to add observations to the transition we use as an identifier. For comprehensive information on treatment year, see Table B.1. Hence, this leaves us with data from all public schools in the following seven municipalities, named in alphabetical order: Gothenburg, Helsingborg, Linköping, Lund, Malmö, Stockholm, and Uppsala.

As for not including any data from independent schools, the reasoning behind is simply that the reform only affected them indirectly. The admission policy of students applying directly to a programme at a specific school was already effective in the private independent school sector.

There is one more exclusion necessary before we can go further. It regards any samples from the individual programme, henceforth referred to as IV. IV was created for students, who upon finishing compulsory school are ineligible to attend upper secondary school. Within Lpo 94, eligibility was conditioned on a passing grade in the three core subjects: Swedish, English, and mathematics. Attending IV has therefore never been much of a choice but rather a means of last resort. Municipal authorities offer the student a spot in the programme after quite a different application process. To include IV statistics is therefore not compatible with the research question, and we have therefore omitted this programme altogether.

Variable	Description
Avg. Female N	Average share of female students over the previous N years
Avg. Foreign N	Average share of students with a foreign background over the previous N years
Avg. High Ed. N	Average share of students from families with post secondary education over the previous N years.
Eng. A	Per cent of students with a fail in English A on their school leaving certificate
Female	Per cent of female students
Foreign	Per cent of students with a foreign background
GPA	Mean grade point average, where the highest possible is 20 and lowest 0
Grad. Rate.	Per cent of students graduated within four years of starting upper secondary school
High Ed.	Per cent of students from families with post secondary education
Math. A	Per cent of students with a fail in Mathematics A on their school leaving certificate
Students graduating	Number of students graduating
Students starting	Number of students starting four years before reported graduation rate
Swe. A	Per cent of students with a fail in Swedish A on their school leaving certificate
τ	Time trend
τ_m	Municipality specific time trend
Town/City name	Municipality dummy
Treatment	A dummy that transitions from 0 to 1 the year of treatment
Year	Year dummy

Table 5.1: *Variable List*

Variable	Obs	Mean	Std. Dev.	Min	Max
Avg. Female 3	3435	57.20	20.11	2.67	100
Avg. Female 4	3435	57.17	20.15	2.67	100
Avg. Foreign 3	3389	26.59	17.16	4	100
Avg. Foreign 4	3389	26.43	16.88	4	100
Avg. High Ed. 3	3790	48.85	20.75	7	100
Avg. High Ed. 4	3790	48.67	20.72	7	100
Eng. A	4125	3.62	6.50	0	52.9
Female	3435	57.49	20.35	2	100
Foreign	3389	27.02	17.77	4	100
GPA	4141	13.65	1.77	8.2	18.4
Grad. Rate	4032	77.64	15.62	3.9	100
High Ed.	3790	49.30	20.86	7	100
Math. A	4136	5.69	9.59	0	70.6
Students graduating	4620	54.27	43.69	1	273
Students starting	4318	64.87	52.96	0	299
Swe. A	4062	1.55	3.43	0	34.2

Table 5.2: *Summary Statistics*

5.3 Subsections of Interest

Within this data set there are additional restrictions that could be of interest. One such restriction is to look at the three largest cities only. Stockholm along with Gothenburg and Malmö are the only three municipalities with a population above 200,000 during the entire period studied. Residential segregation and its implications on the upper secondary school admission systems is likely to have created bigger administrative problems the bigger the municipalities were. Comparing the number of public upper secondary schools in the municipalities in our time frame, leaving out IV and accounting for trivial name changes, the size and availability of options for students become evident, see Table 5.3.

	No. of Schools
Stockholm	31
Gothenburg	17
Malmö	18
Uppsala	11
Linköping	6
Helsingborg	7
Lund	4

Table 5.3: *Schools per municipality*

Going to the core of the reform and the problem it was trying to address, it is especially interesting as well to look at the effects it has had on the two largest programmes in terms of students: the natural and social sciences programmes. They are the two primary programmes for students wanting to continue to higher education and are offered at the widest range of schools.

To summarize, the subsections of interest are described in Figure 5.1 on the following page.

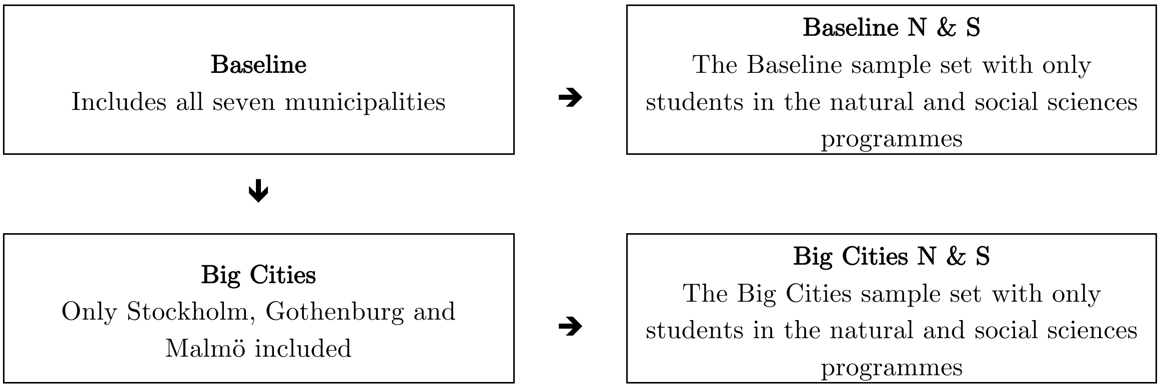


Figure 5.1: Subsections of interest

6 Empirical Strategy

In this thesis, our objective is to identify and quantify peer effects and their impact on academic performance. From previous research, for instance from the academic contributions of Söderström and Uusitalo (2010), we know that the Stockholm public upper secondary school system became more segregated in terms of student ability when the municipality started to implement grade-based admission policies. Assuming external validity of their findings, we can with a relatively high degree of certainty claim that similar developments ought to have occurred as other municipalities in Sweden changed to a grade-based admission system as well.

The sudden nature of these changes in admission policies allows us to identify a close to causal effect of increased student segregation using an event study analysis. Since the changes in admission policies to the public upper secondary schools led to a higher rate of ability sorting between students, while in central aspects holding the other parts of school system fixed, we believe this approach to properly identify the effect of being treated with a new set of peers on student outcomes.

The event study approach used in this thesis is essentially an expanded difference-in-differences analysis with multiple periods as well as multiple treatment and control groups. Our variable of interest, $Treatment$, is a dummy variable that shifts from 0 to 1 at the point in time each municipality changes admission system. At that point, the control groups are all the municipalities in the sample set that have not changed system yet though are going to in our timeframe. This allows for an intercept shift during the period. By testing for significance and examining the coefficient for the dummy variable, we can apprehend the effect of the admission reform on student outcomes and consequently – assuming that our identification strategy holds – infer the impact of peer effects.

The benefit with this empirical strategy is that it allows us to observe the same exogenous shock at a variety of places and in different points in time. In a more simplistic difference-in-differences setting, for example with one treatment group and one control group observed in two points in time, we would worry that something else – apart from the treatment – might contribute to the outcome and hence cause omitted variable bias in the model. With an event study of this kind, this is unlikely since we require the same result to be observed at different places and different points in time in order to be deemed significant and trustworthy.

As described in the section 5, we are forced to limit the scope of this event study to GPA, the fail rates in the three mandatory courses Mathematics A, Swedish A and English A, as well as the overall graduation rate, as defined by the share of students who graduate from the upper secondary school within four years since their beginning. Looking at GPA as an outcome variable enables us to capture a more general effect on the average student after the change in admission policy. The two other types of outcome variables, however, are more likely to capture the specific impact of peer effects on the low-skilled part of the student distribution, since low-skilled students are subject to a larger risk of failing courses or not graduating.

In our primary regression specification we estimate the following model

$$Y_{imt} = \alpha_m + \lambda_t + \beta_1 Treatment_{mt} + \gamma\tau + X_{imt}\delta + \varepsilon_{imt} \qquad (1)$$

where Y_{imt} is an outcome variable (either fail rate, GPA, or graduation rate) in municipality m at the school-specific programme i at time t, α_m is a municipality fixed effect and λ_t is a year fixed effect, $Treatment_{mt}$ is a dummy variable equal to 1 after the change admission policy, τ is a time trend, X_{imt} is a set of time-variant control variables on the demographic composition of the students, and ε_{imt} is the idiosyncratic error of each school, municipality and time period. The set of time-varying control variables that are included in X_{imt} are the share of students whose parents have post-secondary education, the share of students with foreign background and the share of females in each school-specific programme and year. Note that the intercept is given by the municipality fixed effect and the year fixed effect taken together. In order to create a constant reference group, the intercept has consistently been set to the first year for which we have data in the municipality of Stockholm.

Depending on how we think about peer effects different outcomes are plausible in expectation. If we believe that low-skilled students benefit from being treated by the presence of high-skilled students in the classroom, we should expect to see a higher fail rate in the selected courses as well as a lower graduation rate after the reform. The intuition behind this is simple: if peer effects work in this direction, going to school with less skilled classmates should translate into worse educational outcomes. This should be broadly in line with linear-in-means model of how peer effects work described in section 3. On the contrary, if we believe that low-skilled students benefit from going to school with equally skilled classmates, we should expect the fail rate to go down and the graduation rate to increase after the reform in admission system. Such an effect should be more in line with the notion of peer effects stemming from rank concerns, where low-skilled students who previously were desperately behind the other students in the ability distribution now comes closer to the mean and all of a sudden have a realistic competition incentive to try and leapfrog the other students. Note that both these scenarios are consistent with our main hypothesis formulated in section 2.

6.1 Underlying Assumptions

The empirical approach in this thesis rests on the identifying assumption of *parallel trends* between the observed municipalities. That is, absent the change in admission policy, each treated municipality would have had the same trend with regard to the outcome variables as the untreated municipalities. Seeing as our empirical strategy is an expanded difference-in-differences approach, it should not come as a surprise that we have to impose the same identifying assumption.

As have been underlined by Pischke (2007), when we have multiple periods as well as multiple treatment groups, visual inspection of the parallel trends assumption becomes more difficult – as different municipalities are treated at different points in time. Given that treatment has an effect, we should expect to see different municipalities shifting levels at different points in time, making it difficult to infer parallel trends only by visual inspection. We have, however, included graphical representations of the municipal time series evolution of our outcome variables in Appendix C, Figures C.1–4. These graphs essentially confirm the difficulties of visual inspection.

An alternative way to solve this is to conduct placebo testing, by which we include lags and leads of the treatment in the regression model. Of critical importance in such testing is that the leads of the treatment variable, in which we presume that treatment occurred at time $(k_m - j)$ instead of time k_m, will prove insignificant in order for us to claim causality. This will be thoroughly explored in section 8 of this thesis.

Another important assumption underpinning this empirical strategy is that the change in admission policy was uncorrelated with any unobserved municipality features that might have had an impact on the dependent variables. At the heart of this assumption is that the municipalities *only* changed the admission system when they became treated, and not for example implemented new educational measures in coping with declining results. If that were the case, we would have a problem with omitted variable bias in our results, leading us to believe that the impact of policy shift was either smaller or larger than it really was.

Neither must the reason why municipalities changed admission policy be related to past or future values of the dependent variable. This is called the *strict exogeneity* assumption. For example, if the decision to change admission system was a response to weak educational performance in period $t-1$, manifested as a shock to the idiosyncratic error term $\varepsilon_{im,t-1}$, we could be led to believe that the change in policy had positive effect at time t when things were actually going back to normal.

In algebraic terms, let $\boldsymbol{Z}$ denote the matrix of all the vectors in our data set that include our explanatory variables. Given strict exogeneity, the following must hold:

$$E(\varepsilon_{imt} \mid \boldsymbol{Z}) = 0 \rightarrow Cov(\boldsymbol{Z}, \varepsilon_{imt}) = 0 \tag{2}$$

This means that the expected value of the idiosyncratic error ε_{imt}, given the explanatory variables for *all* time periods, must be equal to zero. Otherwise, we would once again have a problem with omitted variable bias in our regressions. This is a strong assumption since it implies that the idiosyncratic error term not only has to be uncorrelated with the explanatory variables in the same time period, but also with future and past observations of the explanatory variables.

However, given that we observe a variety of municipalities that change admission systems at different points in time, we have to some extent safeguarded ourselves from systematic violations of the strict exogeneity assumption. Naturally, individual breaches might occur, but the only way our results will be seriously biased from breaking this assumption is if all, or most, of the municipalities in our sample sets exhibit this problem. This is less likely to happen than if we only studied two municipalities. This will be addressed more thoroughly in section 9 of this thesis.

6.2 Other Considerations

In order to capture the pure impact of peer effects and not risking interference from other possible changes after the event, we have chosen to limit the observations used in our regressions to include all available data up to one year *after* the change in admission policy. In a normalized time setting – where the event is observed at year 0 – this approach means that the regressions will be based on all historical data up to year 1 for each municipality, as is illustrated in Fig. 6.1 below.

