CERGE
Center for Economic Research and Graduate Education
Charles University Prague

Essays on the Effects of Early School Tracking

Miroslava Federičová

Dissertation

Prague, April 2016

Dissertation Committee

Daniel Münich (CERGE-EI, Charles University; chair)

Alena Bičáková (CERGE-EI, Charles University)

Štěpán Jurajda (CERGE-EI, Charles University)

Randall K. Filer (Hunter College, CUNY, New York)

Referees

Steven G. Rivkin (University of Illinois at Chicago)

Torberg Falch (Norwegian University of Science and Technology)

Table of Contents

Abstract

This dissertation explores the different viewpoints on the early-tracking school system and the effect of this early selection on pupils' schooling outcomes. Moreover, as this early selection occurs at the time of high gender differences, all chapters examine the topics also from gender perspective.

The first chapter is focused on the selection process itself and studies the role of grades in explaining the gender difference in application rates to selective schools. This selection is provided mostly according to cognitive skills that are signalled to pupils in the form of grades. Although grades play a very important role in application process, conditional on cognitive skills, grades differ substantially between girls and boys. In this chapter, I propose the model of asymmetric signal of the probability of admission for girls and boys arising from grades. Data about transition from primary to selective schools in the Czech Republic shows that girls apply at significantly higher rates. Leaning on the model, I find that this difference remains the same also after controlling for probability of admission. Furthermore, test scores collected by an international testing program have no effect on gender differences in applications that are, however, in large part explained by grades. This finding is consistent with grades acting as a signal that provides imperfect and incomplete information about the probability of being admitted, and consequently causes the gender difference in application.

The second chapter contributes to the literature on high-stakes testing as a tool for educational accountability. In order to isolate the impact of high-stakes admission exams to selective schools on student academic achievement, I exploit the school policy change in Slovakia that shifted the timing of these exams to the later grade. Using two waves of TIMSS survey data and difference-in-difference methodology I find that high-stakes exams increase ten-year-old students' math test

scores by 0.2 standard deviations on average. This effect additionally accrues by around 0.05 standard deviations for students in the top decile, i.e. students with the highest probability of being admitted to selective schools. Although the effects are similar for both genders, there are indications that girls put more effort into preparing for such exams than boys in a more competitive environment.

The last chapter discusses the impact of negative externalities of early-tracking on pupils' self-confidence. In this chapter, I focus on those pupils that did not apply for selective schools (are left apart from selection process) but at the same time they observe the preparation process of their classmates. As this process is stressful and often involves many months of preparation, it naturally divides many 5th grade classes into two groups: pupils preparing for exams to enter better schools and everyone else, who do not compete for selective schools. I show that this environment has a detrimental effect on the self-confidence of pupils in mathematics who do not apply for selective schools but have peers in their classroom who do apply. This effect is particularly strong for girls. Controlling for the number of successful applicants in classroom further indicates that girls are more likely to experience declines in self-confidence due to the loss of peers than the competitive environment itself.

Abstrakt

Tato dizertace zkoumá rozličné pohledy na školský systém rané selekce a vlivy této rané selekce na výsledky žáků. Navíc, jelikož táto raní selekce nastává v období velikých rozdílů mezi chlapci a děvčaty, všechny články vyšetřují toto téma taky z pohledu genderových nerovností.

První kapitola je zaměřen na samotný proces selekce a zabývá se úlohou známek ve vysvětlení genderových nerovností v míře přihlášek na selektivní školu. Táto selekce je většinou prováděna na základě kognitivních schopností, které jsou signalizované žákům ve formě známek. Ačkoliv hrají známky důležitou roli v přihlášení se, tyto se značně liší mezi děvčaty a chlapci, a to i pokud mají stejné kognitivní dovednosti. V této kapitole navrhuji model asymetrického signálu pravděpodobnosti přijetí pro chlapce a děvčata, který vyvstává ze známek. Údaje z přechodu mezi primárním stupněm vzdělání a víceletým gymnáziem v České Republice ukazují, že děvčata se hlásí ve významně větší míře. Opírajíc se o model, se tento rozdíl nemění ani po zohlednění šancí jednotlivých žáků na přijetí. Navíc, výsledky z mezinárodního šetření nemají žádný vliv na genderový rozdíl v přihlašování se, který je nicméně ve značné míře vysvětlený známkami. Toto zjištění je v souladu s představou, že známky jsou signálem, který nabízí nedokonalou a neúplnou informaci o pravděpodobnosti přijetí, a následně způsobuje genderový rozdíl v míře přihlášek.

Druhá kapitola přispívá do literatury ohledně zkoušek, ve kterých je hodně v sázce, jako nástroj pro odpovědnost ve vzdělání. Abych oddělila dopad přijímacích zkoušek do selektivních škol na studijní výsledky žáků, využívám exogenní změnu školního systému na Slovensku, která posunula tyto zkoušky o jeden rok později. Využitím dat ze dvou vln mezinárodního šetření žáků TIMSS a metody rozdílu-v-rozdílech jsem zjistila, že tyto přijímací zkoušky navýší matematická skóre desetiletých žáků v průměru o 0,2 standardní odchylky. Tento efekt je vyšší o dalších 0,05 standardní odchylky u žáků v nejvyšším decilu, tedy u žáků, kteří mají největší šance být přijati

na výběrové školy. Ačkoliv jsou tyto odhadnuté efekty podobné u obou pohlaví, nacházím indicie, že v konkurenčnějším prostředí vyvíjejí dívky vyšší studijní úsilí než chlapci.

Poslední kapitola pojednává o dopadu negativních externalit rané selekce na sebedůvěru žáků. V této kapitole se zaměřuji na ty studenty, kteří se nehlásí do výběrových škol (tedy jsou vynecháni ze selektivního procesu) ale zároveň pozorují, jak se jejich spolužáci připravují na přijímačky. Jelikož je tento proces stresující a často představuje mnohoměsíční přípravy, rozděluje přirozeně mnoho tříd pátého ročníku na dvě skupiny: ti, kteří se připravují na přijímačky do výběrových škol a ti ostatní, kteří se neúčastní tohoto výběrového procesu. V analýze ukazuji, že toto prostředí má nepříznivý vliv na sebedůvěru v matematice u těch žáků, kteří se nikam nehlásí, ale mají ve třídě spolužáky, kteří se naopak na selektivní školy hlásí. Tento dopad je výrazný zejména u děvčat. Kontrolujíc pro počet úspěšných uchazečů o selektivní školy ve třídě dále naznačuje, že dívky jsou více náchylné na pokles v sebedůvěře z důvodu odchodu spolužáků ze třídy nežli z důvodu samotného konkurenčního prostředí ve třídě.

Acknowledgments

First of all, my special thanks to my supervisor, Daniel Münich, for all his support and encouragement throughout the dissertation process.

I am also very grateful to all the members of my dissertation committee, Štěpán Jurajda, Alena Bičáková, and Randy Filer, and to all my fellow students at CERGE-EI for their advice and comments, and in helping me to formulate and better express my research ideas.

I would also like to thank Steven Rivkin, the head of the Economic Department of the University of Illinois, for his invaluable help and advice during my research stay in Chicago.

Above all, inexpressible thanks goes to my family for their support, love and patience throughout my studies.

Financial support of the Grant Agency of the Charles University (grant no. P402/12/G130) is gratefully acknowledged.

Czech Republic, Prague Miroslava Federičová

April 2016

Introduction

In my dissertation, I focus on the early-tracking school system and its relation with the inequality in educational opportunities and outcomes. As this early selection occurs at the time of large gender disparities, I particularly discuss its impacts on gender inequality. In the first chapter, I explore the gender inequality in application process to selective schools. The second chapter examines the impact of high-stakes admission exams to selective schools on student achievement exploiting the exogenous change of school policy in Slovakia. Turning the attention to those students who did not apply – thus, students that are left behind from selection process –, the third chapter studies the change in their self-confidence due to emerging competitive environment in the classroom.

Early-tracking school system is one of the issues among school policy experts. This school system is characterized by the selection of students to ability differing schools already in the fourth or fifth grade, whereas the comprehensive school system is composed by at least eight years in the same educational institutions. The main critique of selection at early ages is seen in its connection with the inequality of educational opportunity and outcomes generated by differences in parental background. At early ages, pupils are even more affected by the aspirations of their parents than they are in later grades. Thus, the differences in socio-economic background of students may affect not only the pupils' decision to apply to selective schools but also their further educational and labour outcomes.

Several empirical studies examine the overall impact of the early-tracking school system on this inequality. In the first chapter, I contribute to this literature by focusing particularly on the application process that constitutes the first trigger of inequalities. As the selection process to differing ability schools is based on pupils' knowledge, the decision to apply – whether made by parents or pupils themselves – is affected by pupils' perception of their own abilities to success in admission exam. At this early age, grades usually form the only source of information about pupils' abilities, although they are generally perceived as subjective and inaccurate measures of pupils' achievement. Moreover, they produce considerable gender differences in the sense that conditional on cognitive skills girls usually receive better grades than boys. In the first chapter, I show that this gender disparity in grading in large part explains the gender difference in applications to selective schools also after controlling for the predicted probability of admission.

In order to be accepted into the selective schools, pupils undertake high-stakes admission exams, which is a stressful process and often involves many months of preparation. In general, the impact of high-stakes exams have been thoroughly examined as they are seen as an important instrument for enhancing students' motivation and educational accountability. The second chapter, written together with Daniel Münich, contributes to this literature by isolating the impact of teachers' efforts – as they tend to focus on teaching tested skills – from the motivational impact of high-stakes exams on pupils' achievement. In this chapter, we exploit the school reform implemented in Slovakia in 2009 that shifted the transition to selective schools, and thus, also the timing of the high-stakes exams, from the 4^{th} to the 5^{th} grade. This early selection allows us to study the effects of high-stakes exams on pupils' achievement already at the age of 10, i.e. the age characterized by notable developmental changes and culminating gender differences. We show that the presence of high-stakes exams motivates these students to put higher study effort resulting in the increase of their achievement by 0.2 to 0.3 standard deviation. Our results also indicate higher effort in girls in more competitive environments.

Although there is an extensive literature concerning the inequality in education coming from the early-tracking system, relatively little is known about the impact of this selection process on those pupils that did not apply, and hence, are left behind. The third chapter, which is a joint work with Filip Pertold and Michael Smith, argues that early-tracking generates inequalities far before actual placement into selective schools. We suggest that the application, and in the case of girls mainly the admission, of peers negatively affects self-confidence of those who did not apply.

Throughout this thesis, I take advantage of the adolescent testing that has been carried out especially in OECD countries, both at different age levels and in different abilities (e.g. the Trends in International Mathematics and Science Study, commonly known as TIMSS; the Progress in International Reading Literacy Study, PIRLS; etc.). Moreover, I use data from Czech Longitudinal

Study in Education (CLoSE) that follows Czech pupils tested in the 4th grade by TIMSS and PIRLS in 2011 through the next grades. This survey allows the examination of the same panel of pupils before and after the early selection occurs.

1 Chapter

Gender Gap in Application to Selective Schools: Are Grades a Good Signal?

1.1 Introduction

Teachers' evaluations, and more particularly grading, are one of the first – and in the early stage of education usually the only – assessments of their abilities that pupils have access to. Hence, from the beginning of primary education, grades help to form pupils' perceptions of their own study aptitudes and thus indirectly influence a student's choice of school and other future educational decisions.

Despite the importance of the role of grades we actually do not know how grades are created (Matějů and Smith, 2014). Moreover, as the process of grade assignment differs greatly across schools and teachers, they are perceived as subjective (Hoge and Coladarci, 1989) and forming inaccurate predictions of pupils' achievement (Hoge and Coladarci, 1989; Südkamp, Kaiser and Möller, 2012). This inaccuracy seems to be particularly important for boys. Several empirical studies highlight that, conditional on cognitive skills, grades differ substantially between girls and boys (Falch and Naper, 2013; Matějů and Smith, 2014). The gender gap in grades is usually explained – though not fully (Matějů and Smith, 2014) – by noncognitive skills in which boys lag behind girls (Seligman and Duckworth, 2006). As a consequence, girls' grades usually overestimate their true cognitive skills, whereas boys' grades are correspondingly undervalued.

One of the issues that can emerge from this uncertainty in grade assignment is an inappropriate choice of school track. Although grades seem to be an important component of school choice, entering into both the pupil's decision to apply and the school's admission decision, to the best of my knowledge no literature has already addressed the effect of grades in school transition. The goal of this paper is to help to fill this gap and to examine the role of grades in explaining the gender difference in application rates to selective schools.

If abilities tested in admission exams (or required in admission criteria) are not the same as those that teachers use for assigning grades, then grades provide a biased signal of the chances of admission. The bigger the difference between grades and admission requirements, the more inaccurate this signal is. The fact that almost all selective schools administer their own admission tests is evidence that they do not believe that grades provide adequate information regarding their requirements. However, the appropriate choice of school track does not rely only on the correspondence between grades and admission requirements, but also on the extent to which pupils base their educational choice on grades. The fact that grades are usually the only information about pupils' abilities that they have access to makes this question even more exigent. If pupils weight grades too heavily when constructing their own subjective probabilities of admission, those with high grades and lower cognitive skills may overstate their chances and those with lower grades but higher cognitive skills may understate them. Boys appear more likely to be in the latter group.

Using the transition from primary schools to selective schools in the Czech Republic, this paper shows the significant gender gap in applications. This gap is fully explained by gender difference in grades, but controlling for achievement test scores it is even wider. Although girls and boys have the same skill distribution measured by the predicted probability of admission, conditional on that probability girls are still more likely to apply. As only top pupils apply to selective schools, this analysis is further focused on the right hand end of the predicted admission distribution. There is no gender gap in the application rate in the top admission decile. Moreover, in this decile a high percentage of girls and boys earn the best grades from math and Czech language. However, in the second top admission decile, girls apply to selective schools significantly more often than boys, by 9 p.p. At the same time, 83% of girls in this decile earn the best grades from both math and Czech language in comparison with only 57% of boys. Controlling for grades, the 9 p.p. gender difference drops and becomes insignificant. Similar gender differences persist in the third from top admission decile. In other words, the gender gap in application rate is present mostly for marginal pupils and coincides with the gender gap in grades. This is consistent with grade acting as a signal in the application decision.

To study the determinants of application and admission to selective schools I use Czech longitudinal data. Grades in transition to selective schools are presumably more important than in transition to vocational schools or academies that are based more on aptitude than achievement tests. Although grades play an important role in all school transitions, this paper focuses on the specific school transition after primary education. The reason is that in the further transitions, pupils have the possibility to choose between a wider range of tracks, within which some are gender specific. These gender specific outside options may influence gender composition of applicants to other tracks[1]. Regarding the transition process after primary education, girls are shown to apply to selective schools at a significantly higher rate (UIV, 2009). This gender difference in application rate deepens further in the next transition to upper-secondary education.

There may also be other sources of a gender gap in application than grades. Evidence presented in Charness and Gneezy (2012) shows that girls tend to be more risk averse, and the findings in Spencer, Steele and Quinn (1999) suggests that girls face higher testing anxiety due to stereotype-threat. However, this is in contrast with the higher application rate for girls, as these effects may rather deter girls from applying to selective schools. Moreover, regarding the choice of university, Jurajda and Münich (2011) show no gender gap in applications even to the very competitive universities[2]. On the other hand, girls are perceived as more conscientious and more persevering on long-term assignments (Seligman and Duckworth, 2006) which gives girls an advantage in the preparation process for an admission exam and deters boys from applying.