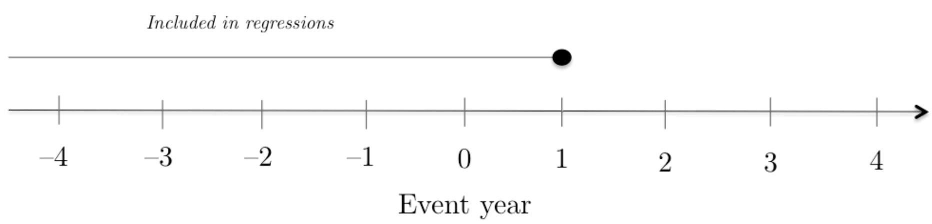


Figure 6.1: *Normalized timeline example*

This is crucial since Karbownik (2014) has shown that teachers respond to admission changes as well by increasing their separation rate from schools that are exposed to negative shocks in terms of skilled students. In other words, teachers are less likely to remain in schools with an increased concentration of low-performing students. Consequently, if we were to analyse the fail rate or the graduation rate based on the entire time series, we would risk capturing the effect of teacher substitution rather than peer effects. This is also the reason behind why we have chosen to observe the fail rate in Mathematics A, English A and Swedish A. These are courses that are most commonly taken and graded the first year of upper secondary school, meaning that the teachers most likely will not have had time to react to any negative – or positive – shock in terms of student allocation yet.

To deal with the concerns of serial correlation between the error terms in our data, we consistently use clustered standard errors at the programme level of each school. As have been argued by Bertrand et al. (2004), standard errors are commonly understated in difference-in-differences analyses. Seeing as our empirical approach builds upon the difference-in-differences framework, this is also the case for our event study. This problem, however, should easily be resolved by clustering. Angrist and Pischke (2009) have pointed out that at least 42 clusters are needed to properly address the issue of serial correlation. Given that we have at least 58 – and most often substantially more – well-defined clusters in our data set, we can be confident that serial correlation will not underestimate our standard errors.

Furthermore, to address the fact that different municipalities, schools and programmes differ in number of students, we consistently use weighted regressions. Different observations are attached with different weights tied to each observation's share of the total number of students. In this setting, larger municipalities, schools and programmes are being assigned more importance in our regressions than do smaller ones.

There is also the issue of grade inflation that needs to be addressed in our empirical strategy. The Swedish National Agency for Education stated in a PM 2012 that their 'overall assessment' was that the Swedish grades have been exposed to inflation since the introduction of the new grading system in the middle of 1990s (Skolverket 2012). This has also been addressed by Vlachos (2010). To account for grade inflation in our regression we have included a time trend τ that is supposed to pick up any such variation stemming from a long general time trend. Though we do not want to downplay the overall importance of performance inflation, we deem our empirical strategy to be sufficient in combating the potential effects it might have on our results. With municipality and year fixed effects as well as general time trend, a substantial part of this development should be captured.

7 Results

From our regressions, we find that in the big cities, the fail rates in Mathematics A and English A and graduation rates increase as an effect of treatment. Graduation rates increase at a generally significant level no matter the sample restriction. Swedish A displays some treatment effect in the baseline sample set restricted to the natural and social sciences programmes. We can therefore reject our null hypothesis in these particular instances. GPA, however, appears to be unaffected by the admission reform. The full regression tables can be found in Tables D.1–5, columns 1–4, in Appendix D, whereas effects of treatment alone are briefly summarized in Table 7.1 below.

The constant, our reference group, presented in the full result tables is defined as the value for Stockholm, the first year of the sample set, 1998/99 for graduation rates and 1997/98 for fail rates and GPA. Though the results might seem small, the effects should be put in relation to the average fail and graduation rates presented in Table D.6. It depends on which year you look, but for Mathematics A and English A, this effect is comparable to an increase somewhere between 30 to 50 per cent. For the effect on graduation rates, it depends on which significant result you look at, but seeing as the graduation rates were high to begin with, the relative increase is not quite as striking.

	Baseline	**Big Cities**	**Baseline N & S**	**Big Cities N & S**
		Fail Rate in Mathematics A		
Treatment	-0.0505	2.128**	0.0628	0.239
	(0.396)	(0.830)	(0.167)	(0.271)
Observations	1,681	951	722	414
R-squared	0.408	0.419	0.247	0.213
No. of Clusters	274	177	95	62
		Fail Rate in English A		
Treatment	0.379	1.276***	-0.130	0.221
	(0.259)	(0.449)	(0.107)	(0.200)
Observations	1,677	947	721	413
R-squared	0.366	0.362	0.196	0.185
No. of Clusters	273	176	95	62
		Fail Rate in Swedish A		
Treatment	0.0281	0.564	0.172*	0.283
	(0.201)	(0.427)	(0.0969)	(0.180)
Observations	1,670	940	719	411
R-squared	0.220	0.218	0.136	0.140
No. of Clusters	273	176	95	62
		GPA		
Treatment	-0.00534	-0.0453	-0.0849	-0.0654
	(0.0713)	(0.130)	(0.101)	(0.174)
Observations	1,681	951	722	414
R-squared	0.767	0.731	0.763	0.677
No. of Clusters	274	177	95	62

	Baseline	Big Cities	Baseline N & S	Big Cities N & S
		Graduation Rate		
Treatment	2.556***	5.037***	1.780*	5.195***
	(0.896)	(1.258)	(1.024)	(1.490)
Observations	1,560	874	689	396
R-squared	0.422	0.398	0.444	0.363
No. of Clusters	256	160	90	58

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 7.1: *Treatment coefficients in the regressions run*

Instinctively, the significant results might seem counterintuitive. How can fail rates increase while graduation rates go up? The fact that we can link treatment to both an increased graduation rate and increased fail rates can be explained easily. What these results suggest is that students that otherwise would have dropped out are now staying in school. Assuming that most students who drop out do so because of difficulties in school, it should not come as a surprise that the share of students graduating with a fail in the studied subjects increases.

The results are generally not as clear as we envisaged. Albeit some significant results, they exhibit an unexpected variation across the different sample sets. Using the regression at hand, the coefficient for treatment differs between the four samples, and is of considerable different sizes. Consistency is only achieved to some degree in the regressions on graduation rates, where they are generally significant and GPA where none is.

7.1 Fail Rates

7.1.1 Mathematics A

Looking at the results for the fail rate in the course Mathematics A, the treatment variable estimate does not present any statistical significance in the majority of our sample sets. Only when we limit the sample to Sweden's three largest municipalities, Stockholm, Gothenburg, and Malmö, do we get an estimate that shows any significance. However, it does not remain significant as we narrow the sample to one with only the natural and social sciences programmes included. The significant treatment variable is to be interpreted as an average constant increase in fail rates for Mathematics A after the treatment of 2.128 percentage points for the big cities, keeping all else equal.

7.1.2 English A

As for the results in the regressions on English A, we are presented with a picture, not too dissimilar from that regarding Mathematics A. Again, only the regression on the sample of just Stockholm, Gothenburg, and Malmö, with all programmes, shows significance for the treatment variable. Treatment in English A in this setting is highly significant and increases the fail rate for the course English A with 1.280 percentage points. Though once again, the variable is anything but stable across the sample sets, it remains positive.

7.1.3 Swedish A

Swedish A breaks from the pattern of the previously presented courses in not showing a significant treatment variable in the big cities. Instead, the baseline regression limited to the natural and social sciences programmes is the only one showing any sort of significance, albeit not particularly strong.

The estimates of the constants are also considerably lower than the previous subjects, indicating that Swedish A is a course more commonly passed than the other two courses presented.

7.2 GPA

When looking at the effect on general performance in the shape of GPA, the picture is clear and consistent. None of the regressions find a statistically significant treatment variable, which causes us to reject the idea of an effect from treatment. Exactly what happens when we allow for sorting by ability is not entirely clear from this result. The only conclusion we can draw with some degree of certainty is that the average student performs just as well in either system.

7.3 Graduation Rates

Regarding the graduation rates, the regressions produce generally significant estimates across the board. The treatment is highly significant in three of the four sample sets. They indicate that treatment – that is, sorting by ability – increases the graduation rates within four years of starting considerably. It is especially high in big cities. That is where the residential segregation is likely to have been worse than in smaller municipalities.

In line with what is shown in Tables A.1 and D.6, graduation rates are generally around 70 to 80 per cent. Perhaps not a surprising feature is that the constants for the natural and social sciences sample sets are considerably higher than those containing all programmes. Worth noting is the big difference between Stockholm on the one hand, represented in our constant, and Malmö and

Gothenburg on the other. Graduation rates are considerably lower in the capital, as can be seen in Table D.4, columns 1–4, in Appendix D.

A highly significant time trend is also observable, though only in the big cities sample sets. Both have a negative coefficient estimating that for each passing year, the graduation rate within four years of starting decreases by roughly one percentage point. Students in the big cities are over time becoming less likely to graduate. This goes against the general development of graduation rates as shown in Figure 4.2.

8 Robustness Checks

8.1 The Independent Schools Hypothesis

The first hypothesis we want to test our results against is if they can be explained by an inflow of systematically different students that, absent the reform, would have gone to independent schools. One can easily make the argument that independent schools, prior to the admission reforms, had a premium attached to its organizational form in the sense that they *could* be applied to with one's GPA, whereas public schools could not. In addition, the easiest way for a talented student residing in a bad neighbourhood to escape inferior schools was to apply to independent schools located elsewhere. For example, this manifested itself in a large spread between the share of students from highly educated families on the independent schools and the corresponding share of students on the public schools among our observed municipalities in the beginning of our time period. This is illustrated in Figure 8.1. After the admission reforms, however, we argue that this premium was removed. Now all schools became available for application, which made the previous function the independent schools had as exit routes for high-skilled students obsolete. Not surprisingly, this also led to a decreasing trend in the share of students from highly educated backgrounds attending independent schools in our data set, which is neatly demonstrated by the convergence between the two curves over time below.

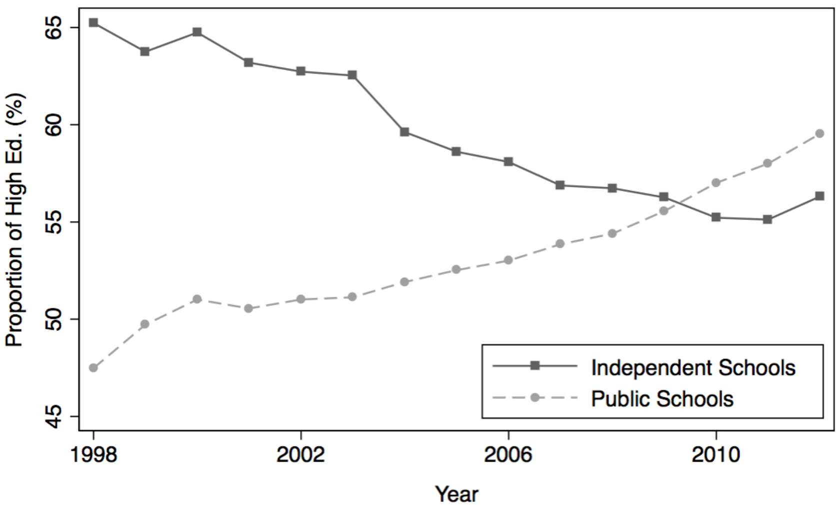


Figure 8.1: *Share of students whose parents have post-secondary education in independent schools vis-à-vis public schools in our baseline sample*

What we worry about is that our findings from the regressions reported in section 7 in reality is the outcome of such student movement between schools, rather than from increased student segregation and possibly then from peer effects. As described above, to isolate the impact of the reform as much as possible we have chosen to run our regressions with data from the public schools, since they were the only schools directly affected by the reform. We do, however, have data on the share of the students in our data set that attended public and independent schools. We can also see the demographic composition of the students in each of the two school forms.

By normalizing the time so that year 0 is equal to the time each municipality changed system, and collapsing our data set to represent the mean of all observed municipalities, we can provide a visual inspection of how the share of students attending independent versus public schools evolved in relation to the shift in admission policy. This can also be done for the demographic composition of the students. In order for our results to remain trustworthy, it would be discouraging if we saw any discontinuous movements in the share of students choosing to attend public schools or in the demographic composition of the students entering public schools during the treatment year.

Fortunately, this is not what we see. Rather, as can be seen in Figures 8.2–4, the proportion of students in our data set that are attending public schools are relatively stable, although on a declining trend. No sudden changes can be traced to the treatment year. This also holds for the demographic composition of the students attending the two different school forms. The trends are steady and no abrupt shifts occurred there either. Hence, based on our data, it does not seem that an influx of systematically different students can explain our results in the previous section.

For the entire time series development of the characteristics data by school type, see Figures E.1–3 in Appendix E.

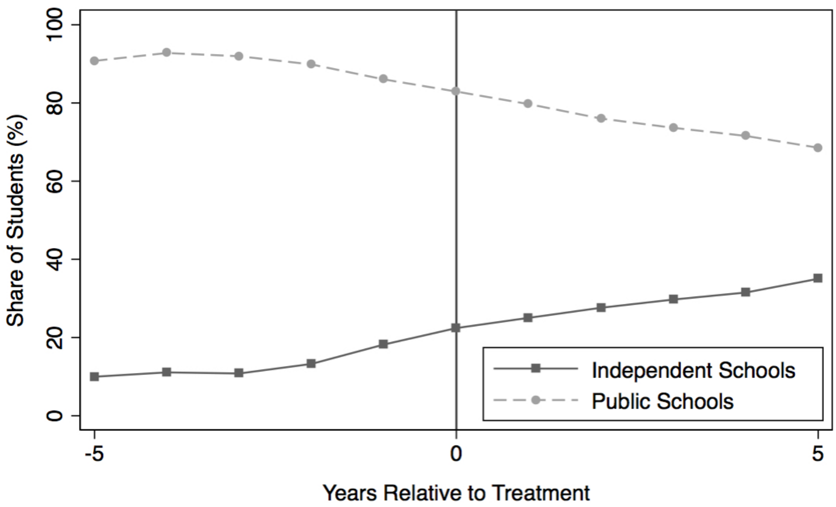


Figure 8.2: *Average share of students attending independent and public schools relative to the treatment year in our baseline sample*

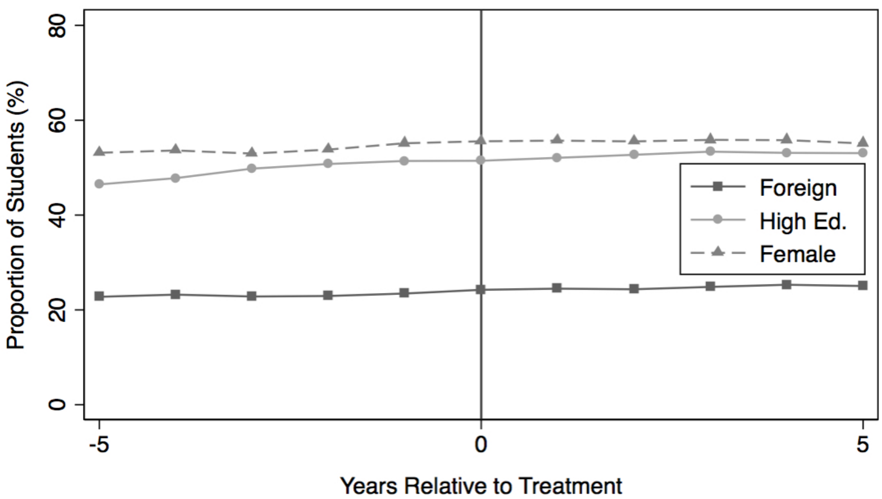


Figure 8.3: *Average demographic composition of students in public schools relative to the treatment year in our baseline sample*

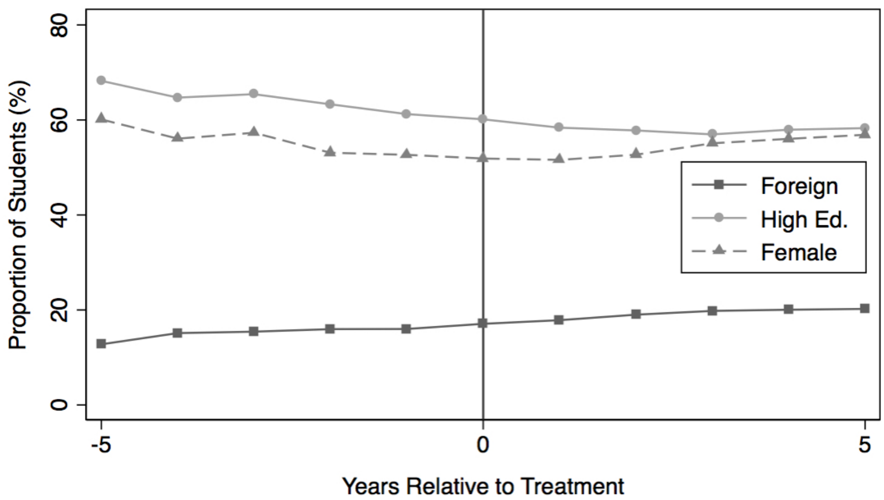


Figure 8.4: *Average demographic composition of students in independent schools relative to the treatment year in our baseline sample*

This conclusion is, however, not entirely robust. It could still be the case that public and independent schools interchanged students between each other during the treatment year, but that they were systematically different in other ways than are captured by looking at the demographic control variables included in our data set. Clearly, far from all individual characteristics are captured in whether one has a foreign background, is female or has highly educated parents. For instance, it is plausible that *ambition* as an individual trait matters greatly for educational performance and that this is not seized properly in any of the control variables previously measured. In a stylized example, it could be the case that public schools at the treatment year received the ambitious students that – without the reform – would have gone to the independent schools. In return, seeing as the number of spots in each school is limited, they might have replaced the unambitious students. Such a reshuffling of students would not necessarily show up in the visual inspection provided above.

In sum, we have indications that nothing drastically happened to the composition of students going to public schools during the treatment year, implying that our results hold in this respect. Given the limitations of our data, though, we cannot entirely reject the independent schools hypothesis.

8.2 Placebo Tests

As explained in section 6.1, the identifying assumption in our empirical strategy is that we have parallel trends between the observed municipalities. Otherwise, it is difficult to claim causality. Seeing as all our observed municipalities are treated during the time frame, but at different points in time, visual inspection of parallel trends assumption is hard to achieve.

An alternative way to examine the case for causality is to run placebo tests, in which we include a number of leads and lags of the treatment variable in the regressions. The intuition behind this is straightforward. If treatment is in fact causing the change in the outcome variable, then the leads – in which treatment is presumed to occur prior to the actual event – should be consistently insignificant.

Let k_m be the true time when $Treatment$ occurs on in municipality m. Then our placebo adjusted model is

$$Y_{imt} = \alpha_m + \lambda_t + \sum_{j=-f}^{q} \beta_j Treatment_{mt}(t = k_m + j) + \gamma\tau + +X_{imt}\delta + \varepsilon_{imt}. \qquad (3)$$

Instead of a single treatment effect, we have now introduced f leads and q lags of the treatment effect. As mentioned above, it is important that the coefficients of the leads should be insignificant. Otherwise, we could have reverse causation. Or possibly, that our results are just the outcome of a completely random process in the municipalities. The issue of whether or not the lags are significant is not as critical for this thesis since it could be the case that treatment has a lagged impact. This would for example be the case if the effect of treatment were accumulating over time, so that β_j would increase in j. In our case, such an effect could be understood as teachers choosing to leave negatively shocked schools. Conversely, in an alternative scenario, it is plausible that the impact of treatment is fading over time, as schools begin to adapt to the new student population and implement measures in order to become more effective in responding to the challenges of the new student cohort. Consequently, a lagged significant impact of treatment would not be nearly as devastating for the robustness of our results as if the leads were proven to be significant.

This allows us to formulate the following conditions in order to test the robustness of our results:

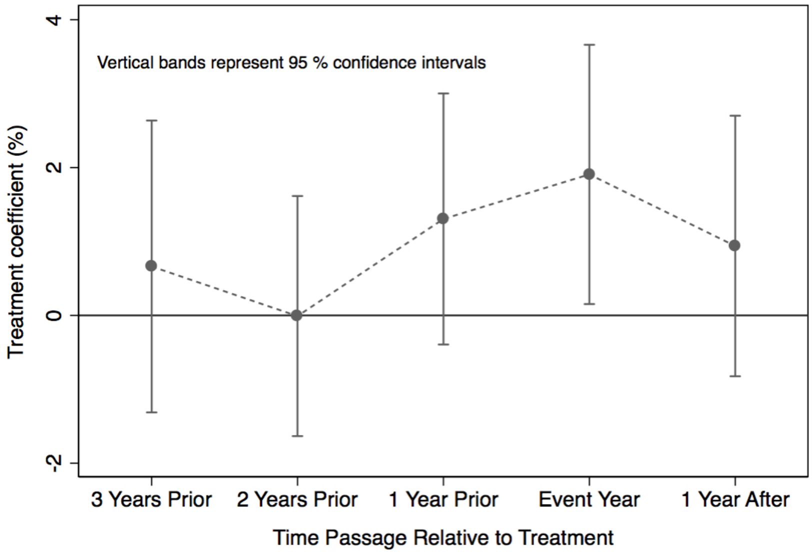
1) $\beta_j = 0 \; \forall j < 0$, and
2) $\beta_j, j \geq 0$ may not be identical.

We present leads no earlier than three years before actual treatment. Further back in time than that would result in problems in regard to our data. A 4-year lead would, for example, result in us only having one year of observations of Stockholm and Helsingborg as being untreated.

8.2.1 Results from Testing

Below we present the results from our placebo testing with regard to the sample sets that in section 7 gave significant treatment coefficients. Full results can be found in Tables E.1–2 in Appendix E.

Starting with the fail rate in Mathematics A in the big city sample set, the placebo test implies that there indeed was a significant impact at the true event year – but not before or after. Hence, our results hold up in this setting, see Figure 8.5 and Table 8.1 below.



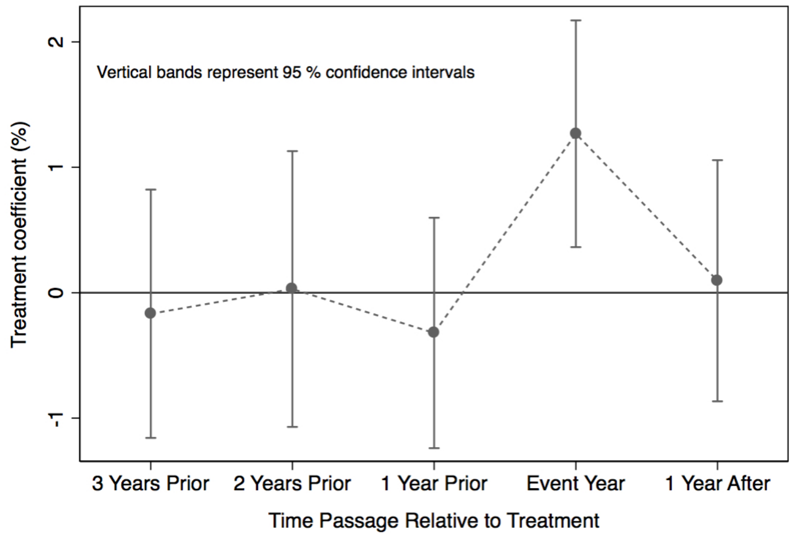
TREATMENT PLACEBO	Big Cities Math. A
3-year lead	0.662
	(1.001)
2-year lead	-0.00899
	(0.823)
1-year lead	1.305
	(0.861)
Event year	1.907**
	(0.889)
1-year lag	0.939
	(0.893)

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Figure 8.5: *Visual illustration of the treatment coefficients for fail rate in Math. A*

Table 8.1: *Numerical treatment coefficients for fail rate in Math. A*

Also when we expose our previous result with regard to the fail rate in English A in the big city sample to placebo testing, our result proves robust, see Figure 8.6 and Table 8.2 below. Once again, the only significant coefficient of treatment is the during the actual treatment year.



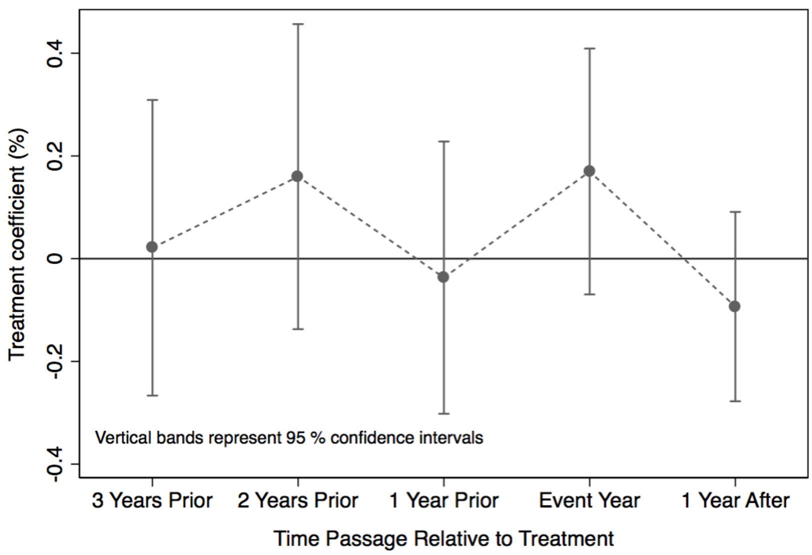
TREATMENT PLACEBO	Big Cities Eng. A
3-year lead	-0.168
	(0.502)
2-year lead	0.0295
	(0.557)
1-year lead	-0.321
	(0.465)
Event year	1.268***
	(0.458)
1-year lag	0.0955
	(0.487)

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 8.6: *Visual illustration of the treatment coefficients for fail rate in Eng. A*

Table 8.2: *Numerical treatment coefficients for fail rate in Eng. A*

When we turn to the third significant treatment variable coefficient result we obtained from our original regression, the fail rate in Swedish A on the natural and social sciences programmes in our baseline sample set, we note a peculiar result. Although all the leads are insignificant, the true treatment coefficient has now turned statistically insignificant, see Figure 8.7 and Table 8.3. Hence, some of the variation that previously was picked up in treatment vanishes when we control for leads and lags in the regression. Since the coefficient for actual treatment is insignificant now, and given that it was close to zero to begin with, we argue that we should reject our previous finding.