[1] After the lower secondary education, pupils in the Czech Republic can choose between academic school ("Gymnasium" in Czech), four-year technical school or three-year vocational school. Whereas technical schools have even attendance rate by sex, in vocational schools two thirds of pupils are boys. If boys have high aspirations to apply to vocational school, its unequal attendance rate by sex can already affect transition after primary education. According to international testing PISA, pupils in vocational schools are left behind the average pupil in their cohort in math test scores by 1 standard deviation and 2 standard deviations behind the pupils from the same cohort in 8-year gymnasium. Since only 20% of pupils apply to selective schools after primary education, this selective track targets primarily high-performing pupils. The assessment data from the transition year to selective schools in 2012 shows that no pupil with a test score lower than 0.75 standard deviations below average applied to selective schools.

[2] However, the higher anxiety during testing among girls in comparison to boys may potentially bias the measure of predicted probability of admission. I address this issue further in the paper.

Nevertheless, explaining all the causes of the gender gap in application rate is not the ambition of this paper, nor is it aimed at uncovering a black box of all factors affecting grade assignment. By using a simple signalling model, followed by an empirical analysis, this paper rather wants to point to the excessive importance of grades in pupils' school choice and its effect on the gender gap in application decisions. The role of grades in decision making may also help to explain other social differences in application decisions, i.e. regarding the socio-economic status of family, minorities, etc.

The paper is organized as follows. The next section reviews the literature on teachers' judgements of pupils' performance and its further effect on schooling decisions. Section 3 formalizes the conceptual framework of the application and admission process. This is followed by the description of the data in Section 4 and by setting out the empirical strategy in Section 5. Section 6 discusses the main results and provides several robustness checks. The study concludes in Section 7 with key findings and policy implications.

1.2 Literature

This section firstly discusses the educational literature and empirical studies concerning the accuracy of teacher judgements. This is followed by an examination of the literature which explore the gender difference in the noisiness of grades according to pupil achievement. This literature helps to uncover what stands behind grades and what is still unknown. The second part of this section is focused on pupils' perceptions and examines literature about the effect of grades on their educational choices.

The effect of teachers' expectations on student achievement was thoroughly developed for the first time in the Pygmalion effect established by Rosenthal and Jacobson (1968). In the following years, many other studies - especially in educational psychology - considered the accuracy of teachers' judgements of student achievement. Hoge and Coladarci (1989) and Südkamp, Kaiser and Möller (2012) reviewed the major empirical studies which emerged mostly in the U.S. over the last forty years. They agree that the correlation between teacher-based judgement and a student academic achievement is relatively high and reaches on average 0.6, although the results vary substantially with values from 0.3 to 0.9. Hoge and Coladarci (1989) concluded that teacher judgments are subjective and susceptible to stereotypes. Moreover, Coladarci (1986) revealed that the accuracy of teachers' judgements differs for low and high achievers with the latter enjoying more accurate judgements.

Further studies attempted to explain the variance in teachers' judgements. Bennett, Gottesman, Rock and Cerullo (1993) considered class-room behaviour as one of the sources of this variance. Examining a sample of U.S. students in the first two years of primary school, they find that a shift in a grade for behaviour by one standard deviation produces only a marginally lower effect on grades than a shift by one standard deviation in academic skills. On the other hand, Feinberg and Shapiro (2010) show teachers' judgements may be based on the average ability of a class. In this case, the variance in average class ability may induce the variance in teachers' judgements. Thus, pupils are not only prone to form their academic self-concepts according to the average ability level of their peers (Marsh, 1984; Marsh, 1987), but teachers also judge pupils this way.

To examine the prevailing gender differences in school achievement Mechtenberg (2009) offers a theoretical model of grading. The model is based on the different perceptions of girls and boys about the meaning of their grades and teachers' responses to these beliefs in grading. As teachers do not want to distort pupils' perceptions of their own academic achievement, they report inaccurate grades only if they are convinced pupils would not internalize these grades. According to Mechtenberg (2009), the pure strategy equilibrium leads to noisy grades in humanities and math for girls, whereas for boys only good grades in humanities are noisy. Hence, their model denotes girls' grades as less accurate than those of boys.

Several empirical studies go along with the Mechtenberg theory. In the study of Norwegian students, Falch and Naper (2013) find girls' grades noisier than boys when results from central exit exams are taken into account. Their finding that girls earn higher grades than boys is further verified in Matějů and Smith (2014), for Czech pupils. They show that girls have on average a two and four times higher probability of earning the top grade from Mathematics and Czech language than boys with the same test scores. This gender difference in grades is usually explained by noncognitive skills in which girls use to outperform boys (Seligman and Duckworth, 2006). However, Matějů and Smith (2014) show that the part of gender gap in grading remains unexplained after controlling for cognitive and noncognitive skills. Betts and Morell (1999) found similar difficulties when they tried to determine the variation in GPA among students in several U.S. universities. Including high school GPA and SAT scores, they also found other relevant factors in the prediction of university GPA such as high school location or socio economic background characteristics.

Although it remains unclear what causes the inaccuracy in teacher's judgements and which factors influence grades, the uncertainty of grades may lead to inequalities in educational opportunities only if grades are considered as accurate measures of pupils' skills and as a main source for important decisions on the educational path. The relevance of grades in explaining the gender gap in the application decision is the main goal of this paper. To the best of my knowledge, no study has so far investigated the effect of grades on decision making during primary and secondary education. Although part of the Matějů and Smith study (2014) is dedicated to the gender difference in the application decision after lower secondary education, they do not attempt to explain the role of grades in this school choice. They find that girls are more likely than boys to apply to selective schools even conditional on elementary school grades. This persistent gender gap in application decisions can be explained by the outside option in the school transition after lower secondary education in which one of the school tracks is gender specific. Moreover, as Matějů and Smith (2014) do not have information about the admission decision, they are not able to specify whether pupils respond appropriately to their admission chances.

On the other hand, there are a number of empirical studies regarding university education that deal with the role of grades received in a course on further study outputs. Considering the issue of grade inflation in U.S. universities, Sabot and Wakeman-Linn (1991) show that difference in grading across departments considerably influences undergraduate choice of major, and hence, they conclude that grades are an important factor in the decision to abandon or to continue the subject. Rask (2010) comes to a similar result and indicates grades as one of the main factors influencing the attrition of students in STEM fields. Using data from Colgate University, Rask and Tiefenhalter (2008) find that women are more sensitive to relative grade than men when choosing economics as an undergraduate major and are less likely to continue in economics if they perform poorly. To explain the possible impact of grades in the pupils' application decision to selective schools I use the signalling model set up in the following section.

1.3 Conceptual framework

Transition between school levels usually consists of two consequential decisions, application and admission[3]. Pupils and their parents choose a school according to their aspirations, their own perceived study aptitudes and perceived probability of being admitted. On the other hand, schools set admission criteria to choose suitable students for their study program. Therefore, the majority of vocational schools and academies place a greater emphasis in admissions on aptitude tests whereas academically selective schools place rather greater emphasis on achievement tests.

In the school system in the Czech Republic, pupils can apply for selective schools at the end of the 5th, 7th and 9th grade. In the 9th grade, all students have to decide between an academic or vocational track. A different situation occurs at the 5th and 7th grade in which only 10% and 2% of pupils from the primary schools, respectively, follow the academically selective track. Other students continue in the same school until the end of the 9th grade. The decision to apply is not obligatory and usually only 20% of pupils take this option.

Using the above mentioned characteristics of the admission process, the following subsections successively model the pupils' probability of being admitted and their application decision as a response to their expected chances of success in the admission process. These theoretical models help to depict the role of grades in these two decisions. Although grades directly enter into both the admission and application decisions through admission criteria, they also indirectly affect pupils' decisions to apply. This indirect impact of grades is described by the signalling model through which pupils predict their success in admission exams, and consequently in admissions. Using this signalling model the last subsection explains the possible occurrence of a gender gap in the application decision.

[3] Regarding the gender difference in school transition this process may be more complicated. In the application decision, the gender gap may occur in the personal decision (i.e. whether to apply or not) and also in the application strategy (i.e. the choice of selective school). In the admission decision, the admission exam and also school admission decision may be gender biased. Here, I assume that selective schools base their admission decision on admission exam score and grades and do not bias it by gender.

1.3.1 Admission decision

In the majority of academically selective schools in the Czech Republic, the admission criteria consist of two elements. The first is the primary school grade average from the last or last two semesters that usually form 25% of the available admission points. The remaining 75% of points relate to results in admission exams. These comprise tests in math, Czech language and general knowledge. Reflecting the above mentioned admission criteria, the probability of admission is defined as a function of admission exam score (*T*) and grade average (*GPA*)[4] as:

$$P(admission) = a * T + b * GPA, \quad where\ a, b \in \langle 0,\ 1 \rangle \quad (1).$$

1.3.2 Application decision

Consider now the application decision. Pupils decide to apply according to their aspirations (*X*) and their own perceived probability of being admitted ($P(.)^e$):

$$P(apply) = g(P(admission)^e, X)$$

Before the deadline for applications, schools officially announce the date of admission exams, the admission criteria – i.e. the constants a and b from Eq. (1) –, and the number of free slots. According to this available information about admission criteria, pupils can form their perceived probability of admission even before the decision to apply. Although pupils know how GPA translates to their probability of admission, they are uncertain about their admission exam scores. Their perceived probability of admission is therefore based on their own grades (GPA) and expectations about their admission exam score (T^e):

$$P(admission)^e = a * T^e + b * GPA$$

To predict their score in the admission exam pupils can use several proxies such as grades, results from national testing, their relative rank in class, etc. If these proxies map the real admission score unambiguously, the perceived probability of admission, $P(.)^e$, coincides with the real

[4] Here, the admission exam score and grade average are linearly transformed to the scale <0, 1>.

probability of admission (recall Eq. (1)). Otherwise, the expected admission score is measured with error that can produce a biased estimate for the real probability of admission. This mechanism is described by the signalling model in the next subsection.

1.3.3 The signalling model

To simplify, assume that students predict their admission score (T^e) based strictly on grade. In other words, they use grade as a signal for the admission exam score and form their perceived probability of admission according to it. Using this assumption, student's decision to apply to selective school is then formed only by GPA and individual aspirations (*X*), i.e. Eq. 2 changes to

$$P(apply) = g(GPA, X)$$

Suppose further that GPA is a function of cognitive skills (*CS*) and other than cognitive skills[5] (*OS*), whereas admission exam is based only cognitive skills[6], such that

$$GPA = p(CS, OS)$$

and

$$T = q(OS).$$

If the distribution of the other than cognitive skills differ for specific groups of pupils (i.e. according to the socio-economic status of the family, gender, age, minority group, etc.), the application decision based strictly on grades can lead to biased beliefs about admission chances. As this paper is focussed on gender differences in the application rate, the next section provides the signalling model with gender specific bias. However, the same approach can be applied to any other group inequalities.

[5] The other skills refer for example to noncognitive skills, such as perseverance. This decomposition of GPA may be understood as the decomposition on skills tested in admission test scores and other skills and individual and/or class characteristics that enter to the grade formation but has no effect on admission test score, i.e. GPA = f (T) + η (where η includes characteristics that do not enter to admission score).
[6] It is clear that the other than cognitive skills also enter to the admission exam score. Thus, this simplification can be modified to a case in which the admission exam score is a function of the both cognitive and other skills, but in a way that other skills would play less role in admission exam score than in GPA.

1.3.4 Gender disparity

Assume that cognitive and other than cognitive skills are two independent random variables[7]. Further assume that the distribution of cognitive skills is equal for girls and boys, but the distribution of other skills is for boys to the left of that for girls. In other words, the cumulative distribution of other skills of boys stochastically dominates the distribution of girls, i.e.

$$F_{OS}^{boys}(x) \leq F_{OS}^{girls}(x) \text{ for all x and with a strict inequality at some x.}$$

If the decomposition of GPA to cognitive and other skills is an additive function, the distribution of GPA for boys is also to the left of that for girls. Thus, conditional on cognitive skills girls earn on average better grades than boys. This result is in line with the literature examining the gender gap in grading[8]. Hence, relying only on the signal from grades in the application decision, the same would hold for admission chances (i.e. girls would apply more often than boys with the same probability of being admitted).

1.3.5 Example

To see this, consider a case in which the both cognitive and other skills have only three realizations {1, 2, 3} where 1 is the highest and 3 is the lowest skill level. The 3 by 3 set of possible combinations of CS and OS can form the corresponding realization of grades:

		CS		
		1	2	3
	1	A	A	B
OS	2	A	B	C
	3	B	C	D

Here, the letter grades “A”, “B”, “C” and “D” corresponds to the American system of grading where A is the best and D the worst grade.

[7] Independence is a very strong assumption since other skills affects knowledge acquisition and thus cognitive skills.
[8] See e.g. Falch and Naper (2013), Matějů and Smith (2014).

Now suppose that CS is equally distributed for girls and boys and is equal to (25%, 50%, 25%). In contrast, let the distribution of OS be different for boys and girls such that girls are more likely to experience the highest skill than boys (i.e. the probability distribution function of OS for boys stochastically dominates the one for girls). Thus, suppose the distribution of other skills to be equal (50%, 30%, 20%) and (30%, 40%, 30%) for girls and boys, respectively. Then, the probability of receiving an A would equal:

$$P_g(GPA = A) = P(CS = 1)P_g(OS = 1) + P(CS = 2)P_g(OS = 1) + P(CS = 1)P_g(OS = 2)$$
$$= 0.25 * 0.5 + 0.5 * 0.5 + 0.25 * 0.3 = 0.45$$

$$P_b(GPA = A) = P(CS = 1)P_b(OS = 1) + P(CS = 2)P_b(OS = 1) + P(CS = 1)P_b(OS = 2)$$
$$= 0.25 * 0.3 + 0.5 * 0.3 + 0.25 * 0.4 = 0.325$$

for girls and boys, respectively. Thus, the probability to receive A is equal to 45% for girls and 32.5% for boys. Using grade as a signal, on average boys will tend to understate their admission chances relative to girls, and hence, apply less often than girls. However, it is clear that not all boys have the same signalling error, and hence the relationship between the signal in the form of grade and true cognitive skills may be nonlinear depending upon the underlying distribution of CS and OS.

At the early stage of school, grades are usually the only signal of abilities that pupils observe. Moreover, several studies[9] show that girls' grades usually overestimate their true cognitive skills, whereas boys' grades are undervalued according to their cognitive skills[10]. Relaxing the assumption that only cognitive skills enter to the admission exam score, if admission exams to selective schools are focused mainly on cognitive skills, using grades as a signal for admission scores, girls overestimate their true probability of admission whereas boys underestimate that probability.

---

[9] See e.g. Falch and Naper (2013), Matějů and Smith (2014), and Seligman and Duckworth (2006).

[10] Gender difference in grades according to cognitive skills is explained in several studies by noncognitive skills, in which boys usually lag behind girls (Seligman and Duckworth, 2006). If the representation of non-cognitive skills in grades is higher than in skills tested by admission exams, then the gender disparity in non-cognitive skills may explain the part of gender differences in application even after controlling for admission.

The discussion above relies on an assumption that pupils' perception of success in the admission exam is fully formed by grades. However, pupils may not believe grades and use other, more adequate signals to predict their probability of being admitted. Then, the gender difference in application to selective schools is affected also by the emphasis that pupils put on grades in order to predict their own probability of being admitted. According to this result, the gender difference in application should be fully explained by controlling for grades. In the following empirical analysis, I focus on the gender difference in application and the extent to which grades explain it.