TREATMENT PLACEBO	Baseline N & S Swe. A
3-year lead	0.0212
	(0.145)
2-year lead	0.160
	(0.150)
1-year lead	-0.0369
	(0.133)
Event year	0.170
	(0.121)
1-year lag	-0.0933
	(0.0928)

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 8.7: *Visual illustration of the treatment coefficients for fail rate in Swe. A*

Table 8.3: *Numerical treatment coefficients for fail rate in Swe. A*

Concerning graduation rates, our previous regressions rendered statistically significant treatment coefficients in all our sample sets. When subjecting these results to the placebo tests, however, only the results from the baseline and the general big city regression can be deemed trustworthy, see Figure 8.8 and Table 8.4. In the other cases, where we only looked at the natural and social sciences programmes, we cannot reject the possibility that the previous outcome merely was resulting from a stochastic process. This is to some extent a counterintuitive finding. As mentioned before, the natural and social sciences programme are the largest programmes and are consequently among the most commonly represented programmes at schools in general. As such, they should intuitively be most affected by the change in admission policy. Therefore, it seems odd that they should fail the placebo tests. Following our formulated conditions for inferring causality, however, we must conclude that there is not sufficient empirical support to claim causality with regard to our previous findings when the sample sets are limited to the natural and social sciences programmes alone.

TREATMENT PLACEBO	Baseline Grad. Rate	Big Cities Grad. Rate	Baseline N & S Grad. Rate	Big Cities N & S Grad. Rate
3-year lead	-0.267	2.318	0.0605	3.179**
	(0.894)	(1.586)	(0.931)	(1.411)
2-year lead	0.241	2.244	-1.490*	0.390
	(0.799)	(1.358)	(0.880)	(1.292)
1-year lead	-0.0414	1.854	0.535	3.757**
	(0.735)	(1.417)	(0.785)	(1.439)
Event Year	2.361***	5.500***	2.135**	5.154***
	(0.824)	(1.328)	(0.904)	(1.341)
1-year lag	0.514	2.768*	-0.628	2.884
	(0.775)	(1.563)	(0.919)	(1.897)

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 8.4: *Treatment coefficients for graduation rate*

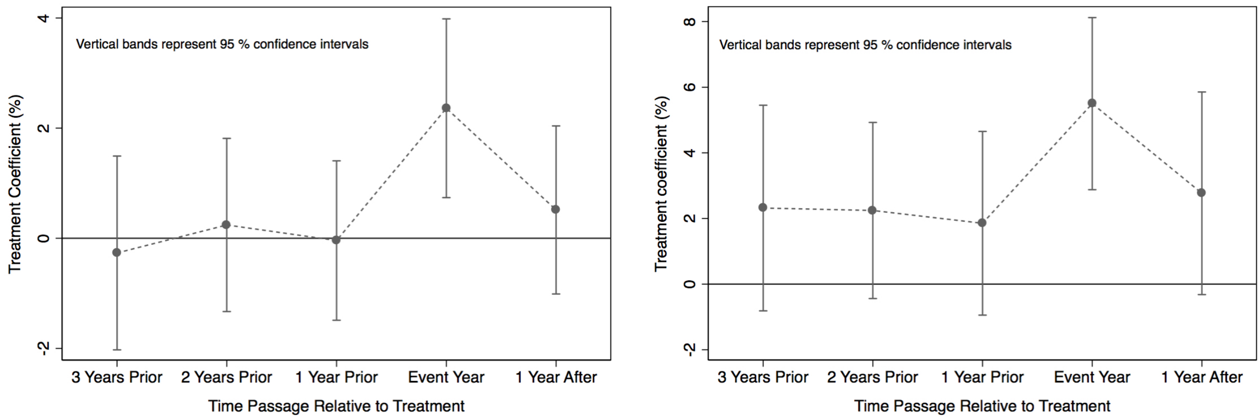


Figure 8.8: *Visual illustration of the robust treatment coefficients for graduation rate*

Panel (a): *Baseline* ***Panel (b):*** *Big Cities*

To summarize, by subjecting our previously significant results to a number of placebo tests, we conclude that the only results that credibly can be related to the change in admission policy – and hence could be associated with peer effects – are the following:

- an increase in the fail rate in Mathematics A in the big cities,
- an increase in the fail rate in English A in the big cities,
- an increase in the graduation rate in all the municipalities in our data set, and
- an even larger increase in the graduation rate in the big cities.

These results have proven to be robust and are strongly significant at the 5 per cent significance level at least. All but Mathematics A are even significant at the 1 per cent level. The careful reader might have noticed that the coefficients for the actual treatment years differ slightly in comparison to the coefficients reported in the result section. This should not come as a surprise since we altered the regression model to include lags and leads. Hence, we also obtain slightly different point estimates.

A sense of caution is however needed when interpreting the results of the placebo tests. These tests provide an indication of causality in the sense that fictitious past treatments do not seem to have mattered, whereas the actual treatments seem to have. This provides, however, little information with regard to whether treatment in fact is an exogenous or endogenous variable in relation to the outcome.

8.3 Municipality Specific Time Trends

An additional way to test the robustness of the results is to introduce municipality-specific time trends in our regressions (Angrist and Pischke 2009). This allows the different municipalities in our sample to follow different trends to a limited degree. Given that the assumption of parallel trends holds, the inclusion of municipality-specific time trends should not do anything substantial with our previous findings. Following this logic, it would be comforting if our treatment coefficients remained as they were before, whereas the robustness could be questioned if they changed dramatically.

How likely is it that the municipalities in our sample exhibit unique time trends? A benefit with our empirical approach is that the fixed differences between the municipalities are controlled for in the model. We also control for a number of time-varying demographic factors: the share of females, the share of students whose parents have post-secondary education, and the share of students with a foreign background. Still we can think of a number of time-varying trends that might be unique to different municipalities, for example the rate by which independent schools are establishing themselves in the municipality or the labour market development in the region in which the municipality is included. Hence, it is not implausible that different municipalities would exhibit different time trends.

By including municipality specific time trends, our model now becomes

$$Y_{imt} = \alpha_m + \lambda_t + \beta_1 Treatment_{mt} + \gamma\tau + \theta\tau_m + X_{imt}\delta + \varepsilon_{imt} \qquad (4)$$

where $\theta\tau_m$ denotes the municipality-specific time trend and all else remaining the same.

If $Treatment$ does not remain significant after this introduction, or changes dramatically, we should be worried about the causal implications of the change in admission policy. Then it could be equally likely that the previously observed effect was the outcome of unique municipal time trends, rather than the causal consequence of the change in admission policy.

8.3.1 Results from Testing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Big Cities	Big Cities	Baseline N & S	Baseline	Big Cities	Baseline N & S	Big Cities N & S
TREATMENT	Math. A	Eng. A	Swe. A	Grad. Rate	Grad. Rate	Grad. Rate	Grad. Rate
Original Regression Model	2.128** (0.830)	1.276*** (0.449)	0.172* (0.0969)	2.558*** (0.896)	5.037*** (1.258)	1.777* (1.023)	5.195*** (1.490)
Municipality Specific Time Trends Included	1.226 (0.828)	1.347** (0.589)	-0.0867 (0.148)	1.064 (0.987)	2.336* (1.333)	1.071 (1.107)	2.139 (1.411)

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 8.5: *Significant treatment coefficients from original regression model with and without municipality specific time trends*

The results on the previously significant treatment variable coefficients are presented in Table 8.5. For most of them we observe stark differences. For full regression results, see Tables D.1–5 in Appendix D.

Beginning with the fail rate in Mathematics A in the big cities regression, the treatment variable is no longer significant having added the municipality specific time trends. This implies that we have to reject the hypothesis that the increased fail rate was a causal response of the change in admission system. Rather, it seems that something else in the municipalities is causing the fail rate to move up.

Surprisingly however, the baseline sample set for the natural and social sciences programmes did turn from being insignificant to being slightly significant, see Table D.1, columns 3 and 7. Hence, we see indications of an effect, but seeing as it was not significant to begin with and is very close to zero, we deem the results not to be sufficiently robust to be given causal importance.

Regarding English A, the significance of the treatment variable did not change majorly having added the municipality specific time trends. The only previously significant treatment variable, found in the big cities sample set, remained significant, albeit having dropped from being significant at the 1 per cent to the 5 per cent level. No previously insignificant treatment variable turned significant by this addition, which is more in line with what we thought we would see.

The significant treatment variable for the baseline natural and social sciences sample set that we had before turned insignificant with the addition of municipality specific time trends. Not only did it turn insignificant, but the coefficient too changed from positive to negative. Swedish A is in one respect a reprise of what we saw for Mathematics A. Once again we obtained a significant treatment variable that was insignificant before, see Table D.3, columns 4 and 8. By the same logic applied above, we are hesitant to accept this as a trustworthy result.

Out of three, only one considerably weaker significant coefficient remained. For the previously so significant treatment on the graduation rates we do however see results in line with what we saw in the placebo test. The big cities sample set still points in the direction of some treatment effect albeit not as significant as before.

To summarize, introducing municipality specific time trends resulted in the following:

- Mathematics A or Swedish A fail rates experience no significant treatment.
- For English A, treatment remains significant in the big cities.
- Treatment on graduation rate remains slightly significant in the big cities.

9 Discussion

9.1 Identification Problem

As our results stand after subjecting them to robustness checks, our empirical strategy has been successful in credibly isolating a causal policy effect and rejecting the null hypothesis in two cases:

1) increasing fail rates in English A in the big cities, and
2) increasing graduation rates in the big cities.

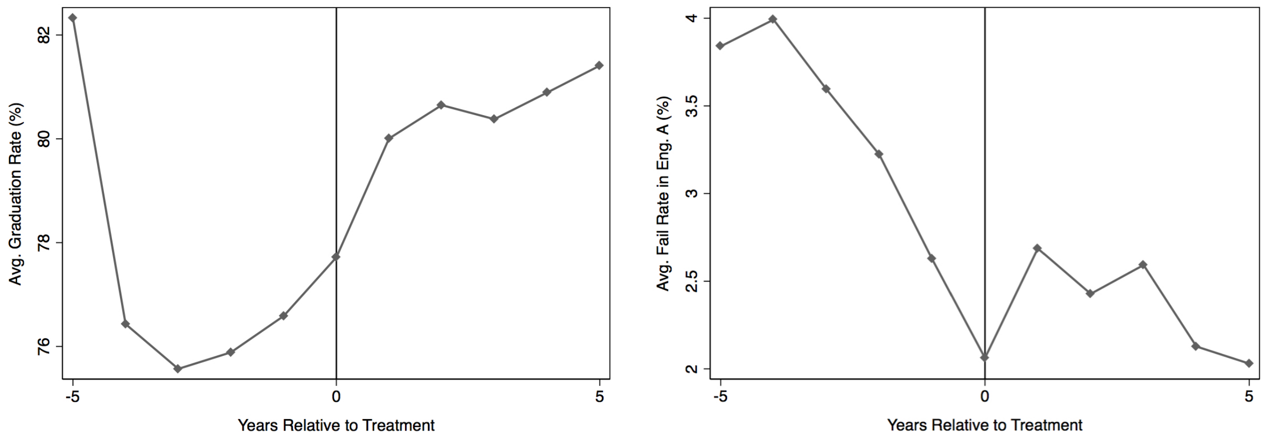


Figure 9.1: *Graphical illustration of key results*

Panel (a): *Graduation Rate* ***Panel (b):*** *Fail Rate in Eng. A*

As described above, one could easily be misled to think that these results are contradictory to each other. However, what these results suggest is that low-skilled students complete their education to higher degree in the sense that they receive grades in all courses. Given that these are students who, absent the reform, would have terminated their education early and are likely to be the least skilled students, it should not come as surprise that the fail rate goes up after the reform. Rather, what these results seem to indicate is that low-skilled students exhibited a higher degree of perseverance after the reform – even though they may have continued to perform badly, they do not seem to have given up as easily as they did before.

These effects are statistically significant and, given the high number of students that enter upper secondary schools each year, they imply a relatively strong impact in absolute terms. Moreover, the students who, prior to the reforms, tried to circumvent the admission system by, for example, registered themselves on a relative's address in order to get in to certain schools, are likely to cause a downward bias on these effects. The results should, however, be interpreted in light of the 18 other times our estimates were dismissed due to insignificance or failed causality tests.

Given the magnitude of our robustness checks, we can nonetheless say – to the best of our knowledge – that the change in admission policies in the big cities seems to have had a causal impact on graduation rate and the fail rate in English A. The crucial question in this thesis, however, is whether these effects credibly can be related to peer effects.

We can think of a number of reasons why this would and would not be the case. The sudden introduction of the reform lends support to that it actually is peer effects that are causing the impact of the policy shift. It implies that nothing substantial but student allocation had time to change systematically in our model. Given that we limit the time span to have an upper limit of one year after the reform in our regressions, we protect ourselves against the eventuality of increased teacher mobility and similar responses to the post-treatment student allocation (Karbownik 2014). Teachers may become more or less effective in the classroom situation depending on the classroom composition, but this would easily fit the notion of indirect peer effects described by Tincani (2015).

What makes it difficult to draw any definitive conclusions in regard to whether peer effects in fact are the underlying cause of these policy effects is all the strong assumptions that we had to impose in section 6. Apart from parallel trends, our identification strategy hinges upon the assumption that the municipalities in our data set did not change anything else besides the admission system when going through with this reform. This is a difficult assumption to test since it would require very detailed information on how every municipality implemented this reform, which given the scope and time frame of this thesis was difficult to attain. It is not entirely implausible, though, that municipalities foresaw increased ability segregation in light of the admission reform, and therefore decided to take action to help the most negatively shocked schools. To the extent this happened we can only speculate, but future studies should attempt to get a more comprehensive understanding of how the municipalities implemented the changes in admission system.

In addition, we also had to impose strict exogeneity in the sense that the decision to change admission system was not correlated with any past, present or future fluctuations of the error term in our models. In practice, this implies that the decision to shift admission system should not have been a reaction to unusually bad outcomes in previous periods. If that were the case, we could be led to believe that there was a positive effect of the admission reform when things in fact were normalizing. This is also a strong assumption that could have been violated.

However, the benefit with our empirical strategy, in which we have different treatment and control groups in different time periods, is that we to some extent have a built-in firewall against sporadic violations of these assumptions. In order for our results to be seriously damaged from breaking these assumptions, they have to be systematically breached in all or most of our observed municipalities. From our state of knowledge, it is difficult to provide an assessment of whether or not this has been the case, but we find confidence in the fact that the likelihood of systematic violations of these assumptions is smaller when we analyse many municipalities rather than just a few. From this perspective it is, however, slightly worrying that the only results that have been verified robust are those from the big city sample set, where only three municipalities are included. In this setting, we cannot argue with the same amount of confidence that our results should not be biased in any substantial way from breaking the underlying assumptions.

To answer our question posed in the beginning of this section, can we credibly say that the observed treatment effect in fact is related to peer effects? The answer is not evident. Given our broad definition of direct and indirect peer effects as well as our adopted model limitations, we argue that we have narrowed down the possible alternative causes to a point where it is difficult to see any other more plausible explanation than peer effects. Then again, it hinges upon that our regressions exhibit no systematic violations of the underlying assumptions. For future studies, better data and more information should be desirable to give a more certain comprehensive answer to this question.

9.2 Selection Bias

Ideally in an event study of this kind, the treatment group should be randomly assigned in order to ascertain that there are no systematic differences between the municipalities in our sample and all other Swedish municipalities not included. Otherwise, there could be a selection into treatment, which could threaten the external validity of our results.

It is evident that this ‘experimental ideal’ is not being reached in our thesis. For one thing, the decision to change admission system was in reality far from being exogenously given, as would have been the case if the reform were imposed by the state, for instance. Rather, the change in admission policy was the outcome of a highly endogenous political process responding to the time-variant and time-invariant issues facing each municipality. Although our results may hold from a policy analysis point of view in regard to the seven municipalities observed, we cannot be certain that these results are transferable to other geographic settings.

We can think of two ways in which our observed municipalities may have been systematically different from other municipalities. First, it is likely that residentially segregated municipalities were more likely to change admission policy in the first place. From previous work (Karbownik 2014), we know that the reason to abandon the proximity-based admission system was often connected to an attempt to alleviate residential segregation. Hence, the political pressure for change is likely to have been greater in municipalities that exhibited a high degree of segregation.

Second, municipalities that exhibited a high number of independent schools also had stronger incentives to change admission system. As described above, prior to the change in admission policy, independent schools had a premium over public schools in the same municipality: they *could* be chosen, whereas public schools could not. If this premium manifested itself in a higher share of students deciding to go to independent schools, municipalities had an economic incentive to equalize the operational terms between the two school forms. Consequently, it is likely that the municipalities in our sample exhibited a higher degree of independent schools in comparison to other municipalities.

This raises questions regarding to which extent we can claim external validity in our thesis. Are we measuring the impact of peer effects in general – or are we measuring the impact of peer effects in segregated environments with a high number of independent schools? This is a difficult question to answer. When communicating with local education officials in writing this thesis, many confirmed that the reason for shifting admission policy broadly was in line with either one or the other of the accounts presented above. This lends support to the criticism that our results might have a selection bias. In what direction and how much is difficult to say, but seeing as this is a threat to the external validity of our results, future research should try to overcome this dilemma by, for example, including the full set of municipalities.

Another point to add is our choice to exclude the municipalities big enough to enact an admission reform but that did not do so within the time frame of the Lpf 94 curriculum, hampering comparability of variables. By excluding them altogether, even as a control group, we run the risk of false positives in our baseline regression runs. However, seeing as all the significant and robust results we had came from the big cities sample set, it is not a bias that has affected our final conclusions.

9.3 Big City Phenomenon

Why do the results remain significant only in the big cities sample set? This is a valid question that goes back to our understanding of what happens in connection to the admission reform at the centre of this thesis. We assume findings from the Stockholm transition, Söderström and Uusitalo (2010) and Karbownik (2014), to remain valid for all other municipalities. Generalizing those results and claiming external validity could be questioned. We argue that the basic functions, increasing segregation by ability and the effects on teacher separation rates, should hold true no matter the municipality. The results, however, lend support to the view that the effects might not be quite as big when the pool of students and teachers and the number of schools grow much smaller than in the case of the big cities.

That this would be the case is logical. A larger group of students leads to the need to have more schools to offer students the programme of their choice. With a high number of schools, not only students, but teachers too, have a wide-ranging choice of which school to attend or work at. If students, when able to apply to a specific school, choose in a way that creates homogenous schools and classrooms, we will see a sorting of a kind that looks vastly different in different municipalities. If the number of schools offering the same programme in this time frame was three or eighteen, as for the case of the natural sciences programme in Linköping and Stockholm respectively, the options are far from similar. This holds, assuming that the student ability distribution is equal across municipalities. In addition, dissatisfaction with your place of work in a big city allows a teacher greater choice than would otherwise be the case in a smaller community.

Though we do not believe that the basic workings should be different depending on municipality size, the effect that an admission reform of this kind has on student sorting is likely not quite as big in smaller municipalities.

9.4 Aggregate Data

The general issue of working with aggregate data within a data sensitive subject such as the identification of peer effects is the fact that we cannot look at the individual level. It limits us from looking at cohorts of students and instead having to analyse entire groups through a mean, more often than not, representing a wide variety of ability and effort. We are limited to look at peer effects through a measure of means, barring us from understanding more intricate levels of behaviour.

As mentioned before, two of our three types of dependent variables help us mainly to understand the impact of peer effects on students in the lower part of the skill distribution. The students in risk of failing courses or not graduating are likely to be the least skilled students. This means that we, with a relatively high degree of certainty, can tell how segregation by ability struck them. It would, however, be equally important to look at what happened to the highest skilled students who instead were treated with the absence of low-performing students after the change in admission policies, in order to obtain a more comprehensive understanding of peer effects. Did they increase their performance even more – or did something else happen? Given the aggregated nature of the data at hand, we cannot know for certain. We only have indications that nothing seems to have happened with average GPA in our sample sets after the admission reforms, but that leaves out a number of interesting movement scenarios *within* the student group as whole. This is obviously an inevitable weakness in our thesis, which future studies should try to resolve by for example running our regression model with longitudinal, registry data.

Another bias that is necessary to address is the lag in reporting the fail rates. As fail rates are not presented in SIRIS until the final school leaving certificates are issued, we cannot know with certainty when and how students achieved a pass grade. This element leaves open the opportunity for failing students to study harder and get a pass in a previous course before leaving school, as some schools allowed for. If anything, this feature would understate the actual effect, strengthening the significant results we have found, and to a degree explaining some of the insignificant results.

9.5 The Problems of Not Measuring Knowledge

A fundamental flaw, which is hard to get around when writing a report on peer effects, is the fact that there is no truly objective way to measure knowledge. From a human capital development perspective, what we are interested in is maximizing each individual's educational potential. In this sense, it is unfortunate that we have to rely on proxies to capture the underlying variable of interest – knowledge. Naturally, GPA, fail rates and graduation rates are inherently linked to innate ability and knowledge, but the correlation is not complete.

Given how the educational sector changed and developed within this period, shifting into a market with considerable competition, we also have reason to suspect that the knowledge proxies were affected, for example through grade inflation. Seeing as we have included a time trend in our regressions to cope with the matter of grade inflation, this should not be of any major concerns in our results. However, one could easily argue that the incentives to inflate grades increased even more as the operational terms between public and independent schools were equalized. Seeing as schools to a larger extent now had to compete in order to attract students, one way to overcome information asymmetries in contact with potential students would be to signal quality by the means of inflating grades. In addition, such an inflationary behaviour is more likely, the higher the number of schools is that were operating in each municipality (Vlachos 2010). Maybe this is what is being captured in the municipality specific time trends? We cannot know, but it is not entirely unlikely.

Given the weaknesses of GPA, fail rates and graduation rates as knowledge proxies, a better dependent variable would have been the results of a standardized test that would remain comparable over time and equivalent across Sweden. If all students had taken for example a test similar to the SAT, then issues of reliability of our proxies would have been minimized. Using the Swedish National Tests that exist would have been optimal, had they been constructed to be comparable over time.

10 Conclusion

To understand peer effects is a vital part in understanding the formation of human capital. If we can optimize classroom organization, there will be considerable macroeconomic gains to reap in the future.

In this thesis, we have conducted an event study in which we have analysed how GPA, fail rates and graduation rates evolved as the seven largest municipalities of Sweden changed admission system to their public upper secondary schools. Seeing as this is likely to have led to a reallocation of students, where high-skilled students were pooled together on certain schools and low-skilled students on others, while holding all other essential parts of the school system fixed, we test the hypothesis that these shifts in admission policy have caused our outcome variables – fail rates, GPA and graduation rates – to change due to a redistribution of peer effects in the local school systems.

Our findings are intriguing. After the reform, more students were graduating than before, but with higher fail rates in courses. Although appearing as a contradiction, this implies that the students who, absent of the reform, should have ended their education early, now endure and complete their education – with the consequence of higher fail rates on the aggregate level. Given that those in risk of not graduating are likely to be the least skilled students, this finding indicates that the change in admission policy had a positive impact on the perseverance of such students. Even though they may have been performing badly – they do not seem to have given up as easily as before.

After subjecting our results to a series of strict robustness tests, these results hold for the three largest municipalities in Sweden – Stockholm, Gothenburg and Malmö – when we look specifically at graduation rate and the fail rate in English A. This should not come as a surprise, given that ability sorting in light of the admission reform is likely to have been greatest in the largest municipalities.

Are these results related to peer effects? We have indications that point in that direction, but more information and better data is likely needed in order to settle this matter without hesitation. We have defined peer effects openly by not limiting them to direct peer-to-peer interaction, but also including teacher adjustments due to the composition of students, so called indirect peer effects (Tincani 2015). By accounting for known changes in teacher mobility, for student characteristics, time trends, year and municipality fixed effects, we have narrowed down the possible causes for change to a point where it is difficult to see any more plausible explanation than that of peer effects.

10.1 Implications for Policy and Classroom Organization

Our results indicate that sorting by ability in the Swedish upper secondary school system was good for students in the lowest part of the skill distribution, in the sense that they increased their graduation rate. Although our data limits us to scrutinize how outcomes evolved for other student cohorts, one should not neglect the considerable economic value of more students graduating.

For one thing, a critical determinant of labour market outcomes in Sweden seems to be whether or not one has a school leaving certificate from the upper secondary school or not. For instance, the employment rate for those who only have primary education is almost half of that in the total population (SCB 2016). Moreover, a series of negative health outcomes are also correlated with not having post-secondary education, for example the risk of dying prematurely (Folkhälsomydigheten 2016). In other words, it is crucial for the well-being of a society that students actually graduate.

This raises questions on how we should organize the composition of students in a class. Without being able to examine the impact ability sorting has on other student groups, it would be unwise to express any far-reaching policy implications. We do not deem our results strong enough to claim external validity across different settings – be they geographic, cultural or time-variant. Nonetheless, our results still indicate that there are gains to be made from trying to optimize group constellations. Specifically, for the least skilled part of the student distribution in Sweden, a certain degree of ability sorting, stemming from a grade-based admission system, seems to have been beneficial.

From an academic point of view, our results can possibly lend merit to the idea that peer effects work through rank concerns, rather than linearly through the mean ability in a group. In a group with a homogenous ability distribution, fewer members will be too far behind the other group members to give up. Although far from established in this thesis, our results suggest that future research should continue to explore this emerging path in terms of how peer effects work.

References

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press. Princeton, NJ.

Bénabou, Roland. 1996. 'Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance'. *American Economic Review*, 86(3): 584-609.

Bertrand, Marianne; Esther Duflo and Sendhil Mullainathan. 2004. 'How Much Should We Trust Difference-In-Differences Estimates?' *Quarterly Journal of Economics.* 119(1): 249–275.

Bhattacharya, Debopam. 2009. 'Inferring Optimal Peer Assignment From Experimental Data'. *Journal of the American Statistical Association*, 104(486): 486-500.

Carrell, Scott E.; Bruce Sacerdote and James E. West. 2013. 'From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation'. *Econometrica*, 81(39): 855-882.

Falk, Armin and Andrea Ichino. 2006. 'Clean Evidence on Peer Effects'. *Journal of Labor Economics.* 24(1): 39-57.

Folkhälsomyndigheten. 2016. *Folkhälsan I Sverige 2016.* Annual reports. Available for download: https://www.folkhalsomyndigheten.se/pagefiles/23257/Folkhalsan-i-Sverige-2016-16005.pdf (15 May 2016)

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'. IFAU – Institute for Evaluation of Labour Market and Education Policy Report 2014:25.

Horton. John J. and Richard J. Zeckhauser. 2016. 'The Causes of Peer Effects in Production: Evidence from a Series of Field Experiments'. Working Paper. Available for download at: http://john-joseph-horton.com/papers/peer_effects.pdf (5 May 2016)

Hoxby, Caroline. 2000. 'Peer Effects in the Classroom: Learning from Gender and Race Variation'. National Bureau of Economic Research Working Paper 7867.