1.4 Data

To analyse the decision to apply to selective schools at early ages, i.e. after primary education, I use data from the Czech Longitudinal Study in Education (CLoSE). One of the three main aims of CLoSE[11] is to map the transition of pupils in the Czech Republic from primary to selective schools at the end of the 5th grade. In CLoSE, pupils tested in math, science and reading skills in the 4th grade by international testing programmes TIMSS and PIRLS[12] in 2011 are followed up during the next grades. In the 5th grade, pupils were thoroughly questioned about their application, preparation and admission process to selective schools. From 177 schools tested in the 4th grade, 163 schools[13] participate in the subsequent round in the 5th grade. Hence, from 4578 students tested in TIMSS 2011, 3681 were again questioning the following year.

1.4.1 Application, preparation and admission process

In 2012, CLoSE conducted a detailed questionnaire about the application, preparation and admission process from 5th graders, i.e. pupils in the transition year. The main characteristics are

---

[11] The other two main assignments of CLoSE are to map the transition of children from kindergarten to primary school and to track the education of the adult population in the Czech Republic. Project CLoSE is financed by the Grant Agency of the Czech Republic.
[12] TIMSS and PIRLS are international testing programs organized in regular 4- and 5-year cycles, respectively. TIMSS measures trends in mathematics and science skills in the 4th and 8th grade, whereas PIRLS tests pupils in the 4th grade for reading comprehension. In 2011, 48 countries participated in TIMSS and PIRLS. In the Czech Republic the representative sample of 177 schools was selected according to region and the extended education status of school.
[13] Half of the 14 missing schools did not participate in the 5th grade because of small number of pupils at school or because more than half of their pupils change school after the 4th grade. The other half of schools did not want to overburden their pupils with other questioning and rejected the further cooperation with CLoSE.

accessible in Table 1.1 in Appendix. In the CLoSE sample, 17% of pupils applied to selective schools and more than half (11%) were admitted. Although girls apply at a 3.8 p. p. higher rate than boys, boys are equally successful in the admission process. Girls and boys also match in the reasons for applying. In most cases, they apply because of their parents and/or themselves. Only 7% of pupils were pushed into application by their parents.

Admission exams to selective schools are very challenging. Almost 50% of pupils who applied prepared daily in the semester prior to the exam. Girls spent more time in preparation than boys, especially by practicing model exercises or by taking extra tuition. They also enjoyed the preparation for admission more than boys. Although both girls and boys cared highly about admission exams, girls were significantly more afraid during the exam[14].

1.4.2 Assessment data

The data from CLoSE allows for the merging of pupils' achievement with their decision to apply for selective school and their success during the admission process. In CLoSE two assessment measures are available. The first measures math and reading skills[15] tested in the 4th grade in TIMSS and PIRLS, respectively. The second measure is final grades from the first semester of the 5th grade from math, Czech and foreign language. In the Czech Republic, the academic year is formed by two semesters, at the end of which pupils receive a final grade from each subject. These final grades attain values on a 5-point scale from 1 – the top grade – to 5 – a score insufficient for transition to the next grade of schooling. Descriptive statistics for the two assessment measures are presented in Table 1.1 in Appendix.

On average, boys significantly outperform girls in math skills by 0.17 standard deviations and girls significantly outperform boys in reading skills by 0.10 standard deviations. If teachers assign grades only according to these two skills, we should observe better grades in math for boys and in Czech for girls. However, girls received significantly better grades than boys from both subjects.

[14] Although girls were more stressed during admission exams, it has no impact on the admission decision. Controlling for cognitive skills measured by TIMSS and PIRLS test score and for primary school GPA, the probability of admission was the same for girls and boys.

[15] Math and reading test scores are normalized to mean 0 and standard deviation 1.

Hence, gender differences in achievement test scores do not correspond with gender differences in assigned grades.

The magnitude of test scores – measured by standard deviation – by which girls and boys differ for particular grades is documented in Table 1.2. Although almost the same proportion of girls and boys have grade 1 from math (around 40%), boys with grade 1 still perform significantly better in math by 0.16 standard deviations than girls. For boys and girls with grade 2 in math, the difference is even higher and equals 0.27 standard deviations. The opposite situation occurs with grades from the Czech language. For pupils with the top grade in Czech, there is no gender difference in test scores. This evokes a consideration that grade 1 in Czech is assigned to boys and girls according to reading test scores[16]. But, for pupils with grade 2 in Czech, boys again significantly outperform girls in reading test scores by 0.12 standard deviations. In other words, grades are assigned to pupils not only according to achievement scores but also according to other skills.

Table 1.2: Representation of grades and gender difference in test scores for particular grades

	Math				Czech (reading test scores)			
Grade	Boys	Girls	Δ (B-G) in scores		Boys	Girls	Δ (B-G) in scores	
1	41%	44%	0.16	(0.06)	26%	43%	-0.00	(0.08)
2	39%	40%	0.27	(0.05)	45%	41%	0.12	(0.06)
3	15%	12%	0.22	(0.13)	22%	13%	0.20	(0.10)
4	4%	3%	-0.08	(0.22)	6%	3%	0.20	(0.22)
5[1]	1%	1%	0.23	(0.38)	1%	0%	-0.26	(0.55)

[1] The insignificant difference in gender gap in test scores for grades 4 and 5, i.e. the worst grades, is probably caused by the small number of pupils within these grades.
Source: Own calculations based on data from CLoSE.
Note: Standard errors are in parenthesis.

Although it is not clear which particular skills enter into these other components of grades, the analysis above shows that girls outperform boys in these skills. Moreover, they affect especially grades in math, whereas in the Czech grades it is apparent only for grades other than 1. The gender difference in these unknown components of grades is crucial for further analysis. If

[16] No gender difference in assignment of grade in Czech according to cognitive skills may be understood also in the sense that boys in this group form the same perceived probability of admission as girls, and hence as they receive the same signal from grades. Examining the application decision particularly for this group of pupils, I found no gender difference even after controlling for math and reading test scores. Recall, that the gender gap in application decision for the whole sample reaches 3.6 p.p. and controlling for cognitive skills it is 6 p.p.

they are important for selective schools, and hence influence the grades and admission decisions in a similar way, their gender difference should not induce the gender gap in application to selective schools[17]. However, if these other components of grades are irrelevant for selective schools, leaning on grades in the application decision can lead to gender inequality in applications conditional on admissions.

The detailed distribution of grades in math and Czech by gender is depicted separately for quartiles of reading and math skills in Figure 1.1 and Figure 1.2, respectively. Regarding reading skills, 71% of girls in comparison to only 52% of boys in the top quartile earn the best grade in Czech. This difference in grades persists through all the quartiles of reading skills. The gender difference in grade assignment within ability quartiles is lower for math, although girls still achieve the top grades at a higher rate than boys in the same quartile. The very high representation of best grades in lower quartiles can be explained by the fact that there is no national guideline for grading[18], and thus, low achievers can also attain the top grade. Moreover, it again suggests that grades also follow other skills than those presented by the TIMSS and PIRLS test score distribution, and these give girls the advantage in achieving better grades than boys in the same ability quartile.

This is in line with previous studies which show grades also to be formed by noncognitive skills, and the distribution of these for girls is to the right of that for boys[19] (Matějů and Smith, 2014; Seligman and Duckworth, 2006). The different amount of cognitive skills needed for particular grades for girls and boys is crucial for the signal that pupils observe from their grades, and consequently for their probability of being admitted and their decision to apply. Hence, gender

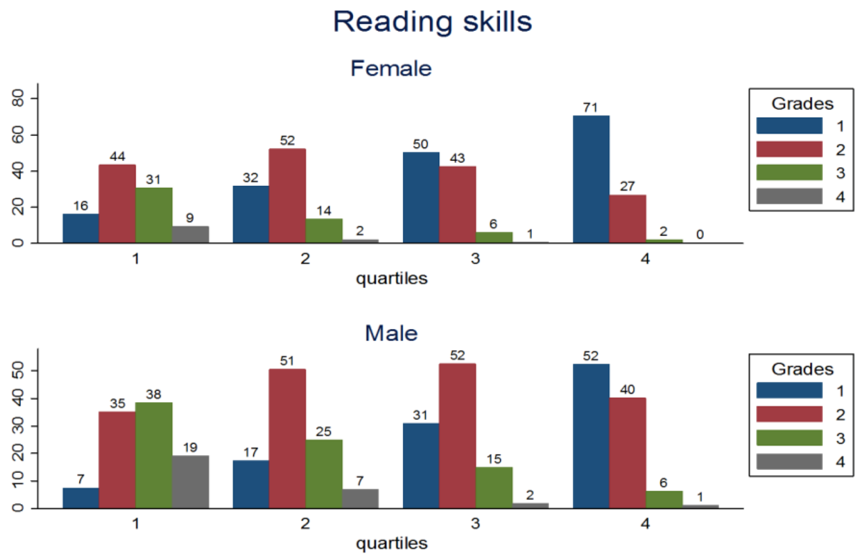
[17] Here, I assume that pupils make their application decision according to grades.

[18] There are only general rules for teacher's evaluations. For detailed information see the decree of the Ministry of Education, Youth and Sports in the Czech Republic, n. 48/2005 (§14-17).

[19] If there is any gender disparity between international testing and curricula taught at school, this can overestimate the gender difference in grades assignment conditional on achieved test scores. This situation may occur if TIMSS and PIRLS test different skills that are taught at school, and these are aimed at a particular gender. Since TIMSS and PIRLS are based on curricula knowledge, there should only be a negligible difference between abilities measured by these tests and in class. Other sources of gender disparity may arise from the gender specific test-taking environment. In high-stake exams, girls usually underperform boys due to anxiety they experience during testing (Niederle and Vesterlund, 2010). Hence, the test anxiety can lead to lower test scores for girls according to their real abilities. However, the anxiety effect disappears in a less intense competitive environment (Jurajda and Münich, 2011). As TIMSS and PIRLS are low-stake testing and have no impact on grades or school transition, I assume the girls' performance is not affected in this particular test-taking environment.

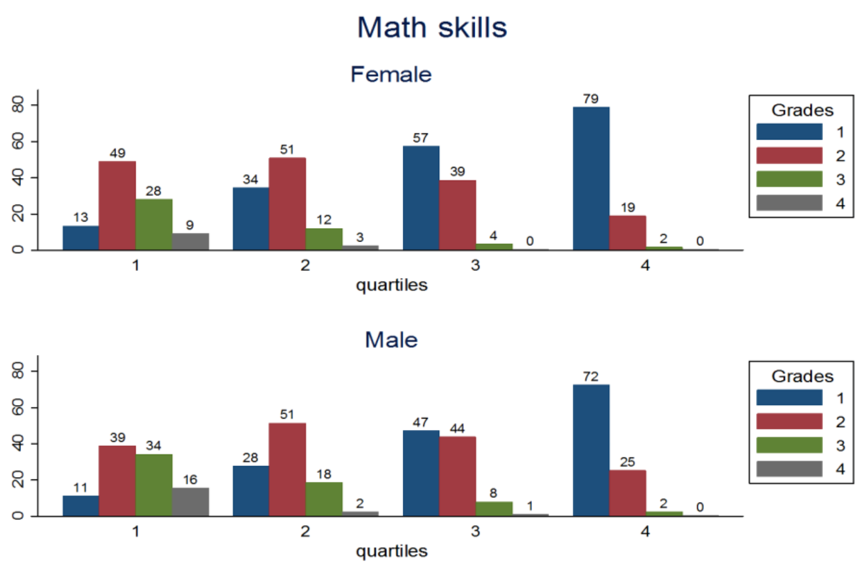
difference in grading, conditional on test scores observed in our sample can be one of the sources of gender difference in application to selective schools.

Figure 1.1: Distribution of grades in Czech according to quartiles of reading test scores, by gender



Source: Own calculations based on CLoSE.

Figure 1.2: Distribution of grades in Math according to quartiles of math test scores, by gender



Source: Own calculations based on CLoSE.

1.5 Methodology

The gender difference in application to selective school can be explained by two potential causes. Firstly, girls have a higher probability of admission, so they also apply at a higher rate than boys. In this case, the probability of admission distribution for girls lies to the right of that for boys. The second cause considers the gender difference in grading. If, conditionally on the probability of admission, the average grade for girls is higher than the average grade for boys, weighting the application decision on grades would deter boys from applying, in comparison with girls with the same admission chances. These two causes are not mutually exclusive as grades directly enter into the admission decision and hence affect the probability of admission.

The analysis firstly targets the gender difference in application decision and the ability measures that can affect it. To examine further the application decision conditional on the probability of admission the analysis turns to the predicted probability of being admitted. Recall that the transition to selective school is not mandatory, and usually only 20% of the cohort in a selection year applies and slightly more than half of them is admitted to selective schools. The estimation of the probability of admission is thus restricted for pupils who decided to apply for selective schools. According to the conceptual framework, the probability of admission is explained by the two admission criteria, GPA and admission exam scores:

$$P(admission = 1) = \Phi(\alpha_0 + \alpha_1 GPA + \alpha_2 T + \epsilon), \quad (1)$$

where *GPA* is a grade average from three subjects, i.e. mathematics, Czech and foreign language, and *T* is the admission test score. As the results from admission exams are not available in data, the math and reading test scores from TIMSS and PIRLS, respectively, are used as a proxy for admission test scores. These scores are normalized to the zero mean and standard deviations equal to 1. To address measurement error in test scores I use an instrumental variable approach. The mathematics test score is in the model instrumented by science test scores measured by TIMSS in the 4^{th} grade. Concerning reading skills, PIRLS divide reading test items by the process of reading comprehension to: a) interpreting, integrating, and evaluating, and b) retrieval and straightforward inferencing items[20]. In the model (1), test scores from items using straightforward inference are instrumented by items using interpreting process.

As all variables that enter to the admission decision model (1) are available for the whole sample, the admission probability can also be predicted for students who did not apply. This enables the estimation of pupils' application decisions by taking into account their probability of admission. The conceptual framework suggests that pupils apply to selective schools according to their individual aspirations and the perceived probability of being admitted. Following this framework the gender difference in application is modelled as:

$$P(apply = 1) = \Phi(\beta_0 + \beta_1 Boy + \beta_2 P(admission)^e + \beta_3 X + \rho), \tag{2}$$

where *Boy* is a dummy variable equal to one for boys and zero for girls, $P(admission)^e$ is the predicted probability of admission from model (1), and X is the set of individual characteristics, such as age, the number of books at home (set of dummies for particular intervals of the number of books), and parental education (two dummies for whether each completed tertiary education). Here, the coefficient of interest β_1 estimates the difference in application probability for girls and boys with the same predicted probability of admission and aspirations modelled by individual characteristics. In the conceptual framework, the possible non-zero estimate of β_1 is explained by the gender difference in perceived probability of being admitted. This occurs if pupils rest their expectations of admission scores on measures that disadvantage one gender compared to the real admission scores.

To examine whether the gender difference in application conditional on predicted probability of admission is due to grades, I decompose gender differences into those within admission probability deciles and those between deciles. This approach particularly addresses pupils with very high probability of admission, and likely also the very high probability of application. In order to gain a better understanding of the underlying sources of any gender differentials within deciles,

---

[20] For detailed description of reading test items and their division according to a comprehensive process, see Mullis, Martin, Kennedy, Trong, and Sainsbury (2009).