Joensen, Juanna Schrøter and Helena Skyt Nielsen. 2015. 'Spillovers in Educational Choice'. Working Paper. Available for download at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2548702 (12 February 2016)

Kandel, Eugene and Edward P. Lazear. 1992. 'Peer Pressure and Partnerships'. *Journal of Political Economy.* 100(4): 801-817.

Karbownik, Krzysztof. 2014. 'Do changes in student quality affect teacher mobility? Evidence from an admission reform'. IFAU – Institute for Evaluation of Labour Market and Education Policy Working Paper 2014:15.

Kremer, Michael. 1993. 'The O-Ring Theory of Economic Development'. *Quarterly Journal of Economics*, 108(3): 551-575.

Manski, Charles F. 1993. 'Identification of Endogenous Social Effects: The Reflection Problem'. *Review of Economic Studies*, 60(3): 531-542.

Mas, Alexandre and Enrico Moretti. 2009. 'Peers at Work'. *American Economic Review.* 94(3): 656-690.

Pischke, Jörn-Steffen. 2007. *Empirical Methods in Applied Economics.* Lecture Notes. October. Available for download at: http://econ.lse.ac.uk/staff/spischke/ec524/evaluation4_07.pdf (1 May 2016).

Sacerdote, Bruce. 2001. 'Peer Effects with Random Assignment: Results for Dartmouth Roommates'. *Quarterly Journal of Economics*, 116(2): 681-704.

Sacerdote, Bruce. 2014. 'Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?' *Annual Review of Economics.* 6(1): 253-272.

SCB. 2016. *Statistikdatabasen.* Befolkningen 15–74 år (AKU) efter arbetskraftstillhörighet, utbildningsnivå och kön. År 2005–2015. Available for download at: http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__AM__AM0401__AM0401P/NAKUBefUtbNivAr/?rxid=bc5c0125-0e0b-44f7-8ac3-7b6bfac6149b# (10 May 2016)

Skolverket. 2012. *Betygsinflation – betygen och den faktiska kunskapsutvecklingen.* PM. Dnr 2012:387. Available for download at: http://www.skolverket.se/publikationer?id=2804 (1 May 2016).

Skolverket. 2015a. *Slutbetyg i grundskolan, våren 2015.* PM. Dnr 5.1.1–2015:1103. Available for download at: http://www.skolverket.se/publikationer?id=3528 (30 April 2016).

Skolverket. 2015b. *Beskrivande data 2014 – förskola, skola och vuxenutbildning.* Rapport 420. Available for download at: http://www.skolverket.se/publikationer?id=3418 (30 April 2016).

Skolverket. 2016a. *Jämförelsetal.* Database. Available for download at: http://www.jmftal.artisan.se/databas.aspx#tab-0 (5 May 2016).

Skolverket. 2016b. *Uppföljning av gymnasieskolan.* Regeringsuppdrag – uppföljning och analys av gymnasieskolan. Available for download at: http://www.skolverket.se/publikationer?id=3642 (4 May 2016).

Skolverket. 2016c. *Skolverkets Internetbaserade Resultat- och Kvalitetsinformationssystem (SIRIS).* Analysstöd. View at: http://siris.skolverket.se/siris/ris.get_image?pin_dok_namn=genom45ar07 (4 May 2016).

Skolverket. 2016d. *Skolverkets Internetbaserade Resultat- och Kvalitetsinformationssystem (SIRIS).* Database. Available for download at: http://siris.skolverket.se/siris/f?p=SIRIS:62:0::NO::: (4 May 2016).

Söderström, Martin and Rope Uusitalo. 2010. 'School Choice and Segregation: Evidence from an Admission Reform'. *Scandinavian Journal of Economics*, 112 (1): 55-76.

Tincani, Michela M. 2015. 'Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence'. Working Paper. Available for download at: http://www.homepages.ucl.ac.uk/~uctpmt1/Tincani_jmp.pdf (5 May 2016).

Vlachos, Jonas. 2010. 'Betygets värde'. Konkurrensverket, Stockholm. Uppdragsforkningsrapport 2010:6. Available for download at: http://www.konkurrensverket.se/globalassets/aktuellt/nyheter/betygets-varde.pdf (10 May 2016).

APPENDICES

A Institutional Background

GY 11	Grad. Year	Fail Rate in Math. 1 (%)	Fail Rate in Eng. 5 (%)	Fail Rate in Swe. 1 (%)	Graduation Rate within Four Years (%)	GPA (0–20)
	2015	2.3	1.5	1	78	14.0
	2014	3.6	1.7	1	–	14.0
LPF 94	Grad. Year	Fail Rate in Math. A (%)	Fail Rate in Eng. A (%)	Fail Rate in Swe. A (%)	Graduation Rate within Four Years (%)	GPA (0–20)
	2014	–	–	–	76	–
	2013	3.3	2.2	1.2	77	14.0
	2012	3.6	2.1	1.2	77	14.0
	2011	3.1	2.1	1.1	76	14.1
	2010	3.2	2.3	1.2	76	14.0
	2009	3.1	2.3	1.1	76	14.1
	2008	3.7	2.8	1.2	76	14.0
	2007	3.8	2.5	1.1	75	14.1
	2006	3.8	2.6	1.1	75	14.1

Table A.1: *Average GPA, fail rates in selected courses and overall graduation rate, Sweden 2006–2015 (Skolverket 2016b, Skolverket 2016d)*

Proximity	Randomized	Hybrid
Gothenburg	Helsingborg	*Gothenburg*
Jönköping	Linköping	
Lund	*Norrköping*	
Malmö		
Norrköping		
Stockholm		
Umeå		
Uppsala		
Västerås		

Table A.2: *Admission systems prior to school choice*[1]

[1] Note that the municipalities typed in italics have implemented more than one change in admission system.

B Data

	First Academic Year with Grade Based Admission Policy	First observation in Fail Rate statistics	First Observation in Four-Year Graduation Rates Statistics	Curriculum in which Admission Reform Was Implemented
Stockholm	2000/01	2002/03	2003/04	Lpf 94
Gothenburg	2002/03	2004/05	2004/06	Lpf 94
Malmö	2001/02	2003/04	2004/05	Lpf 94
Uppsala	2008/09	2010/11	2011/12	Lpf 94
Linköping	2007/08	2009/10	2010/11	Lpf 94
Västerås	2011/12	2013/14	2014/15	Gy 11
Örebro	2011/12	2013/14	2014/14	Gy 11
Helsingborg	2000/01	2002/03	2014/15	Lpf 94
Norrköping	2011/12	2013/14	2014/15	Gy 11
Jönköping	2011/12	2013/14	2014/15	Gy 11
Umeå	2011/12	2013/14	2014/15	Gy 11
Lund	2001/02	2003/04	2004/05	Lpf 94

Table B.1: *Admission policy reform year and observations in time series*

	Math. A	Eng. A	Swe. A	GPA	Stud. Grad.	Grad. Rate	Stud. Start.	Female	Foreign	High Ed.
Math. A	1									
Eng. A	0.630	1								
Swe. A	0.487	0.475	1							
GPA	-0.607	-0.567	-0.445	1						
Stud. Grad.	-0.256	-0.261	-0.176	0.319	1					
Grad. Rate	-0.408	-0.386	-0.295	0.529	0.357	1				
Stud. Start.	-0.244	-0.254	-0.160	0.292	0.777	0.252	1			
Female	0.061	0.044	-0.024	0.067	-0.125	-0.100	-0.13	1		
Foreign	0.076	0.148	0.060	-0.217	-0.237	-0.248	-0.119	-0.018	1	
High Ed.	-0.578	-0.552	-0.382	0.833	0.418	0.553	0.400	-0.141	-0.288	1

Table B.2: *Correlation matrix for variables used*

In order to account for collinearity, we constructed this correlation matrix. For variables appearing in the same regression, the only considerable exception to the rule of low correlation is the correlation between High Ed. and GPA.

C Empirical Strategy

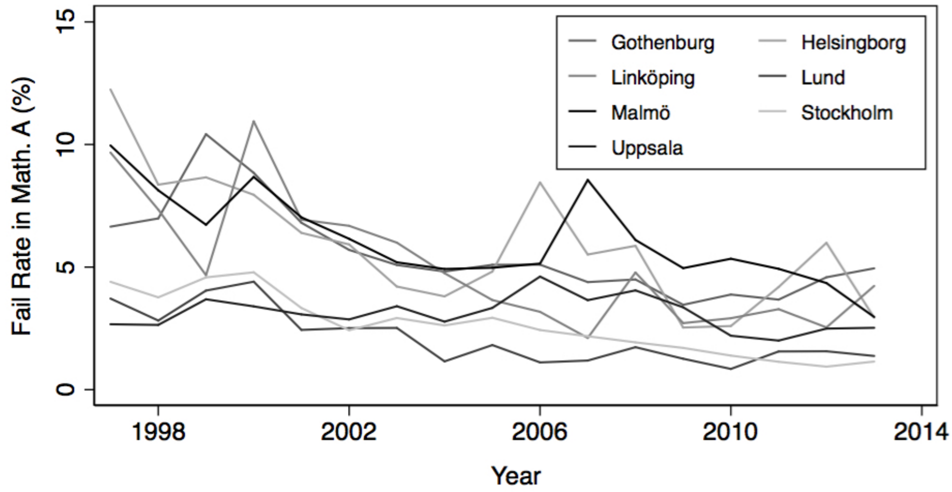


Figure C.1: *Average fail rate in Mathematics A over time*

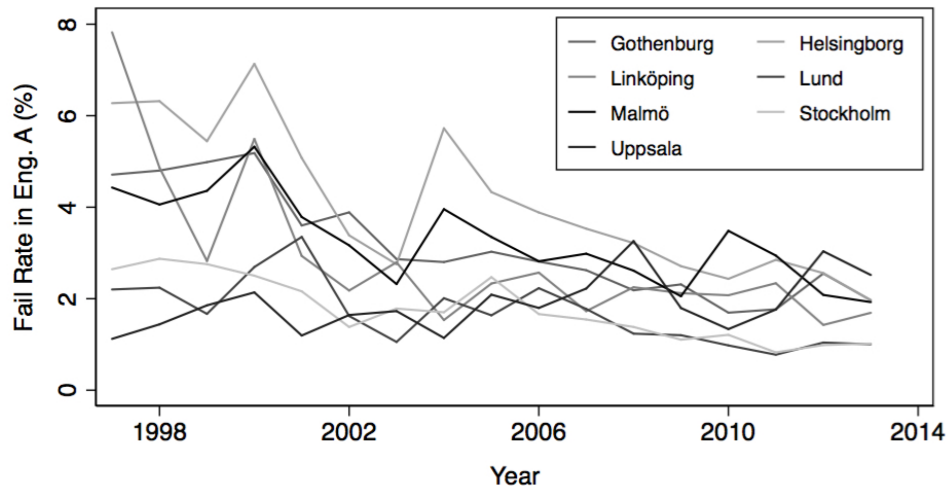


Figure C.2: *Average fail rate in English A over time*

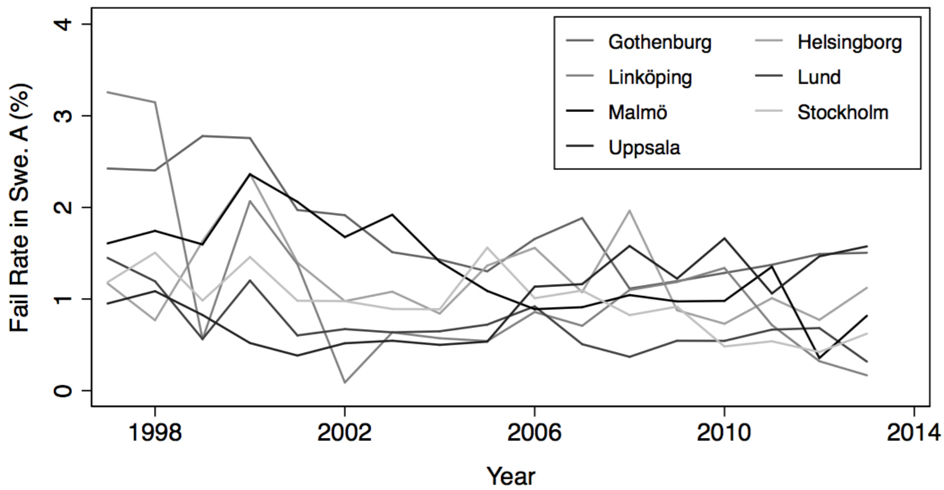


Figure C.3: *Average fail rate in Swedish A over time*

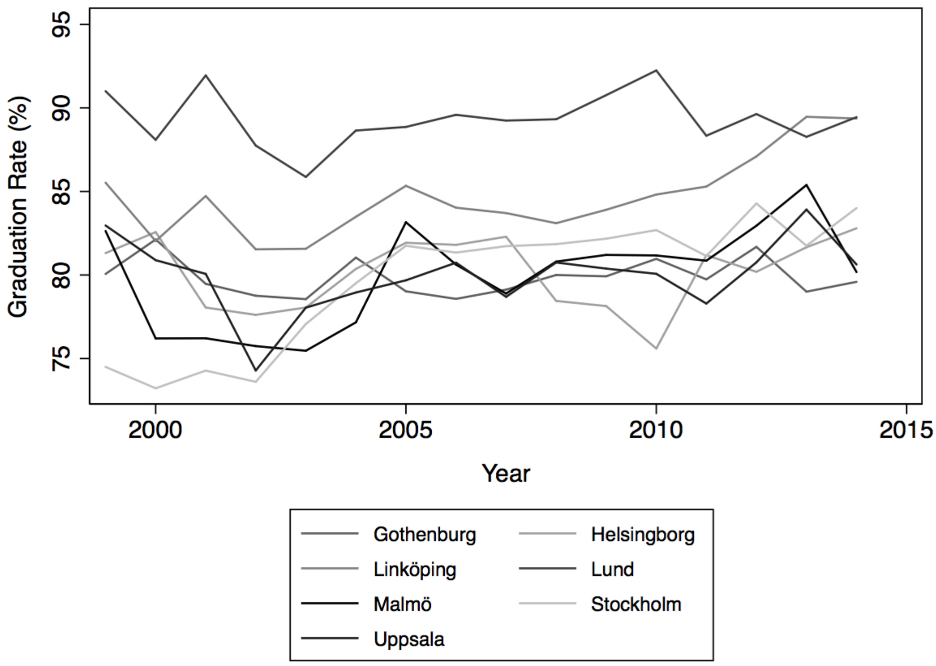


Figure C.4: *Average graduation rate over time*

D Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Big Cities	Baseline	Big Cities	Baseline	Big Cities	Baseline	Big Cities
			N & S	N & S			N & S	N & S
VARIABLES	Math. A	Math. A	Math. A	Math. A	Math. A	Math. A	Math. A	Math. A
Treatment	-0.0505	2.128**	0.0628	0.239	0.369	1.226	0.383*	0.287
	(0.396)	(0.830)	(0.167)	(0.271)	(0.483)	(0.828)	(0.220)	(0.250)
Avg. Female 3	0.00864	0.000948	0.0167***	0.0179***	0.00820	0.000573	0.0172***	0.0178***
	(0.0176)	(0.0295)	(0.00389)	(0.00606)	(0.0177)	(0.0295)	(0.00390)	(0.00601)
Avg. Foreign 3	-0.0751***	-0.118***	-0.0153*	-0.0112	-0.0748***	-0.119***	-0.0137*	-0.0100
	(0.0277)	(0.0358)	(0.00804)	(0.00936)	(0.0277)	(0.0360)	(0.00799)	(0.00927)
Avg. High Ed. 3	-0.243***	-0.311***	-0.0531***	-0.0513***	-0.243***	-0.311***	-0.0521***	-0.0505***
	(0.0186)	(0.0339)	(0.00647)	(0.00977)	(0.0186)	(0.0340)	(0.00640)	(0.00958)
Gothenburg	0.240	0.187	-0.337	-0.352*	1.671	0.949	-0.570**	-0.487*
	(0.926)	(1.090)	(0.215)	(0.204)	(1.487)	(1.493)	(0.273)	(0.277)
Malmö	-0.207	-0.429	-0.485**	-0.463*	0.762	0.0998	-0.0513	-0.0443
	(1.148)	(1.150)	(0.227)	(0.232)	(1.684)	(1.639)	(0.422)	(0.428)
Helsingborg	-2.388**		-0.541**		-0.935		-0.144	
	(1.149)		(0.230)		(1.506)		(0.313)	
Linköping	-0.606		0.122		0.439		0.253	
	(0.870)		(0.181)		(1.182)		(0.312)	
Lund	0.778		0.258		1.674		0.329	
	(0.842)		(0.167)		(1.167)		(0.230)	
Uppsala	-1.028		-0.000515		-1.632		-0.360	
	(0.841)		(0.202)		(1.064)		(0.247)	
τ	0.0896	-0.332*	-0.00646	0.0950	0.0927	0.0719	-0.129	0.0678
	(0.0714)	(0.193)	(0.0212)	(0.0905)	(0.207)	(0.297)	(0.0837)	(0.0941)
Constant	18.63***	23.16***	3.611***	3.386***	17.94***	22.83***	3.470***	3.267***
	(2.083)	(3.511)	(0.611)	(0.875)	(2.098)	(3.522)	(0.610)	(0.860)
Municipality Specific Time Trends	NO	NO	NO	NO	YES	YES	YES	YES
Year Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,681	951	722	414	1,681	951	722	414
R-squared	0.408	0.419	0.247	0.213	0.412	0.419	0.264	0.222
No. of Clusters	274	177	95	62	274	177	95	62

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table D.1: *Regression results on fail rates in Mathematics A from runs with and without municipality specific time trends*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Big Cities	Baseline	Big Cities	Baseline	Big Cities	Baseline	Big Cities
			N & S	N & S			N & S	N & S
VARIABLES	Eng. A	Eng. A	Eng. A	Eng. A	Eng. A	Eng. A	Eng. A	Eng. A
Treatment	0.379	1.276***	-0.130	0.221	0.381	1.347**	-0.223	0.253
	(0.259)	(0.449)	(0.107)	(0.200)	(0.320)	(0.589)	(0.160)	(0.299)
Avg. Female 3	-0.0139	-0.0172	0.000678	-0.00308	-0.0140	-0.0172	0.000782	-0.00305
	(0.0110)	(0.0201)	(0.00342)	(0.00538)	(0.0111)	(0.0202)	(0.00349)	(0.00547)
Avg. Foreign 3	0.000365	-0.00488	0.0137**	0.0133*	0.000554	-0.00503	0.0135**	0.0129*
	(0.0164)	(0.0197)	(0.00597)	(0.00744)	(0.0167)	(0.0201)	(0.00590)	(0.00734)
Avg. High Ed. 3	-0.141***	-0.157***	-0.0244***	-0.0266***	-0.141***	-0.157***	-0.0244***	-0.0269***
	(0.0110)	(0.0189)	(0.00497)	(0.00806)	(0.0110)	(0.0191)	(0.00500)	(0.00811)
Gothenburg	-0.157	-0.0173	-0.333**	-0.258	0.381	-0.0676	-0.164	-0.255
	(0.523)	(0.580)	(0.162)	(0.199)	(0.799)	(0.820)	(0.266)	(0.289)
Malmö	-0.721	-0.712	-0.214	-0.187	-0.551	-0.889	-0.249	-0.333
	(0.542)	(0.580)	(0.198)	(0.212)	(0.903)	(0.923)	(0.287)	(0.315)
Helsingborg	-0.467		-0.0980		-0.299		-0.138	
	(0.880)		(0.312)		(0.911)		(0.399)	
Linköping	-0.357		0.0131		0.175		0.218	
	(0.529)		(0.147)		(0.719)		(0.246)	
Lund	1.336**		0.132		1.178*		0.200	
	(0.581)		(0.159)		(0.635)		(0.264)	
Uppsala	-0.561		-0.236		-0.772		-0.201	
	(0.520)		(0.148)		(0.700)		(0.231)	
τ	0.146*	-0.226*	0.0120	-0.0425	0.176	-0.261	0.0568	-0.0547
	(0.0759)	(0.115)	(0.0156)	(0.0457)	(0.145)	(0.206)	(0.0664)	(0.102)
Constant	10.99***	11.91***	2.008***	2.255***	10.81***	11.97***	1.958***	2.306***
	(1.335)	(2.359)	(0.524)	(0.783)	(1.338)	(2.338)	(0.561)	(0.814)
Municipality Specific Time Trends	NO	NO	NO	NO	YES	YES	YES	YES
Year Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,677	947	721	413	1,677	947	721	413
R-squared	0.366	0.362	0.196	0.185	0.367	0.362	0.199	0.186
No. of Clusters	273	176	95	62	273	176	95	62

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table D.2: *Regression results on fail rates in English A from runs with and without municipality specific time trends*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Big Cities	Baseline N & S	Big Cities N & S	Baseline	Big Cities	Baseline N & S	Big Cities N & S
VARIABLES	Swe. A	Swe. A	Swe. A	Swe. A	Swe. A	Swe. A	Swe. A	Swe. A
Treatment	0.0281	0.564	0.172*	0.283	-0.279	0.409	-0.0867	0.357*
	(0.201)	(0.427)	(0.0969)	(0.180)	(0.297)	(0.471)	(0.148)	(0.202)
Avg. Female 3	-0.0125**	-0.0120**	0.00251	0.00454	-0.0128**	-0.0121**	0.00264	0.00454
	(0.00569)	(0.00589)	(0.00271)	(0.00421)	(0.00574)	(0.00592)	(0.00266)	(0.00423)
Avg. Foreign 3	-0.0130	-0.0160	0.00494	0.00503	-0.0130	-0.0164	0.00496	0.00542
	(0.00946)	(0.00988)	(0.00596)	(0.00650)	(0.00944)	(0.00997)	(0.00603)	(0.00659)
Avg. High Ed. 3	-0.0613***	-0.0736***	-0.0148***	-0.0176***	-0.0613***	-0.0738***	-0.0147***	-0.0174***
	(0.00636)	(0.00786)	(0.00325)	(0.00430)	(0.00632)	(0.00792)	(0.00328)	(0.00439)
Gothenburg	-0.0529	-0.0175	-0.209*	-0.186*	0.460	0.128	-0.135	-0.274**
	(0.278)	(0.319)	(0.107)	(0.110)	(0.448)	(0.451)	(0.132)	(0.136)
Malmö	-0.324	-0.367	-0.160	-0.183	-0.209	-0.471	0.00937	-0.0885
	(0.294)	(0.297)	(0.117)	(0.116)	(0.484)	(0.503)	(0.206)	(0.212)
Helsingborg	-0.874***		-0.0832		-0.582		-0.200	
	(0.322)		(0.176)		(0.468)		(0.224)	
Linköping	-0.583*		0.144		0.391		0.674***	
	(0.329)		(0.0998)		(0.409)		(0.134)	
Lund	0.190		0.0756		0.356		0.309	
	(0.231)		(0.115)		(0.362)		(0.202)	
Uppsala	-0.556**		-0.0951		-0.380		0.0561	
	(0.249)		(0.106)		(0.368)		(0.211)	
τ	0.0747*	-0.0952	0.00358	-0.0454	0.260**	-0.0312	0.125**	-0.0794
	(0.0386)	(0.0821)	(0.0226)	(0.0312)	(0.111)	(0.136)	(0.0611)	(0.0583)
Constant	5.811***	6.273***	1.246***	1.110**	5.554***	6.264***	1.128***	1.092**
	(0.762)	(0.774)	(0.391)	(0.499)	(0.760)	(0.805)	(0.402)	(0.509)
Municipality Specific Time Trends	NO	NO	NO	NO	YES	YES	YES	YES
Year Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,670	940	719	411	1,670	940	719	411
R-squared	0.220	0.218	0.136	0.140	0.225	0.218	0.156	0.143
No. of Clusters	273	176	95	62	273	176	95	62

Robust standard errors in brackets

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table D.3: *Regression results on fail rates in Swedish A from runs with and without municipality specific time trends*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Big Cities	Baseline N & S	Big Cities N & S	Baseline	Big Cities	Baseline N & S	Big Cities N & S
VARIABLES	Grad. Rate	Grad. Rate	Grad. Rate	Grad. Rate	Grad. Rate	Grad. Rate	Grad. Rate	Grad. Rate
Treatment	2.558***	5.037***	1.777*	5.195***	1.064	2.336*	1.071	2.139
	(0.896)	(1.258)	(1.023)	(1.490)	(0.987)	(1.333)	(1.107)	(1.411)
Avg. Female 4	-0.0646**	-0.0715**	-0.0832**	-0.132**	-0.0666***	-0.0737**	-0.0825**	-0.133**
	(0.0251)	(0.0350)	(0.0381)	(0.0605)	(0.0251)	(0.0353)	(0.0384)	(0.0611)
Avg. Foreign 4	-0.122***	-0.0868**	-0.113*	-0.147*	-0.124***	-0.0877**	-0.112*	-0.146*
	(0.0384)	(0.0415)	(0.0572)	(0.0800)	(0.0387)	(0.0417)	(0.0569)	(0.0804)
Average High Ed. 4	0.327***	0.366***	0.271***	0.213**	0.326***	0.366***	0.270***	0.213**
	(0.0293)	(0.0422)	(0.0533)	(0.0916)	(0.0295)	(0.0422)	(0.0546)	(0.0920)
Gothenburg	7.080***	8.321***	5.710***	6.853***	11.03***	10.57***	10.27***	9.353***
	(1.689)	(1.897)	(1.959)	(1.970)	(2.255)	(2.318)	(2.538)	(2.301)
Malmö	8.905***	9.510***	6.944***	7.390***	11.51***	11.36***	10.61***	9.813***
	(1.495)	(1.579)	(1.533)	(1.681)	(2.031)	(2.056)	(1.823)	(1.828)
Helsingborg	8.460***		6.167***		11.31***		7.757***	
	(1.505)		(1.438)		(2.043)		(1.681)	
Linköping	7.329***		5.202***		10.54***		8.540***	
	(1.552)		(1.585)		(1.815)		(1.699)	
Lund	9.542***		8.789***		12.31***		11.35***	
	(1.719)		(1.696)		(2.563)		(2.575)	
Uppsala	3.227*		2.543*		5.674**		4.011*	
	(1.742)		(1.391)		(2.602)		(2.304)	
τ	-0.340	-1.195***	-0.156	-1.167***	0.885*	-0.00431	0.683	0.167
	(0.217)	(0.305)	(0.176)	(0.435)	(0.516)	(0.521)	(0.464)	(0.519)
Constant	66.45***	62.97***	72.24***	78.35***	64.58***	62.09***	70.18***	77.27***
	(3.290)	(4.317)	(5.724)	(9.383)	(3.302)	(4.274)	(5.733)	(9.364)
Municipality Specific Time Trends	NO	NO	NO	NO	YES	YES	YES	YES
Year Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,560	874	689	396	1,560	874	689	396
R-squared	0.422	0.398	0.444	0.363	0.428	0.401	0.461	0.371
No. of Clusters	256	160	90	58	256	160	90	58