I further estimate the extent to which grades account for those differences. Thus, the probability of application conditional on predicted probability of admission is estimated as:

$$P(apply = 1) = \Phi(\gamma_0 + \gamma_1 Boy + \gamma_2 Admit + \gamma_3 Boy * Admit + \gamma_4 GPA + \sigma),$$

where *Admit* is the set of dummy variables for admission probability deciles, and GPA is, in this specification, a set of two dummy variables equal to one for pupils with grade 1 from math and Czech[21]. Controlling for the probability of being admitted, coefficients of grades should be insignificant. Moreover, if the signal of a grade does not cause gender disparity in the application decision conditional on the predicted probability of admission, i.e. the hypothesis is not true, the coefficient γ_3 should not change by including grades in the regression.

Since the school transition is aimed at top pupils, some pupils have an almost zero chance of being admitted and hence do not consider applying. As the functional form of estimated models can be sensitive to the values of pupils at the low edge, I omit these pupils from analysis. These are pupils with math and reading test scores more than 0.75 standard deviations below average and pupils with grade four, i.e. the worst grade, from math or Czech. Although considering the whole sample does not yield considerably different estimates, I use this subsample in the following analysis.

1.6 Results

First, the analysis examines the gender gap in the application decision and the available achievement measures that can explain it. However, the gender gap in application may be only the consequence of a gender difference in the probability of admission. Thus, the analysis further focuses on the application decision conditional on the predicted probability of admission. Finally, I show the role of grades in explaining the gender difference in application conditional on predicted probability of admission. Particular attention is dedicated to pupils in the top admission deciles, i.e. pupils with the highest probability of being admitted, and thus, with the highest probability of applying.

[21] 99% of pupils who applied to selective schools got grade 1 or 2 from mathematics and 98% from Czech.

1.6.1 Application decision

The analysis is firstly focused on gender difference in application rate and the ability measures that affect it. Table 1.3 reflects successively the impact of grades and test scores on the decision to apply. All specifications presented in separate Columns of Table 1.3 show that the probability of applying is a function of both grades and test scores. However, these two achievement measures differ in the rate to which they explain the gender difference in application. The baseline model presented in Column (1) confirms that on average girls apply to selective schools at a significantly higher rate than boys by 3.6 p.p. Controlling for math and reading test scores in Column (2) does not explain this gender difference at all. In fact, the gender difference in application is even higher and reaches 6 p.p. On the contrary, this gender difference in application is fully explained by grades (Column 3).The latter also holds after including both assessment measures to the regression showed in Column (4). The linear probability model presented in Column (5) leads to similar results as the probit model[22].

Table 1.3: Application decision to selective school, marginal effects after probit

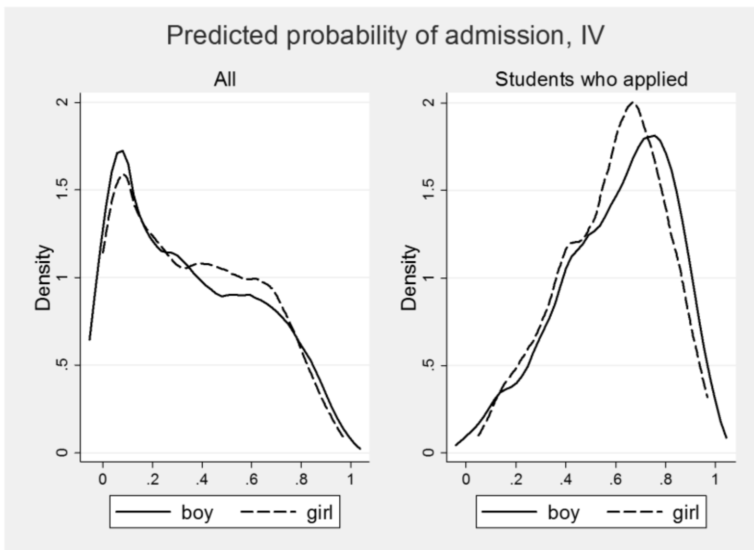
Application	(1) probit	(2) probit	(3) probit	(4) probit	(5) LPM
Boy	-0.036**	-0.060***	0.004	-0.021	-0.017
	(0.016)	(0.016)	(0.017)	(0.017)	(0.015)
Grade math=1			0.178***	0.109***	0.099***
			(0.020)	(0.021)	(0.020)
Grade Czech=1			0.187***	0.142***	0.162***
			(0.020)	(0.020)	(0.022)
Math score		0.128***		0.083***	0.079***
		(0.013)		(0.013)	(0.012)
Reading score		0.096***		0.064***	0.063***
		(0.012)		(0.013)	(0.012)
N	3149	3049	3105	3005	3005

Note: All regressions include a set of individual characteristics described above. Standard errors in parentheses clustered on class level (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

[22] Using the interactions between grades and gender, no significant gender difference in response to grades was found. By including socio-economic variables to the model, the estimated coefficients of interest stayed mostly unchanged. These results are available upon request.

These results are consistent with the two possible explanations. The first suggests that grades are more important in the admission process, and the lower application rate of boys is just a response to their lower probability of being admitted. The second explanation is based on the signal of admission scores that grades may provide to pupils before they decide to apply. This signal may be systematically wrong in ways that lead to lower application rates of boys relative to girls with a similar probability of admission. In the next subsections, I examine the determinants of the probability of admission to gain a clearer understanding of the relative importance of each of these explanations.

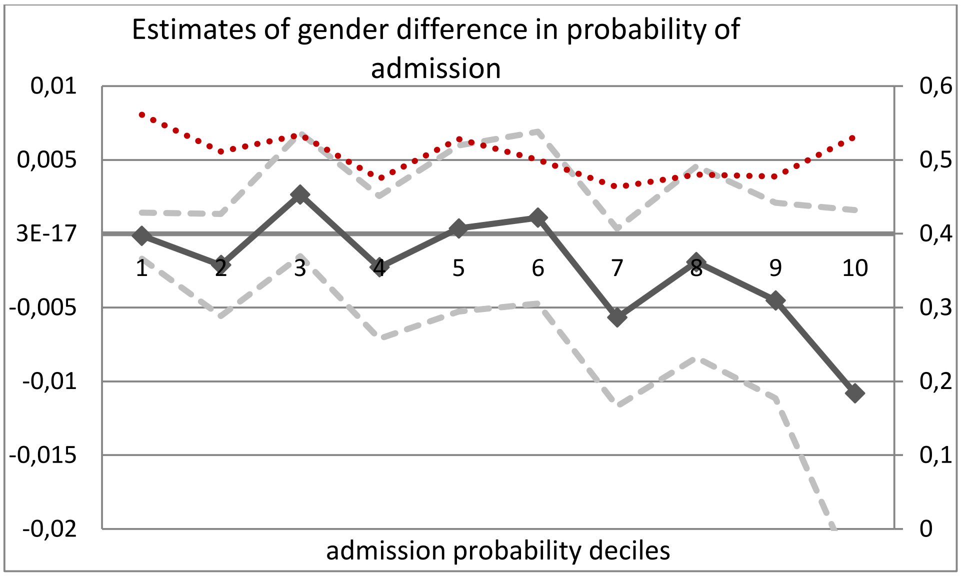
Figure 1.3: Distribution of predicted probability of being admitted, by gender and application



1.6.2 Probability of admission

The first potential explanation of gender difference in application rate is based on the gender difference in distribution of probability of admission. The admission criteria to selective schools consist of the results in the admission exam and of the primary school grade average. Using the grade average from Math, Czech and foreign language and test scores from math and reading skills, the probability of admission is firstly estimated for the pupils who applied and further predicted for the whole sample. Table 1.4 in Appendix presents the results for the model of the probability of admission. Column (1) confirms the importance of the two assessment measures in admission results. The instrumental variable approach in Column (2) – used to remove the measurement error in test scores – provides similar results. For further analysis, I use the predicted probability of admission from the instrumental variable approach.

Figure 1.4: The estimates and 95% confidence intervals of gender difference in probability of admission, by probability admission deciles (all pupils)



Note: Positive values refer to higher probability of admission for girls and negative for boys. The dotted line reflects the proportion of boys among pupils who applied in each admission decile.

The distribution of predicted probability of admission is depicted separately for girls and boys in Figure 1.3. On the left, the admission probability for the whole sample suggests a higher appearance of boys at both tails and a greater number of girls in the middle of the distribution. However, on average girls and boys do not differ in the probability of admission. The gender difference in probability of admission on average is also not confirmed for pupils who decided to apply, although the boys' distribution seems to be to the right of the girls (the graph in Figure 1.3, on the right). The gender difference in admission probability is also not confirmed within admission deciles (Figure 1.4). The dotted line in Figure 1.4 shows that girls and boys are almost evenly distributed in admission deciles. These results indicate no gender difference in the probability of admission[23]. Thus, the lower application rate of boys is not a response to their lower

[23] Although there is no gender difference in the predicted probability of admission, the measure of predicted probability of admission may have different gender specific sources. For example, the higher anxiety during testing among girls may motivate girls to increase their admission chances by getting better grades. However, comparing the change in GPA between the 4th and 5th grades among those who applied and who did not apply did not show any gender specific differences. Further analysis also failed to demonstrate that cognitive skills – measured by reading and math test scores

admission chances as is suggested in the first possible explanation of the gender gap in application rate.

1.6.3 Probability of application conditional on probability of admission

In the previous section, the analysis showed that the probability of application is significantly higher for girls than for boys. However, they do not differ in their distribution of admission probability. In the following section, I examine whether the gender difference in application rates persists after controlling for the predicted probability of admission. Figure 1.5 depicts the girls' and boys' application rates within the admission probability deciles. According to what might be expected, the application rate declines gradually from the top deciles, but with several differences between girls and boys. In the top decile, girls and boys apply to selective schools with the same probability of 60%. The gender difference occurs in the next two lower deciles in which boys apply less often than girls almost by 10 p.p.

The three top deciles are crucial for application to selective schools, as usually only 20% of pupils decide to apply and 10% are admitted. Using a finer division[24] of the admission probability in Figure 1.6, the deeper decline in probability of application can be observed between the top 15 and 25% of pupils, again with considerable differences for girls and boys. While the sudden drop in the application rate arises for girls after the top 25% of pupils, for boys it is after the top 15% of pupils. This provides evidence indicating that boys correspond differently to their real probability of admission than girls. This difference may be explained by the gender biased signal that pupils receive from grades in order to predict their admission score, and further, to form their own perceived probability of admission.

– and GPA have different weights for admission decision for girls and boys. In other words, girls have the same probability of being admitted as boys with the same test score, i.e. higher stress during exams seems to have no effect on girls' admission chances given their true cognitive skills. These results are available upon request.

[24] The further results are not sensitive to the finer division of admission probability, and therefore, the next analysis is presented only for deciles.

Figure 1.5: Application rates in the predicted admission probability deciles, by gender

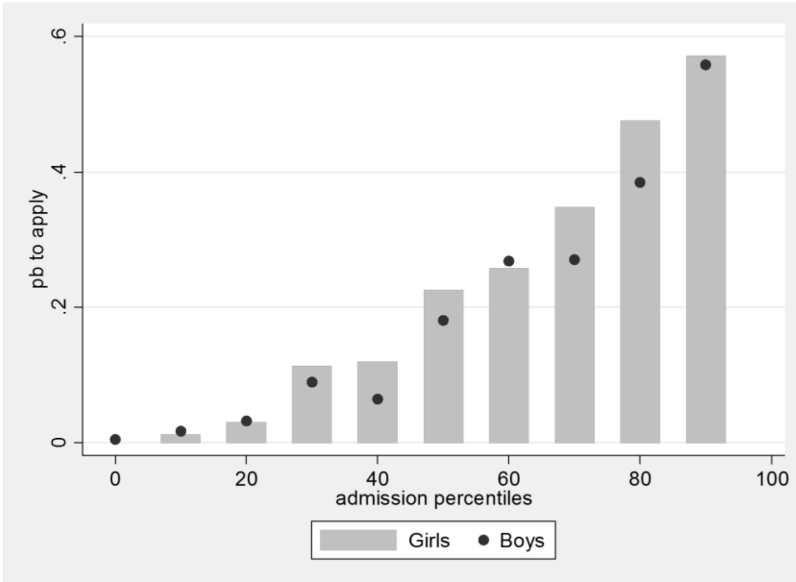
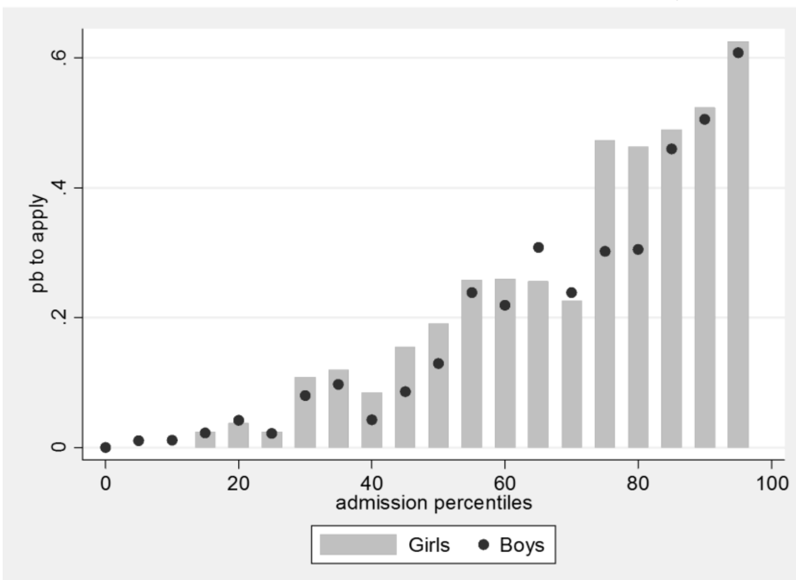


Figure 1.6: Application rates in the predicted admission probability percentiles, by gender



The gender difference in application conditional on admission observed in Figure 1.5 and 1.6 also persists on average (see Table 1.5 in Appendix). The admission probability is expressed in the regression by the dummy variables for each decile, excluding the two low deciles as a base characterized by almost zero application rate and hence zero gender gap in application. As in the previous results, the application rate grows gradually across the admission deciles and this holds for all specifications. The model presented in equation (2), i.e. the application regression controlling only for the predicted probability of admission and individual characteristics, is estimated and its results are presented in Column 1. The results reveal that gender difference in application persists also after controlling for predicted probability of admission (see Table 1.3). In the next columns of Table 1.5, the estimates of interest are the interactions between admission deciles and the gender variable. Controlling only for individual characteristics in Column 2, in the top decile, girls and boys react to the admission chances by the same application rate, while in the subsequent decile – the 9th decile – girls are significantly more likely to apply to selective school by 9%. To examine whether grades or scores can explain this gender difference, I control for these variables in Column 3 and Column 4, respectively. Conditional on grades, the gender difference in the 9th decile drops and becomes insignificant, while conditional on test scores it stays at the 9%. These results are again consistent with grade acting as a signal for the application decision. This signal does not fully correspond to the admission chances in a way that makes it disadvantageous for boys to apply.

This specification however also considers pupils with a zero chance of admission, and hence, with zero aspirations to apply. Further analysis is therefore focused on pupils in the top admission deciles.

1.6.4 The role of grades in application decision

Although girls and boys at the top admission decile (10th decile) apply at the same rate – i.e. respond to their real probability of admission similarly –, this does not hold for pupils at the boundary, i.e. pupils in the 9th and 8th admission decile. A potential explanation is offered by the role of grades in decision making. Firstly, Table 1.6 examines the effect of grades assignment on the application decision within the admission deciles. Comparing the application rates among pupils with the best grades from both subjects, the gender difference in application rate disappears in all three top admission deciles[25]. Thus, if girls and boys earn the best grades from math and Czech they decide to apply to selective school at the same rate in the three top admission deciles.