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table D.4: *Regression results on Graduation Rates from runs with and without municipality specific time trends*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Big Cities	Baseline	Big Cities	Baseline	Big Cities	Baseline	Big Cities
			N & S	N & S			N & S	N & S
VARIABLES	GPA	GPA	GPA	GPA	GPA	GPA	GPA	GPA
Treatment	-0.00534	-0.0453	-0.0849	-0.0654	0.0572	-0.131	0.0537	-0.0618
	(0.0713)	(0.130)	(0.101)	(0.174)	(0.0795)	(0.128)	(0.0892)	(0.159)
Avg. Female 3	0.0132***	0.0135***	0.00716**	0.00437	0.0132***	0.0135***	0.00747**	0.00436
	(0.00174)	(0.00257)	(0.00339)	(0.00582)	(0.00176)	(0.00258)	(0.00343)	(0.00584)
Avg. Foreign 3	0.00951***	0.0114***	0.0321***	0.0259***	0.00972***	0.0117***	0.0331***	0.0260***
	(0.00315)	(0.00407)	(0.00677)	(0.00883)	(0.00313)	(0.00406)	(0.00675)	(0.00883)
Average High Ed. 3	0.0698***	0.0718***	0.0963***	0.0873***	0.0700***	0.0720***	0.0972***	0.0874***
	(0.00224)	(0.00369)	(0.00613)	(0.00953)	(0.00226)	(0.00369)	(0.00621)	(0.00956)
Gothenburg	0.171	0.183	0.324*	0.302	0.169	0.240	0.313	0.289
	(0.132)	(0.140)	(0.188)	(0.188)	(0.177)	(0.178)	(0.267)	(0.254)
Malmö	0.388***	0.392***	0.437***	0.392**	0.620***	0.671***	0.465**	0.436**
	(0.134)	(0.132)	(0.162)	(0.159)	(0.189)	(0.190)	(0.202)	(0.203)
Helsingborg	0.268*		0.666***		0.135		0.438	
	(0.159)		(0.193)		(0.221)		(0.278)	
Linköping	-0.200		-0.226*		-0.339*		-0.402**	
	(0.157)		(0.132)		(0.205)		(0.191)	
Lund	-0.209		-0.387**		-0.300		-0.616**	
	(0.180)		(0.175)		(0.242)		(0.268)	
Uppsala	0.112		-0.0898		0.0621		-0.289	
	(0.160)		(0.171)		(0.215)		(0.228)	
τ	0.0320*	0.102***	0.0405*	0.0826*	0.00485	0.147***	-0.0416	0.0804
	(0.0164)	(0.0280)	(0.0216)	(0.0439)	(0.0427)	(0.0456)	(0.0506)	(0.0568)
Constant	8.712***	8.627***	6.752***	7.675***	8.701***	8.541***	6.749***	7.662***
	(0.198)	(0.304)	(0.615)	(0.932)	(0.211)	(0.308)	(0.634)	(0.950)
Municipality Specific Time Trends	NO	NO	NO	NO	YES	YES	YES	YES
Year Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,681	951	722	414	1,681	951	722	414
R-squared	0.767	0.731	0.763	0.677	0.769	0.733	0.766	0.677
No. of Clusters	274	177	95	62	274	177	95	62

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table D.5: *Regression results on GPA from runs with and without municipality specific time trends*

Academic Year	Math. A	Eng. A	Swe. A	GPA	Grad. Rate
1996/97	6.651	3.969	1.702	12.777	-
1997/98	5.594	3.739	1.734	13.067	-
1998/99	6.362	3.492	1.426	13.417	80.431
1999/00	6.749	4.034	1.844	13.494	78.944
2000/01	5.030	2.997	1.315	13.862	78.738
2001/02	4.307	2.439	1.118	14.006	77.245
2002/03	4.027	2.162	1.074	14.169	78.490
2003/04	3.538	2.507	0.967	14.246	80.634
2004/05	3.780	2.698	1.131	14.297	82.077
2005/06	4.029	2.384	1.164	14.401	81.659
2006/07	3.774	2.228	1.139	14.452	81.354
2007/08	3.813	2.159	1.081	14.380	81.788
2008/09	2.788	1.788	1.000	14.464	81.967
2009/10	2.678	1.759	0.967	14.540	82.331
2010/11	2.687	1.676	0.929	14.512	81.499
2011/12	2.823	1.842	0.797	14.551	83.603
2012/13	-	-	-	-	83.270
2013/14	-	-	-	-	83.292

Table D.6: *Mean fail rates, GPA, and graduation rate over years in our sample*

E Robustness Checks

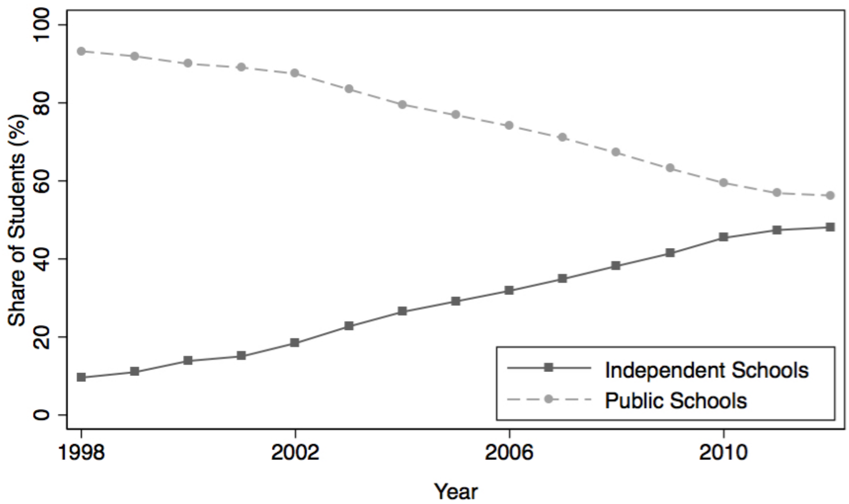


Figure E.1: *Share of students in public and independent schools over time in our sample*

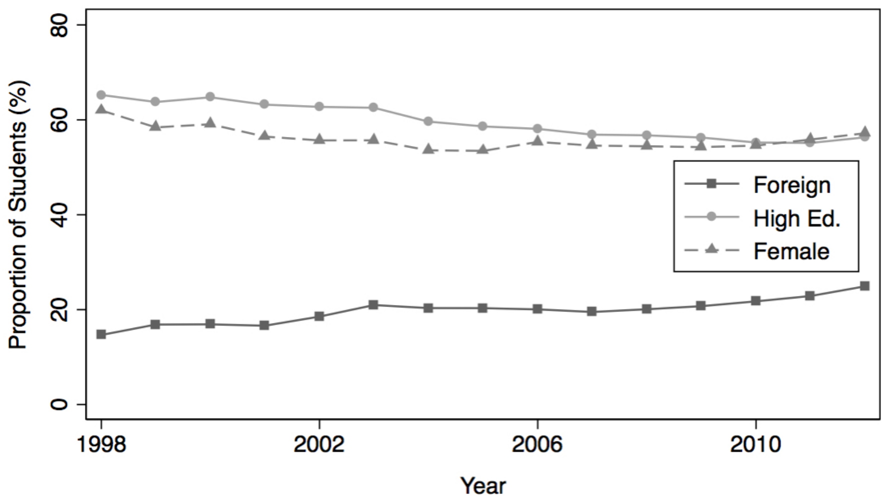


Figure E.2: *Student composition in independents schools over time in our sample*

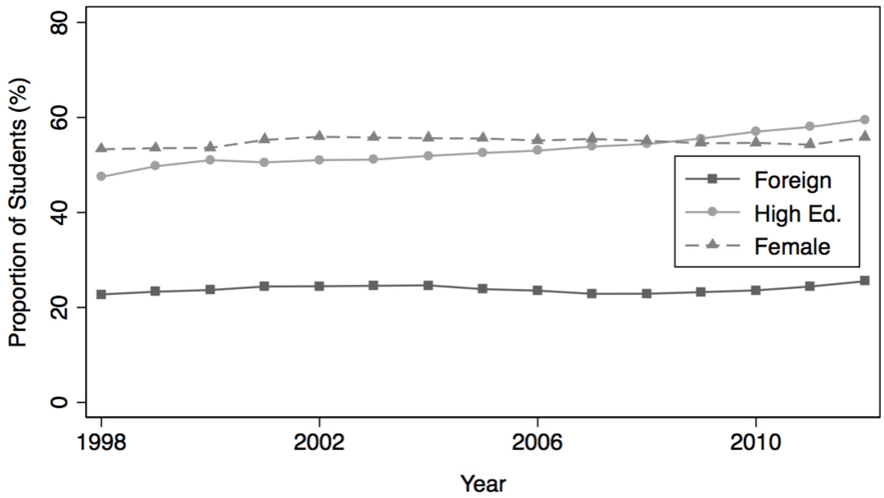


Figure E.3: *Student composition in public schools over time in our sample*

	(1)	(2)	(3)
	Big Cities	Big Cities	Baseline N & S
VARIABLES	Eng. A	Math. A	Swe. A
3-year lead	-0.168	0.662	0.0212
	(0.502)	(1.001)	(0.145)
2-year lead	0.0295	-0.00899	0.160
	(0.557)	(0.823)	(0.150)
1-year lead	-0.321	1.305	-0.0369
	(0.465)	(0.861)	(0.133)
Treatment	1.268***	1.907**	0.170
	(0.458)	(0.889)	(0.121)
1-year lag	0.0955	0.939	-0.0933
	(0.487)	(0.893)	(0.0928)
Avg. Female 3	-0.0172	0.000598	0.00243
	(0.0203)	(0.0295)	(0.00273)
Avg. Foreign 3	-0.00484	-0.119***	0.00478
	(0.0199)	(0.0360)	(0.00604)
Avg. High Ed. 3	-0.157***	-0.312***	-0.0149***
	(0.0190)	(0.0340)	(0.00326)
Gothenburg	-0.128	0.830	-0.190*
	(0.786)	(1.486)	(0.112)
Malmö	-0.776	-0.107	-0.147
	(0.613)	(1.185)	(0.117)
Helsingborg			-0.0836
			(0.177)
Linköping			0.214
			(0.140)
Lund			0.0865
			(0.116)
Uppsala			-0.0240
			(0.156)
τ	-0.172	-0.710	-0.00343
	(0.273)	(0.535)	(0.0287)
Constant	11.96***	22.93***	1.237***
	(2.308)	(3.481)	(0.392)
Municipality Specific Time Trends	NO	NO	NO
Year Dummies	YES	YES	YES
Observations	947	951	719
R-squared	0.362	0.420	0.138
No. of Clusters	176	177	95

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table E.1: *Placebo regression results for fail rates*

	(1)	(2)	(3)	(4)
	Baseline	Big Cities	Baseline N & S	Big Cities N & S
VARIABLES	Grad. Rate	Grad. Rate	Grad. Rate	Grad. Rate
3-year lead	-0.267	2.318	0.0605	3.179**
	(0.894)	(1.586)	(0.931)	(1.411)
2-year lead	0.241	2.244	-1.490*	0.390
	(0.799)	(1.358)	(0.880)	(1.292)
1-year lead	-0.0414	1.854	0.535	3.757**
	(0.735)	(1.417)	(0.785)	(1.439)
Treatment	2.361***	5.500***	2.135**	5.154***
	(0.824)	(1.328)	(0.904)	(1.341)
1-year lag	0.514	2.768*	-0.628	2.884
	(0.775)	(1.563)	(0.919)	(1.897)
Avg. Female 4	-0.0647**	-0.0736**	-0.0823**	-0.133**
	(0.0252)	(0.0352)	(0.0381)	(0.0609)
Avg. Foreign 4	-0.122***	-0.0889**	-0.111*	-0.150*
	(0.0385)	(0.0415)	(0.0572)	(0.0807)
Avg. High Ed. 4	0.327***	0.366***	0.272***	0.210**
	(0.0295)	(0.0422)	(0.0536)	(0.0930)
Gothenburg	7.093***	10.73***	5.478***	9.422***
	(1.677)	(2.202)	(1.949)	(2.037)
Malmö	8.900***	10.67***	6.825***	8.625***
	(1.482)	(1.628)	(1.527)	(1.710)
Helsingborg	9.531***		8.818***	
	(1.726)		(1.703)	
Linköping	8.424***		5.586***	
	(1.698)		(1.612)	
Lund	7.319***		5.107***	
	(1.561)		(1.602)	
Uppsala	3.199		1.920	
	(2.043)		(1.677)	
τ	-0.332	-2.582***	-0.0603	-2.643***
	(0.258)	(0.793)	(0.241)	(0.904)
Constant	66.47***	62.21***	72.25***	77.75***
	(3.288)	(4.235)	(5.746)	(9.422)
Municipality Specific Time Trends	NO	NO	NO	NO
Year Dummies	YES	YES	YES	YES
Observations	1,560	874	689	396
R-squared	0.422	0.402	0.446	0.373
No. of Clusters	256	160	90	58

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table E.2: *Placebo regression results for graduation rates*