---

[25] Similar results come from comparison of girls and boys with the same grade in Czech. In this group, boys actually apply by 1.4 p.p. more often to selective schools than girls. This group is specific in the way that these boys and girls do equally well in available reading test score (Table 2). Thus, the boys with the best grade in reading should have the same other skills not tested in the achievement test than the girls with the best grade in Czech; and hence, girls and boys in this group should be affected by the grade in Czech equally. This can explain the slightly higher application rate for boys with the best grade in Czech.

In other words, girls and boys seems to follow the signal of grades to predict their own admission probability. Moreover, using this signal they further form their application decision equally.

Table 1.6: Application rates to selective school in admission deciles, by gender and grades

		Predicted admission deciles							
Grade		7		8		9		10	
Math	Czech	Girl	Boy	Girl	Boy	Girl	Boy	Girl	Boy
Yes*	Yes	0.31	0.42	0.35	0.33	0.51	0.46	0.59	0.59
Yes	No	0.16	0.19	0.36	0.25	0.21	0.31	0.60	0.33
No	Yes	0.24	0.33	0.39	0.25	0.60	0.00	0.20	-
No	No	0.24	0.20	0.25	0.20	0.00	0.27	0.00	0.50
Overall		0.26	0.27	0.35	0.27	0.48	0.38	0.57	0.56

*Yes report grade 1 and No grade other than 1

Note: Data for group with frequency less than 15% are in grey.

Table 1.7: Frequencies in admission deciles, by gender and grade

		Predicted admission deciles							
Grade		7		8		9		10	
Math	Czech	Girl	Boy	Girl	Boy	Girl	Boy	Girl	Boy
Yes*	Yes	0.45	0.29	0.60	0.37	0.83	0.57	0.93	0.85
Yes	No	0.20	0.35	0.20	0.35	0.10	0.31	0.03	0.13
No	Yes	0.13	0.06	0.10	0.07	0.05	0.03	0.03	0.00
No	No	0.22	0.30	0.11	0.21	0.02	0.09	0.01	0.02
Overall		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Yes report grade 1 and No grade other than 1

Thus, the gender gap in the application decision may occur if the best grades are distributed differently for girls and boys in the admission deciles. To detect this channel of gender gap in application, Table 1.7 describes the distribution of grades in admission deciles for girls and boys separately. In the top admission decile, the variation in grades is very low. Nearly all, i.e. 93% of girls and 85% of boys, achieve the best grade from math and Czech language. As grades are mostly the same for these girls and boys, the gender biased signal of grades does not create the gender difference in perceived probability of admission, and thus, neither in the application rate. On the contrary, the variation in grades for girls and boys is broadening out in the lower deciles in a way that the probability of obtaining the best grades is, for boys, diminishing sharply with descending admission deciles. This variation causes a systematically wrong signal for the application decision in ways that lead to lower application rates of boys relative to girls with similar admission probability. Particularly, in the decile second from the top, 83% of girls earned the best grades from both subjects while there are only 57% of boys with the best grades. In the third from the top decile (8th decile), the gender difference in grade assignment is similar to the 9th decile, here with 60% of girls compared to 37% of boys with the best grades from math and Czech language. This gender difference in grading persists also after controlling for test scores (Figure 1.7 in Appendix). In the 10th and 9th admission decile, the probability of achieving the best grade from both subjects is significantly higher for girls than for boys with the same test scores by 7 p.p. and 14 p.p., respectively.

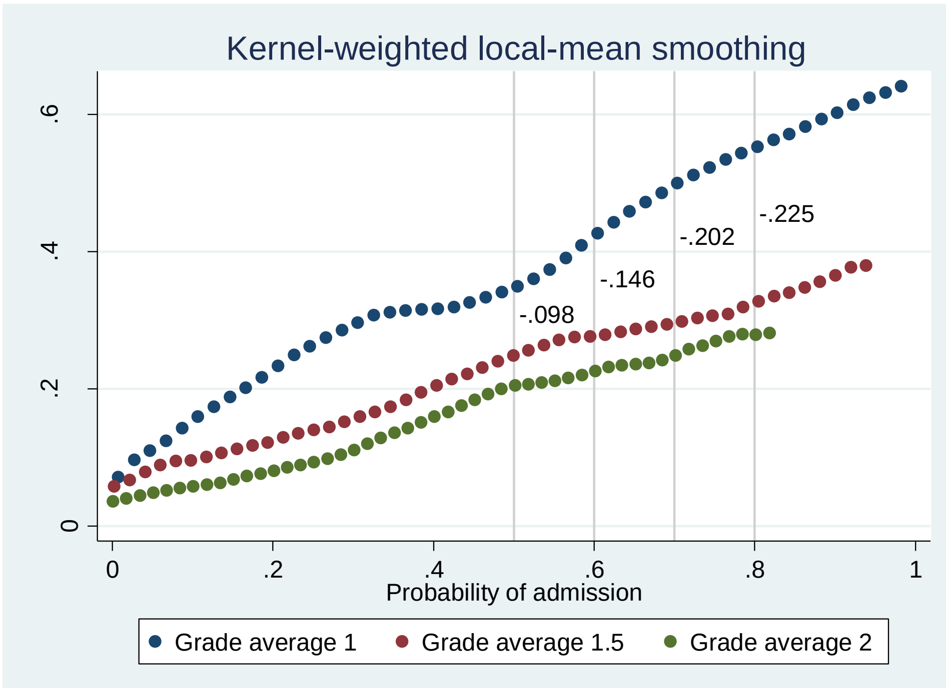
Thus, considering the top three admission deciles, in these the gender difference in application rates correspond to the gender difference in grades assignment. This is persistent also after controlling for test scores. Hence, the results are consistent with skills not tested in available test scores influence grades, and as a consequence, bias their signals in a way that skews the pupils' beliefs of their own admission chances and may cause the gender gap in application decision.

1.6.5 Robustness checks

The previous analysis is based on the assumption that grades play an important role in decision making, especially because of the lack of any other signals about pupils' skills. Two further approaches are applied to strengthen this assumption. Firstly, I show the importance of grade on application conditional on admission. Then, two features of the admission process are used to reveal that pupils' decisions to apply rely on grades.

By using the Kernel smoothing of the application decision, Figure 1.8 depicts the probability of application for pupils with different GPA conditional on admission probability. Here, GPA is computed as an average of grades from math and Czech. The application rate is again increasing with higher probability of admission. However, the growth in application rate is steeper for pupils with GPA equal to 1, i.e. with the best grade from both subjects, relative to pupils who earn the worse grade from one of these two subjects and the best grade from the other. These pupils achieve a GPA equal to 1.5. Their application rate is similar to pupils with GPA 2, i.e. with a worse grade from both subjects, at each point of a predicted admission probability. The difference in slopes among pupils with various GPA's induces the broadening of disparity in application rate with increasing admission probability. Thus, the gap in application rate between pupils with GPA equal to 1 and 1.5 yields 15% for pupils with a predicted admission probability of 0.6 and even 20% for pupils with a predicted admission probability of 0.7. In other words, although the probability of admission is the same for these pupils, the ones with at least one worse grade apply at an extensively lower rate. These results provide evidence indicating that grades highly predict the application decision even after controlling for predicted probability of admission.

Figure 1.8: Application rates conditional on probability of admission, by GPA



I further explore the emphasis that pupils put on grades using the two features of the transition to selective schools in the Czech Republic. In the Czech Republic, the admission criteria are not imposed by the government, and hence, each school can set their own weights on grades and admission exams in the admission decision. On the average, admission decision is based from 25% on grades and 75% on admission exams. However, schools differ considerably in admission criteria. Figure 1.9 in Appendix depicts the variation in weights placed on grades that academically selective schools adopt for their admission criteria. The Figure represents the weights on grades on district level. Controlling for the weighted average[26] of weights placed on grades in each district, I find no difference in the influence of grades on the application decision (Table 1.8 in Appendix). In other words, pupils in districts with very low weight on grade in the admission

[26] The weight on grades in the admission process for each district is constructed as a weighted average of weights on grades from all selective schools in the district. Schools are weighted according to the number of students admitted.

criteria rely on grades at the same rate as pupils in districts with a higher weight on grades. This result indicates that even if the weight on grades in the admission decision is negligible, pupils still use their grades to predict their probability of being admitted, and hence, to decide whether apply or not. This may be explained by the lack of any other information about admission chances that pupils possess other than grades[27].

The second feature of school transition is based on the construction of admission exams. In 2012, 45% of selective schools used the standardized admission exams provided by a private company SCIO. The company offers the possibility of trying their type of admission exams in advance and of comparing a pupil's own results with the real results of admission exams from previous years. Thus, even before pupils decide to apply they can appraise their own admission chances without relying on grades. Using the variation across districts in administrating Scio exams, I examine the difference in the role of grades in the application decision for pupils who are or are not exposed to Scio exams. The Czech districts are divided to three groups according to Scio exams: districts with less than 10% of schools, districts with between 10 to 90% of schools, and districts with more than 90% of schools administrating Scio exams[28]. The results are presented in in Appendix in Table 1.9. Conditional on Scio exams, pupils do not apply at different rates (Column 1). The more important result in Column 2 suggests that pupils do not put significantly higher emphasis on grades in districts with no Scio exams in comparison to districts in which all schools use Scio exams for admissions. It indicates that the possibility of appraising the pupils'

---

[27] The result in Column 1 of Table 8 suggests that with increasing weights on grades the application rate declines significantly. There are two potential explanations. Districts with higher excess demand for selective schools put lower weights on grades because of insufficient variation in grades among applicants to make the admission decision only according to grades. Using the excess demand, the effect of weights on grades on the application decision is reduced to half. On the other hand, pupils in districts with higher emphasis on grades can be deterred from applying if teachers in these districts impose higher demands on grades. Examining the probability of earning the best grade from math and Czech, this probability is significantly higher for pupils in districts with very small weights on grades in the admission process than for pupils in districts with a high emphasis on grades. Conditioning on cognitive skills the difference reaches 10%. This result supports the importance of grades in the application process to selective schools. The results from both analyses are available upon request.

[28] The percentage of schools in districts administrating Scio exams is computed as a weighted average of schools administrating these exams to all selective schools in districts. Schools are again weighted according to the number of students admitted. Figure 1.10 in Appendix depicts the variation in Scio exams across Czech districts.

admission chances even before the application decision does not affect the application process to selective schools and pupils still believe grades and rely on them in the same rate[29].

1.7 Conclusion

In this study, I evaluated the gender difference in application to selective schools. It was shown that although girls and boys have the same distribution of admission probability girls apply at a significantly higher rate than boys. The main potential explanation of this gender specific application behaviour is based on the lack of information provided to pupils in a transition year. Especially at early grades, pupils' perceptions of their own academic skills rely mostly on teacher evaluations. However, no general guidelines regulate the grading standards, and thus, teachers are prone to subjective judgements (Hoge and Coladarci, 1989). Adding to this the fact that in some countries there is no nation-wide standardized testing, the majority of pupils have only a vague notion of how good they are compared to the rest of their peers in the country. Similarly, Bennett et al. (1993) conclude in their study that "the data reinforce the need to supplement teacher judgements with other objective evidence of academic performance when important decisions about students are to be made." (p. 353).

The goal of this paper is not to set a framework for grading or to criticize how teachers evaluate pupils in the class, but to refer to the problem of inefficient allocation of pupils among school tracks assuming that admission probability is a good measure of the expected return to selective schooling for boys and girls. In the conceptual framework, I show that in the case of lack of information, relying on grades disadvantages boys in the application to selective schools. As boys achieve worse grades than girls – partly because of noncognitive skills in which boys lag behind girls – they form their perceived probability of being admitted to the left of that for girls conditional on real admission probability. Thus, some boys do not apply although they should, and some girls apply although their real chance of admission is low.

Using the longitudinal study CLoSE in the Czech Republic, this paper reveals the role of grades in explaining the gender difference in application. Controlling for the probability of

[29] The insignificant results for Scio exams could be caused by the fact that Scio exams are available to the public, and hence, even pupils in districts without Scio exams in admission can try to do them. Thus, their presence in a district may not signal any difference in relying on grades in the application decision.

admission, girls apply to selective schools at a significantly higher rate than boys, and the difference is even higher at the right tail of distribution of admission probability. On the other hand, there is no gender difference in the application rate for girls and boys with the best grades from math and Czech. The best grades from subjects tested in admission exams seems to be the most important signal for pupils' decisions to apply to selective schools. This applies to girls as well as to boys. However, the allocation of best grades varies extensively between girls and boys at the right end of the admission distribution. Here, the probability of achieving the best grade from math and Czech is significantly lower for boys by 7 to 14 p.p. than for girls even conditional on cognitive skills. Thus, this paper explains the significant part of the gender difference in application rate by the gender difference in grading conditional on test scores and by the unreasonably high emphasis on grades in the application decision.

This result addresses important policy questions about the effective allocation of pupils to school tracks. Already Hastings and Weinstein (2008) point to the importance of information for school choice. They find that "providing parents with direct information on school test scores resulted in significantly more parents choosing higher-scoring schools for their children."(p. 1375). The results in the paper suggest that providing pupils with more adequate information than grades about their own admission chances could reduce the gender gap in application rate to selective schools.

1.A Appendix 1

Table 1.1: Descriptive statistics

	All pupils				Pupils who applied			
	All	Girls	Boys	Δ (B-G)	All	Girls	Boys	Δ (B-G)
Apply to selective school	16.9	18.9	15.1	**-3.8 [1.23]**				
Admitted	10.7	11.9	9.6	**-2.3 [1.01]**	63.2	62.9	63.6	0.7 [3.57]
Math TIMSS test score	0 (1)	-0.09 (0.97)	0.08 (1.02)	**0.17 [0.04]**	0.71 (0.80)	0.60 (0.79)	0.85 (0.79)	**0.25 [0.06]**
Reading TIMSS test score	0 (1)	0.05 (0.99)	-0.05 (1.01)	**-0.10 [0.05]**	0.70 (0.81)	0.71 (0.79)	0.69 (0.83)	-0.02 [0.06]
Grade Math*	1.81 (0.86)	1.78 (0.82)	1.84 (0.89)	**0.07 [0.03]**	1.19 (0.41)	1.18 (0.39)	1.21 (0.43)	0.02 [0.03]
Grade Czech*	1.96 (0.87)	1.78 (0.81)	2.12 (0.90)	**0.34 [0.03]**	1.30 (0.52)	1.21 (0.42)	1.42 (0.60)	**0.21 [0.04]**
Father – university education (%)	18.3	17.8	18.9	1.1 [1.3]	46.3	46.0	46.7	0.7 [3.7]
Mother – university education (%)	14.8	14.6	15.0	0.4 [1.2]	35.9	33.8	38.2	4.4 [3.6]
Positive attitudes towards learning (%)								
Self-confidence	47.7	45.3	50.0	**4.7 [1.6]**	72.1	70.3	74.3	4.0 [3.3]
Positive attitudes	55.1	52.8	57.2	**4.4 [1.6]**	76.1	75.2	77.1	1.9 [3.2]
Perseverance	59.2	59.2	59.2	0.0 [1.6]	73.4	74.8	71.7	-3.2 [3.3]
Liking mathematics	60.3	54.6	65.7	**11.1 [1.6]**	65.1	63.6	66.8	3.3 [3.5]
Liking Czech language	33.0	39.1	27.3	**-11.8 [1.5]**	42.2	51.4	31.3	**-20.1 [3.6]**

Table 1.1: Descriptive statistics - continued

	All pupils				Pupils who applied			
	All	Girls	Boys	Δ (B-G)	All	Girls	Boys	Δ (B-G)
Preparation for admission exams (at least once a week) (%)								
At home					73.6	73.9	73.3	-0.6 [3.3]
With tutor					46.4	51.3	40.7	**-10.5 [3.7]**
Course at my primary school					14.5	12.9	16.3	3.3 [2.6]
Course at academic school					30.1	27.0	33.7	**6.7 [3.4]**
Scio tests					66.2	70.7	61.0	**-9.7 [3.5]**
Prepare daily in last semester					45.3	49.0	40.9	**-8.1 [3.7]**
Reasons for applying								
Me and parents wanted					38.2	37.6	38.8	1.1 [3.6]
My parents wanted but I did not					6.9	5.5	8.5	3.0 [1.9]
Scio tests at admission exams					64.7	63.1	66.4	3.3 [3.5]
Number of observations (%)	3,682	1,817	1,865		737	395	342	
		48.5	51.5			53.6	46.4	

Source: Own calculations based on TIMSS 2011, PIRLS 2011 and CLoSE.
Note: Standard errors are in brackets and standard deviations are in parenthesis. Significant gender differences at the .05 level are in bold.
*Higher mean represents worse grades. Grades are on scale 1-5 with 1 as the best grade.

Table 1.4: Probability of admission to selective school, marginal effects after probit

Admission	(1) All	(2) IV[1]
Reading score	0.041*	0.044
	(0.024)	(0.027)
Math score	0.172***	0.196***
	(0.023)	(0.031)
Grade average	-0.224***	-0.201***
	(0.048)	(0.050)
N	704	704

Standard errors in parentheses (* p < 0.1, ** p < 0.05, *** p < 0.01).
[1] Instrument for reading and math test score

Table 1.5: Application decision to selective school conditional on the deciles of predicted probability of being admitted, LPM

Application	(1)	(2)	(3)	(4)
Boy	-0.033**	0.002	0.002	0.016*
	(0.014)	(0.008)	(0.008)	(0.008)
Predict. admission=3	-0.000	-0.003	-0.011	0.025
	(0.011)	(0.014)	(0.015)	(0.017)
Predict. admission=4	0.071***	0.084***	0.059**	0.122***
	(0.020)	(0.030)	(0.030)	(0.032)
Predict. admission=5	0.056***	0.082***	0.040*	0.133***
	(0.019)	(0.024)	(0.024)	(0.031)
Predict. admission=6	0.157***	0.195***	0.116***	0.252***
	(0.025)	(0.037)	(0.039)	(0.043)
Predict. admission=7	0.196***	0.194***	0.098***	0.259***
	(0.027)	(0.033)	(0.037)	(0.042)
Predict. admission=8	0.239***	0.274***	0.154***	0.352***
	(0.028)	(0.037)	(0.043)	(0.046)
Predict. admission=9	0.349***	0.390***	0.244***	0.478***
	(0.030)	(0.038)	(0.046)	(0.045)
Predict. admission=10	0.437***	0.461***	0.303***	0.585***
	(0.032)	(0.042)	(0.046)	(0.058)
Boy * Predict. ad.=3		0.004	0.012	0.006
		(0.020)	(0.020)	(0.021)
Boy * Predict. ad.=4		-0.024	-0.009	-0.023
		(0.038)	(0.038)	(0.038)
Boy * Predict. ad.=5		-0.050*	-0.031	-0.054*
		(0.029)	(0.029)	(0.029)
Boy * Predict. ad.=6		-0.074	-0.041	-0.075
		(0.046)	(0.046)	(0.046)
Boy * Predict. ad.=7		0.008	0.032	0.009
		(0.051)	(0.051)	(0.051)
Boy * Predict. ad.=8		-0.071	-0.037	-0.070
		(0.054)	(0.055)	(0.054)
Boy * Predict. ad.=9		-0.085*	-0.056	-0.086*
		(0.050)	(0.051)	(0.050)
Boy * Predict. ad.=10		-0.048	-0.037	-0.055
		(0.047)	(0.047)	(0.047)
Grade math=1	No	No	Yes	No
Grade Czech=1	No	No	Yes	No
Scores	No	No	No	Yes
N	3274	3274	3263	3274
R^2	0.256	0.257	0.271	0.262

Note: All regressions include a set of individual characteristics described above. Standard errors in parentheses clustered on class level (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 1.8: Application decision to selective school conditional on weights on grades in admission criteria, marginal effects after probit

Application	(1) probit	(2) probit
Weights 1 <0, 0.2)[1]	0.140***	0.019
	(0.041)	(0.052)
Weights 2 <0.2, 0.4)	0.071*	-0.000
	(0.043)	(0.054)
Weights 1 * Grade math		0.070
		(0.051)
Weights 1 * Grade Czech		0.089
		(0.056)
Weights 2 * Grade math		0.035
		(0.052)
Weights 2 * Grade Czech		0.092
		(0.058)
Grade math=1		0.118***
		(0.040)
Grade Czech=1		0.086*
		(0.047)
N	2604	2563
R^2		

Note: Standard errors in parentheses. (* p < 0.1, ** p < 0.05, *** p < 0.01)

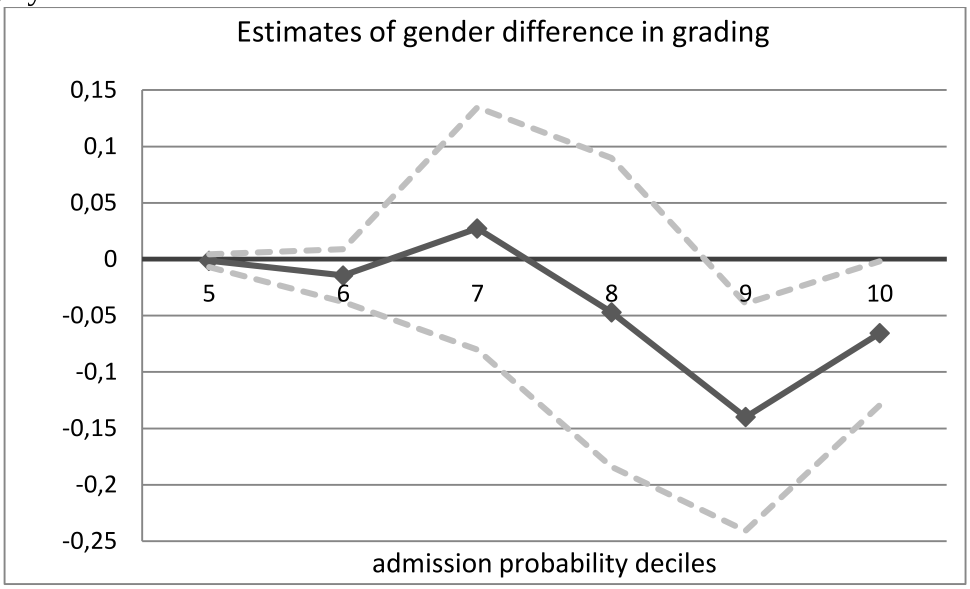
[1] This result is not sensitive to the division of interval. Several partitions were used yielding the same results.

Table 1.9: Application decision to selective school conditional on Scio exams, marginal effects after probit

Application	(1) probit	(2) probit
Scio 1 <0,0.1)	-0.031	0.026
	(0.037)	(0.048)
Scio 2 <0.1,0.9)	0.045	0.042
	(0.031)	(0.042)
Grade math*Scio 1		-0.019
		(0.053)
Grade Czech*Scio 1		-0.064
		(0.063)
Grade math*Scio 2		0.039
		(0.047)
Grade Czech*Scio 2		-0.026
		(0.051)
Grade math=1		0.150^{***}
		(0.038)
Grade Czech=1		0.199^{***}
		(0.044)
N	2604	2563

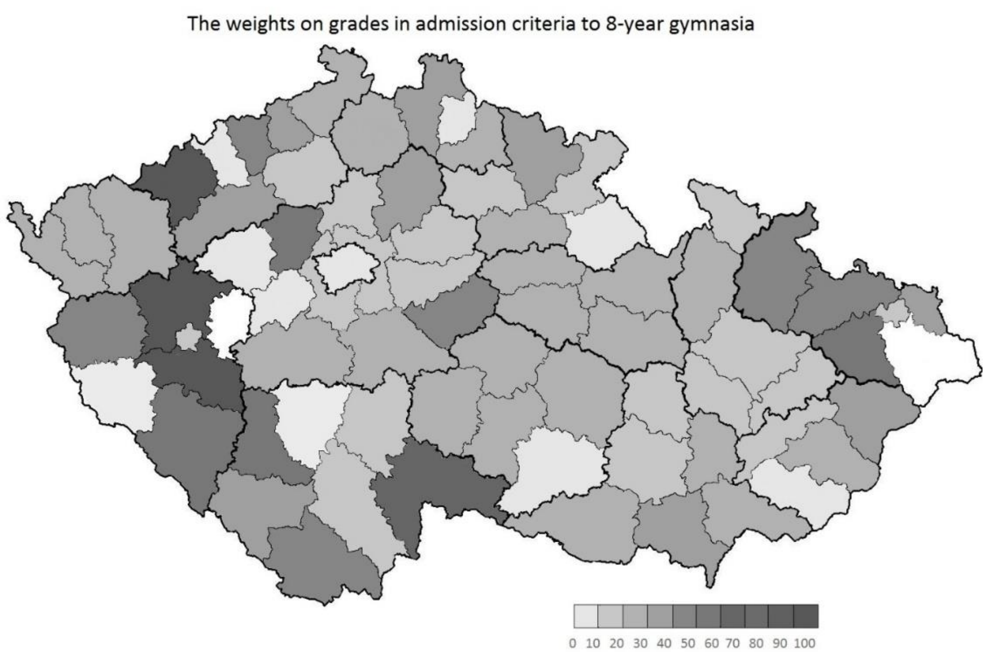
Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Figure 1.7: Estimates and 95% confidence interval of gender difference in grading conditional on test scores, by admission deciles



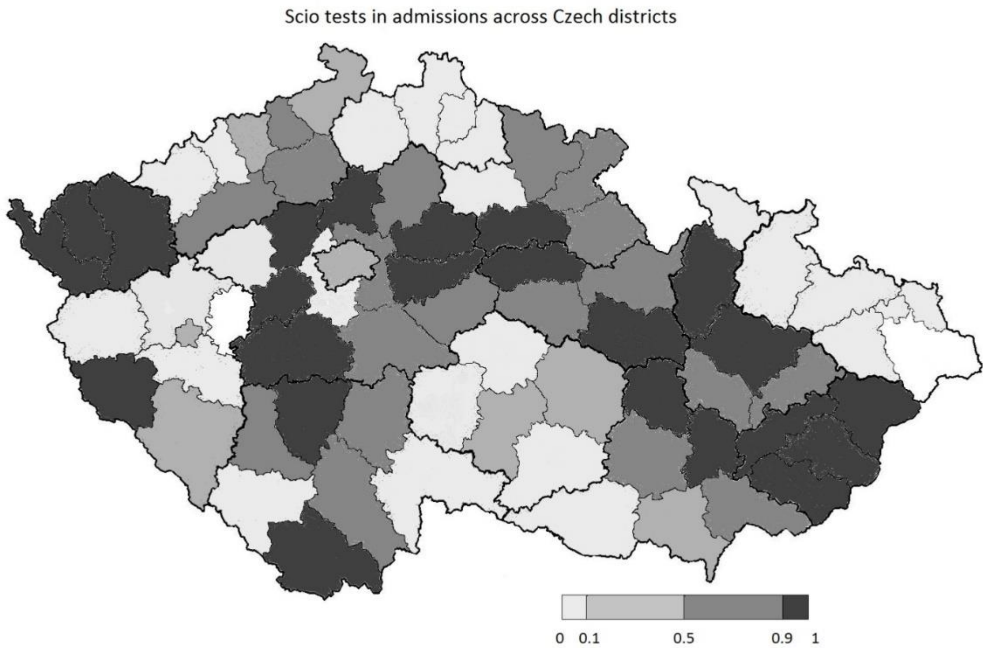
Note: Positive values refer to higher probability of achieving the best grades for girls than for boys.

Figure 1.9: The weight placed on grades in admission criteria to 8-year gymnasium, by Czech districts



Note: Two districts in white are districts with no selective school after primary education.

Figure 1.10: The weighted average of schools using Scio exams in admission decision, by Czech districts



Note: Two districts in white are districts with no selective school after primary education.

2 Chapter

The Impact of High-Stakes School-Admission Exams on Study Effort and Achievement: Quasi-experimental Evidence from Slovakia

Co-authored by Daniel Münich

2.1 Introduction

Literature in psychology, pedagogy, and sociology provides rich evidence that students' motivation to learn is a key co-determinant of their educational outcomes (Vansteenkiste, Simons, Lens, Sheldon and Deci, 2004; Elliot, McGregor and Gable, 1999). Grading and evaluation schemes are typical examples of external incentivizing factors that could motivate students to learn, and increase their achievement. The impact of these schemes on student achievement is dependent on whether the students perceive the assessments and evaluation exercises as relevant for their future, or as a mere task to be fulfilled (Ryan and Deci, 2000). One example of such external factors is high-stakes testing, which is considered an important instrument for enhancing students' motivation and educational accountability (Jacob, 2005). Most empirical studies have focused on exit exams at the end of key schooling stages—as an example of high-stakes testing—and most of these studies demonstrate an increase in student effort and academic achievement[1].

[1] See Jürges, Schneider, Senkbeil, and Carstensen (2012); Jürges, Schneider and Büchel (2005); Jürges and Schneider (2010), Piopiunik (2014); Woessmann (2002); Bishop (1997).

However, empirical evidence of the effect of exit exams does not distinguish between the effects of students' increased motivation to learn and increased effort on the part of teachers and school administrators, who are also incentivized (Jürges, Schneider and Büchel, 2005).[2] The results of school exit exams are often publicly known, closely watched, and create strong incentives for teachers and principals. There is abundant evidence that teachers tend to focus on teaching their students the skills that will be tested directly, at the expense of cultivating more complex skills, which are not tested directly or cannot be tested easily[3]. Moreover, while students usually sit exit exams at the end of the lower and upper secondary stages of education, i.e. at the ages of 15 and 18, relatively little is known about the formation of motivation to learn at the age of 10, i.e. at the time when pupils transfer from primary to lower secondary education. This age is characterized by notable developmental changes and culminating gender differences in brain development (Lenroot et al., 2007).

In this study, we enrich the existing literature by separating the impact of teachers' incentives from the effect of the students' motivation to learn. In particular, we explore whether and to what extent high-stakes admission exams to selective schools affect student effort and consequently achievement. Our identification strategy exploits the quasi-experimental design of the 2009 school reform in Slovakia, which postponed the transition to selective schools by one year, from the end of the 4th to the end of the 5th grade. To identify the effect of high-stakes admission exams, we use student achievement in the 4th grade from the international survey TIMSS before and after the Slovak reform. In 2007, i.e. before the Slovak reform, the 4th graders were tested by TIMSS in May, and at the beginning of June the same students then had to sit admission exams for selective schools. In 2011, i.e. after the Slovak reform, TIMSS examined the 4th graders in May, as before, but the students were expected to sit selective school admission exams a whole year later. Our hypothesis is that this shift in the timing of competitive school selection lowered the 4th graders' motivation to learn, leading to lower achievement in TIMSS tests in 2011. Since test scores

[2] See Bishop (1997) and Bishop (1999) for the detailed description of incentives of students, parents, teachers, and school administrators induced by high-stakes testing.

[3] Jürges et al. (2012) show a positive effect of central exit exams only on exam-specific knowledge together with no effect on knowledge not included in tested fields. A similar result is shown by Jacob (2005). When the test results are relevant also for teachers and school administrators, it seems that teachers teach predominately the curriculum that is tested.

between 2007 and 2011 could exhibit a longer-term trend for reasons unrelated to the reform, we include data for students in the Czech Republic in our study: Czechs were also tested by TIMSS in both 2007 and 2011; however, in both years they took admission exams to selective schools at the end of the 5th grade, i.e. according to the timescale adopted in Slovakia after the reform. We use the difference-in-differences methodology to identify the causal effect of admission exams on achievement. Because of the lack of data we cannot investigate a common pre-treatment trend in TIMSS test scores[4] in these countries, as it is a usual identification assumption in the difference-in-difference approach. Therefore we further exploit the variance in excess demand for selective schools across Czech and Slovak districts, and apply difference-in-difference-in-difference approach.

Unlike school exit exam results, the results of high-stakes admission exams to selective schools are not publicly released, and are not perceived as an indicator of school quality, therefore they do not drive teacher motivation. Moreover, in the Czech and Slovak case studied, there is anecdotal evidence that the primary schools consider the earlier transition to selective schools to be a threat to them, as they risk losing their best students.[5] As a result, schools and teachers do not usually provide additional tutoring or special teaching to prepare students for the admission exams (Straková and Greger, 2013; Federičová and Münich, 2014). Any effect the high-stakes admission exams have on student learning can thus be attributed to increased effort on the part of the students themselves, and their parents.

High-stakes admission exams at early ages are part of the admission process to academic junior high schools that exist in most central European countries, such as Germany, Austria, the Czech Republic, Slovakia, and Hungary. Selection for these schools takes place at the end of primary education. There is anecdotal evidence that some parents and students take these exams very seriously. Using data from a unique longitudinal study[6] in the Czech Republic, Federičová

[4] A common pre-treatment trend is however indicated in PISA test scores that examine 15-year-old students in math in three-year cycles. See Figure 2.6.
[5] Currently, schooling ministries in both countries consider the reduction of enrolment in academic junior high schools, since the driving reason is perceived as an adverse impact of brain-drain on mainstream schools.
[6] The Czech Longitudinal Study in Education (CLoSE) follows the cohort of students in the Czech Republic who were tested in their 4th grade by TIMSS and PIRLS 2011. The data include detailed information about the students' intention

and Münich (2014) examine the scale and scope of preparation for these exams. They document that among 20% of those who apply to academic junior high schools, nearly 85% devote some time to their preparation every week during the last semester before the admission exam, and among them, more than half prepare for the exams almost every day. At the same time, only a small group of students (15%) have the opportunity to attend a preparatory course at their school. This confirms the schools' low interest in supporting their students' transition to the selective schools. Moreover, Federičová and Münich (2014) also show that the vast majority (90%) of students apply to selective schools on the basis of their own or their parents' interest, and only half of their teachers agree with their decision.

Our estimates imply that high-stakes exams increase ten-year-old students' math test scores by 0.2 standard deviations. Using the quantile regression, this effect additionally accrues by 0.05 standard deviations for students in the top decile of the math test score distribution, i.e. students with the highest probability of being accepted to selective schools. Furthermore, we identify some differences in behavior between girls and boys. Although the overall effect of the admission exams is about the same on both girls and boys, girls seem to put greater effort into succeeding at the admission exams than boys in districts that have a more competitive selection process.

Our paper is structured as follows: section 2 reviews the literature on the role of motivation, and the existing empirical findings on the effect of high-stakes testing on student achievement. Section 3 describes the selection process for academic junior high schools in the Czech Republic and Slovakia. Sections 4 and 5 discuss the data and the identification strategy. Section 6 presents the results, and section 7 concludes.

2.2 Literature review

In the field of educational psychology, motivation plays an important role in exerting efforts and enhancing educational outcomes (Crumpton and Gregory, 2011; Hidi and Harackiewicz, 2000); in the long-run, it can also affect individuals' labor market outcomes (Dunifon and Duncan, 1998). This is reflected in the commonly accepted definition of motivation that defines it as a force that

to apply to an academic junior high school, their preparation for the admission exams and their results in 2012. For a further description of the CLoSE survey see Straková and Greger (2013).

moves one to act (Ryan and Deci, 2000). Classic motivational theory distinguishes two types of motivation, intrinsic and extrinsic. Intrinsic motivation is linked to individual interests and has a positive impact on student achievement. Extrinsic motivation accrues from external factors such as rewards, deadlines, competition or performance evaluations, and can either foster or undermine the impact of intrinsic motivations (Crumpton and Gregory, 2011).

Keeping the dichotomy of intrinsic and extrinsic motivation in mind, Ryan and Deci (2000) formulated a new self-determination theory (SDT). SDT defines several types of extrinsic motivation which differ in their level of self-determination. In other words, the distinguishing feature of each type is the extent to which the individual perceives external factors as important for their own sake. Some external factors motivate individuals only to obtain rewards or to avoid punishments. These incentives disappear with the disappearance of the source of motivation and hence do not motivate individuals to acquire new skills or to achieve personal goals. On the other hand, some external factors are personalized and individual. Although this extrinsic motivation is not driven by any inherent interest—as in the case of intrinsic motivation—it can enhance intrinsic motivation when students perceive the external factors as being important for their present or future outcomes.

High-stakes exams, such as those related to school admissions or school grades, represent specific external motivational factors. Although they are not usually intended to foster motivation *per se*, they do so because they are widely perceived as prestigious and important for future school outcomes.[7] Therefore, they can facilitate intrinsic motivation, leading to greater achievement. Several empirical studies have found that high-stakes exams of this kind have a positive effect on student achievement. Jacob (2005) shows that a new policy in the Chicago schooling district, which introduced high-stakes testing in several grades as a requirement for transition to higher grades increased students' average math and reading test scores by 0.2 to 0.3 standard deviations. Meanwhile, a cross country comparison of nation-wide standardized exit exams studied by Woessmann (2002) shows that these yield even higher effects on math literacy, with a magnitude around 0.4 standard deviations. Using the variation in the schooling systems across German states,

[7] In particular, more than 90% of students in the CLoSE survey perceive admission to selective schools to be a gateway to university education.

Jürges et al. (2012) find that students in states with standardized exit exams outperform students in states without such exams by 0.26 standard deviations in the subjects tested, although they find no effect on literacy in areas not tested as part of the central exit exams. They interpret their findings as the accountability feature of high-stakes exams fostering the monitoring of schools and teachers either by parents or school administrators. Contrary to this, high-stakes admission exams to selective schools primarily motivate the students themselves and their parents, rather than schools and teachers. This is because schools have no interest in losing their best students.

Intergenerational transmission literature[8] provides evidence of the important role that parents play in the interaction of extrinsic motivations. In our case, they are influential in pushing their children to apply for selective schools and to prepare for the admission tests. Moreover, there are good reasons to expect a different impact on students of different genders. At the time of early school selection—usually at the age of ten or eleven—boys are less mature and less responsive to parental authority than girls (Wilder and Powell, 1989). Although the effect of parental background seems to be stronger in the case of boys at this stage, girls are more aware of the importance of the admission exams and put greater effort into preparing for the admission tests. Hence, one might expect that selection procedures at younger ages would increase girls' achievement more than boys'. However, we do not observe corresponding gender imbalances in the composition of selective schools' student population. This may be explained by the stressful nature of the high-stakes admission exams and a distinct perception of stress by each gender. As documented by both Gneezy and Rustichini (2004) and Jurajda and Münich (2011), high-stakes admission tests may cause psychological stress that works against girls' success. As a consequence, the girls' greater sense of purpose does not necessarily translate into their greater success in admissions.

2.3 School selection in the Czech Republic and Slovakia

Our identification strategy relies on a policy intervention in Slovakia. In addition to Slovak data, we use data from the Czech Republic to control for common trends. Slovakia and the Czech

---

[8] On the impact of family background on children's academic outcomes, see Cunha and Heckman (2007); Feinstein (2003); and Anger and Heineck (2010). On the impact of parents' attitudes on the development of children´s attitudes and on the formation of their goals and motives, see Grolnick and Ryan (1989); Friedel, Cortina, Turner and Midgley (2007).

Republic are both successors of the former Czechoslovak Federation, dissolved in 1993. Although both countries have developed separately since then, they have remained extraordinary similar in terms of their demographic structures and institutions, including their school system (Table 2.1).

Table 2.1: Characteristics of the school system in the Czech Republic and Slovakia

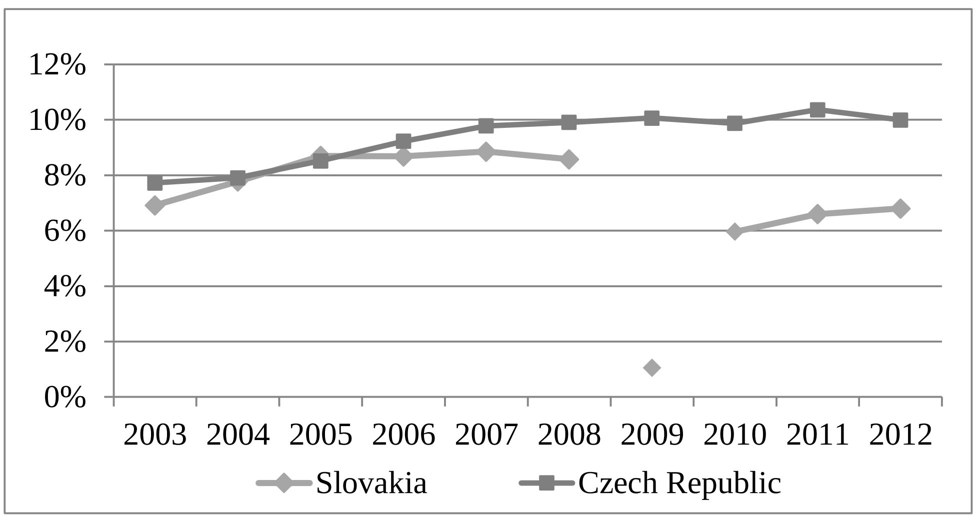
Characteristics of school system	Czech Republic		Slovakia	
	2007	2011	2007	2011
School entry age[1]	6	6	6	6
-number of years of compulsory education	9	9	10	10
-number of grades in primary and lower secondary education[2]	9	9	9	9
The grade of the first academic selection	6	6	5[3]	6
The average number of pupils in class				
-in primary education	19	19	19	17
-in lower secondary education	20	19	22	19
The teacher-pupil ratio				
-in primary education	1:17	1:17	1:16	1:14
-in lower secondary education	1:12	1:11	1:12	1:11
Owner structure of primary and lower secondary education				
-% of students in private schools	0,6%	0,8%	0,6%	0,9%
-% of students in church schools	0,6%	0,7%	4,8%	5,1%
Government expenditure on edu. (% of GDP)[4]	4,5%	4,7%	4,4%	4,6%
Teacher salaries in primary and lower secondary schools (in EUR)	877	1 091	545	802

Source: National statistical offices.

[1] Pupils can enrol into Basic School if they reach six years of age by September 1 of a given school year. Children can enrol at the age of 7 if recommended by pedagogy advisor or upon parental request.
[2] It was 8 years for both countries before 1998.
[3] The 2009 school reform in Slovakia that shifts forward the transition to selective schools by one year.
[4] Government expenditure data from OECD (2013). Years of reference for data are 2005 and 2010.

Both systems allow for the early tracking of students into selective schools. In particular, at the end of the 5th grade, a small proportion of students are given the opportunity to move from the so-called *Basic School* (lasting 8 - 9 grades) into a so-called *Academic School*.[9] Studies at Academic School last eight years and cover the lower- and upper-secondary levels. In both countries, school types similar to these Academic Schools had been traditional before the Second World War, but were then closed by the communist regime in the early 1950s. After the Velvet Revolution in 1989, Academic Schools were re-established in both countries. The only difference between the two countries was the grade at which students entered the Academic Schools. In the Czech Republic, Academic Schools admitted students after completion of their 5th grade of Basic School, whereas in Slovakia students entered Academic Schools after their 4th grade. The 2009 school reform in Slovakia pushed the enrolment grade forward by one year, bringing it in line with the Czech system. As part of this reform, the Slovaks also reduced the proportion of students enrolled in the Academic Schools from almost 9% (in 2009) to 6% (Figure 2.1).

Figure 2.1: The enrolment in the first grade of Academic Schools (lower secondary) as a share of an age cohort



Source: National statistical offices.
Note: The sudden drop in the time series of Slovakia was caused by the 2009 school reform.

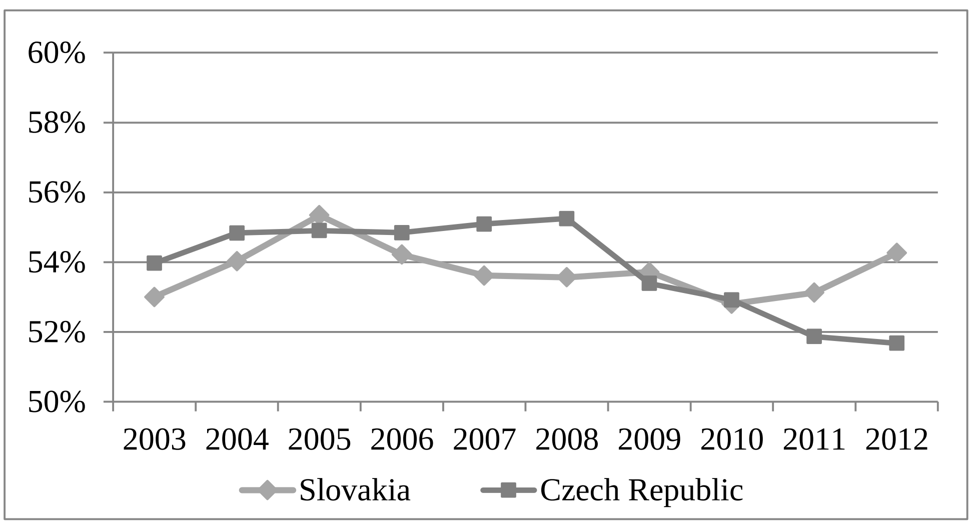
Although the curriculum taught at the Academic Schools is the same as that taught in the corresponding grades at mainstream Basic Schools, it is taught in greater depth and intensity at the Academic Schools, where higher demands are made of the students. Compared to Basic Schools,

[9] In Slovak / Czech, Basic School is called *základná škola / základní škola*, Academic School is called *osemročné gymnázium / osmileté gymnasium*.

Academic Schools are traditionally perceived as being more prestigious, of better quality, securing peers providing greater positive spillovers, and as a result, attending one of these schools considerably increases a student's chance of being admitted to a high quality university (Dearden, Ferri and Meghir, 2002).[10] Since the number of places at Academic School is kept fixed, Academic Schools have faced substantial excess demand ever since their re-establishment in the 1990s. This excess demand generates extrinsic motivation for some students to study harder, in order to pass the demanding admission exams.

Since early 1990s, in both countries, the share of students in Academic Schools has grown graduallyFigur). This positive trend during over the past two decades was autonomous due to a demographical decline vis-à-vis the fixed number of slots in Academic Schools. The only intentional change in selectivity rate was brought about by the 2009 school reform in Slovakia. In terms of gender composition, girls only slightly outweigh boys in Academic School student bodies. Over the past decade, around 55% of Academic School students have been girls (Figure 2.2).

Figure 2.2: The proportion of girls in the first grade of Academic Schools (lower secondary)



Source: National statistical offices.

[10] Using a natural experiment in Northern Ireland, Guyon, Maurin and McNally (2012) show demonstrate the effect of Academic schools on the success in national examinations at age 16 and 18.

In the school year 2011/2012, there were 147 Academic Schools in Slovakia and 275 in the Czech Republic. Admission exams for Academic Schools take place at the beginning of June in Slovakia, and at the end of April in the Czech Republic. Prior to the 2009 reform in Slovakia, TIMSS testing coincided with these admission exams, while since the reform, TIMSS testing has preceded Academic School admission tests by about one year, in both countries. In both countries, each Academic School uses its own admission criteria which they judge, in most cases, on the basis of written tests. Mathematics and the national language are the most commonly tested subjects. The tests are usually based on curricular knowledge, as is also the case with TIMSS testing. Some Academic Schools also use general study aptitude tests or adopt additional criteria such as the student's grade performance at Basic School or achievement in mathematical and other Olympiads.

2.4 Data

TIMSS is an international[11] survey which tests 4th graders' mathematics and science skills in four-year cycles.[12] We employ two successive rounds of TIMSS data: the 2007 round, i.e. preceding the school reform in Slovakia, and the 2011 round, i.e. following the reform. We also use data from the Czech Republic, which serves as a control group.

Figure 2.3 (in Appendix) depicts the kernel densities of TIMSS math test scores[13] in both countries in both years, and Table 2.2 reviews the corresponding basic distributional characteristics. In 2007 the Czech Republic lagged behind Slovakia in terms of average scores, but its average test scores grew five times faster than Slovakia's, until 2011, when average scores in both countries were very similar. More importantly, whereas this shift in the Czech Republic is

[11] TIMSS assessment data are comparable across all participating countries, since all students who take part in TIMSS are tested on the same set of test items. Moreover, various procedures during the preparation of test items ensure compatibility of the test results across all countries (Mullis, Martin and Foy, 2008).

[12] The assessment data from the first cycle of TIMSS in 1995 were scaled by setting the mean test scores from all participating countries to 500 and the standard deviation to 100 (Olson, Martin and Mullis, 2008). In each successive TIMSS cycle, the assessment data were placed on a scale from the previous cycle to provide accurate measures of trends across all cycles of TIMSS.

[13] TIMSS provides five plausible values that represent student math test scores (Foy, Galia and Li, 2008). To compute descriptive statistics of math test scores and regression coefficients and their standard errors, we apply Rubin's (1987) methodology for multiple imputations.

similar across all percentiles of the TIMSS math test score distribution, in Slovakia the shift does not appear in the top quartile.

Table 2.2: Math test scores in Slovakia and the Czech Republic in 2007 and 2011

Percentiles	Slovakia			Czech Republic		
	2007	2011	Δ 2011-07	2007	2011	Δ 2011-07
5^{th}	350 (10.2)	363 (8.5)	13 (14.3)	361 (7.1)	387 (5.8)	25 (9.5)*
10^{th}	389 (9.8)	401 (7.2)	12 (13.4)	392 (7.0)	419 (4.9)	27 (8.7)*
25^{th}	446 (4.2)	460 (3.7)	15 (6.0)*	440 (4.5)	467 (2.8)	27 (5.3)*
50^{th}	502 (3.0)	513 (3.7)	11 (5.1)*	490 (4.0)	514 (2.4)	24 (4.3)*
75^{th}	553 (3.5)	560 (2.9)	7 (5.0)	536 (3.0)	560 (2.3)	24 (3.6)*
90^{th}	597 (4.6)	600 (4.3)	3 (6.6)	576 (2.9)	598 (2.3)	22 (2.9)*
95^{th}	623 (5.2)	626 (6.1)	2 (8.9)	597 (3.0)	621 (3.7)	24 (4.6)*
overall	496 (4.5)	506 (3.8)	10 (6.4)	486 (2.8)	511 (2.4)	25 (3.3)*
N	4,919	5,150		4,199	4,580	

Source: Own calculations based on TIMSS 2011 and TIMSS 2007.
Note: Standard errors are in parentheses (*p<0.1).

This pattern is in line with our key hypothesis that additional studying effort due to preparation for Academic School admission exams (present in Slovakia in 2007 but not in 2011) increases educational achievement among those who apply for admission.

To account for the possibility that the strength of extrinsic motivations could be dependent on the students' demographic and socio-economic family backgrounds, we control for individual characteristics. In particular, following common practice, we use the number of books at the student's home as a proxy for their socio-economic family background.[14] TIMSS reports the number of books on a discrete scale. The distributions of book possession are similar in both countries and exhibit similar, relatively small changes between the two years. Table 2.3 in Appendix reviews the descriptive statistics of TIMSS individual background characteristics in the two testing years for both countries.

[14] See e.g. Brunello and Checchi (2007) which shows that number of books at home and level of parents' education both have similar effects on student achievement.

2.4.1 Excess demand

We explore whether the presence of high-stakes admission exams gives rise to extrinsic motivations fostering students' study effort and increasing their educational achievement as measured here by TIMSS math test scores. If selective admission to Academic Schools creates additional study incentives, greater incentives and therefore a greater impact on student achievement should be observed in districts experiencing higher relative excess demand (measured as the ratio of applications to admissions minus 1) and negligible incentives and impact in districts with zero excess demand. To test this, we collected district level data from school registers and merged them with the TIMSS data.[15] Figures 2.4 and 2.5 in Appendix depict the levels of and changes in demand, supply and relative excess demand for Academic Schools between the years 2007 and 2011, at the district level. The systematic drop in supply in almost all Slovak districts in 2011 was due to the 2009 school reform. Some Academic Schools reduced overall student numbers or class sizes after the reform, resulting in an overall drop in supply by 28%. However, demand for Academic Schools in Slovakia dropped too (by 24%), so that the relative excess demand remained about the same. It should be noted that the demand reported is the revealed one by applications submitted. In Slovakia, the drop in enrolments led to lower number of applications submitted. By contrast, while supply remained relatively stable between the two TIMSS rounds in the Czech Republic, demand increased by 20%. This led to higher relative excess demand (up to around 30%) driven primarily by increased demand in districts which had already faced high demand and high relative excess demand in 2007.

2.5 Methodology

Our identification strategy employs the quasi-experimental feature of the 2009 school reform in Slovakia, i.e. an exogenous shift in the timing of admission exams for selective Academic Schools, by one year. Our key hypothesis is that the reform reduced the incentives for some students in the 4th grade to study, leading to lower levels of achievement in the TIMSS tests in 2011 in Slovakia.

---

[15] As the excess demand incorporates application behavior, and hence may be endogenous to student achievement, we use the excess demand from years 2008 and 2011. As in Slovakia it is not mandatory to report the number of applications, we further add the missing data by contacting Academic Schools personally. Because some of these schools are closed now or did not archive this data, for 4% of observations excess demand remains missing.

To control for possible trends in test scores between 2007 and 2011 that could be unrelated to the 2009 reform, we use data on the equivalent cohort of students in the Czech Republic as a control group. These data enable us to apply a difference-in-differences methodology (DID) controlling for time- and district-specific effects (Wooldridge, 2002).

To examine the impact of admission exams on achievement via increased student studying effort, we estimate the following model:

$$S_{it}^{d} = \alpha + \beta_1 DD^{SVK} * DY^{2007} + \gamma_1 DY^{2007} + \gamma_2 DD^{d} + \gamma_3 X_{it} + \epsilon_{it}^{d} \tag{1}$$

where S is the math score[16] achieved by finishing 4th grade student i in year t and in district d; DD^{SVK} is a dummy equal to 1 for Slovak districts and 0 for districts in the Czech Republic; and DY^{2007} is a time dummy equal to 1 for the year 2007, i.e. before the 2009 reform, and 0 for the year 2011. Thus, 2011 and the Czech Republic represent the base captured by the intercept α. Furthermore, we control for time (DY^{2007}) and district (DD^{d}) specific effects. We also control for observed individual and school characteristics in vector X, such as the students' age and number of books at home, and the size of the municipality in which the school is located. The key coefficient of interest is β_1 which captures the effect of the high-stakes admission exams (treatment) on the Slovak student achievement in 2007, i.e. before the reform. To allow for different impacts on boys and girls, we estimate Eq. (1) separately for each gender.

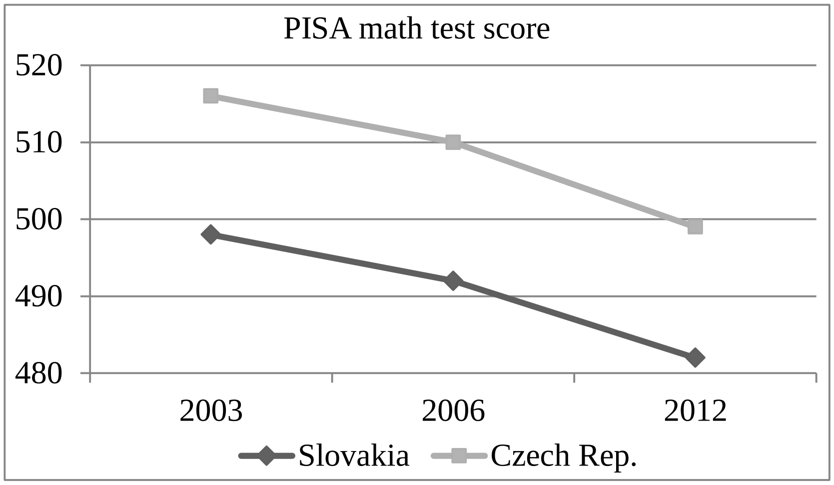
Our key identifying assumption is that in the absence of treatment, the coefficient β_1 would be equal to zero. Specifically, this means that on average and conditional on our covariates, the TIMSS test scores in the Czech Republic and Slovakia would have followed the same trends if the Slovak reform had not been introduced. As we explained in Section 3, following the division of Czechoslovakia the two countries' school systems remain very similar. We document this with test score trends from the PISA international survey,[17] which tests 15-year-old students (Figure 2.6). Unfortunately, there are no other surveys available that would enable us to provide additional

[16] The mean of the math test score distribution is normalized to 0 and the standard deviation to 1.

[17] We should show here that TIMSS test scores in the two countries have evolved with parallel trends after the 2009 reform in Slovakia. However, the TIMSS test scores for years later than 2011 are not yet available.

evidence.[18] But except the shift in the time of selection, and the drop in enrolment to Academic Schools, there were no other notable changes in Slovakia which would have had an impact on student achievement, especially on students at the upper half of the test score distribution.

Figure 2.6: Time trends in math test scores (PISA) in the Czech Republic and Slovakia



Source: PISA 2003, 2006, and 2012.

As the main identification assumption of the difference-in-differences approach cannot be explicitly exposed in TIMSS data, we further exploit the variance in excess demand to selective schools and employ difference-in-difference-in-difference strategy. Together with the presence of high-stakes admission exams, the degree of excess demand may foster incentives and increase student performance in order to secure admission. This establishes variation in competitiveness across districts and in time that may capture the possible unobserved factors that influence the time trend in test scores in the Czech Republic or Slovakia. Moreover, given that the excess demand is correlated with the interaction term, the estimate of the interaction coefficient in the previous specification of Eq. (1) may be biased due to the omitted variable. To relax the common trend assumption and to obtain unbiased estimates, we consider an augmented specification of Eq. (1) controlling for relative excess demand:

---

[18] The majority of students affected by the 2009 school reform in Slovakia reached 15 years of age in 2014. Using the PISA survey data from before and after the reform, i.e. PISA 2012 and future PISA 2015, we will be able to examine the effect of later selection, estimating whether one more year with the top students enhanced or worsened the level of achievement among those students who remain in Basic School.

$$S_{it}^{d} = \alpha + \beta_1 DD^{SVK} * DY^{2007} + \gamma_1 DY^{2007} + \gamma_2 DD^{d} +$$

$$+ \beta_2\, ED(DD^{SVK} * DY^{2007} + DY^{2007} + DD^{d}) + \gamma_3 X_{it} + \eta_{it}^{d}, \quad (2)$$

where *ED* is relative excess demand[19] for Academic Schools in district *d* and time *t*. Note that the degree of selectiveness should only matter in Slovakia before the reform, and hence, the coefficient γ_4 should be close to zero. The coefficient for the interaction term between treatment and excess demand (β_2) captures the change in achievement due to more intensive studying prompted by the more competitive environment.

Further, it should be noted that only 20% of students in the Czech Republic and 10% of students in the Slovak Republic apply for a place at an Academic School. One might suppose that the high-stakes admission exams these students sit have a greater impact on their own effort and achievement than they have on other students via, for example, peer-effects. Unfortunately, the TIMSS data do not contain information as to which individual students applied for Academic Schools and which did not. We use two alternative approaches, taking into account that the probability of applying for a place at Academic School is higher at the upper tail of the initial skills distributions.

First, we estimate a model (2) using a quantile regression.[20] The quantile regression model estimates the treatment effect at different parts of the conditional distribution of TIMSS test scores. Thus, in contrast to the linear regression model which estimates the conditional mean function, the quantile regression model estimates the conditional quantile function (Hao and Naiman, 2007).

Second, we estimate the percentile distance of each student from the admission threshold and enter it as an additional covariate to models (1) and (2). We assess the distance by estimating the probability of a student being admitted to Academic School using data from an *ad hoc* follow-up survey of the Czech TIMSS 2011 cohort at the end of their 5th grade (one year after TIMSS testing)

[19] Excess demand is measured as the ratio of applications to admissions minus 1, i.e. zero excess demand indicates districts in which the same number of students applies for places as are admitted.
[20] For a detailed description of quantile regression see Koenker (2005).

which contains information as to whether the student was admitted to an Academic School[21]. We use estimated parameters of a Probit model admitted / not admitted to predict the corresponding admission probabilities for Slovak students (Table 2.4 in Appendix). By ranking the students by their predicted probabilities within each district and taking into account the district-specific number of places at Academic School, we compute the absolute percentile distance of each student from the admission threshold in both directions[22]. We further use the distance variable to estimate the specification from model (2) on the subset of students who are close to the admission margin from both sides, i.e. whose distance variable is less than 5 percentage points.

2.6 Results

Table 2.5 presents the estimated coefficients and their standard errors for different specifications of the difference-in-differences models (1) and (2) for girls and boys separately. The *treatment* variable is the dummy for Slovakia in 2007. To interpret the estimated coefficients more easily, we standardized the test scores so that in 2007 they have zero mean and a standard deviation equal to one.[23] In the base-line model presented in Column (1), the average treatment effect is significant and equal to 0.11 and 0.14 standard deviations for girls and boys, respectively. Controlling for individual and school characteristics in Column (2), the treatment increases to about 0.19 standard deviations with no gender difference. Hence, the presence of admission exams to Academic Schools enhances students' average math test scores by around 0.2 standard deviations.[24]

[21] The longitudinal Czech project CLoSE ran a follow-up survey for the TIMSS 2011 student cohort one year later, and asked detailed questions concerning their preparation for Academic School admission exams and their results. The CLoSE data enable us to estimate a Probit model of the probability of being admitted to Academic School for the Czech sample of TIMSS 2011 students.

[22] This threshold is specified for each district by the proportion of students in the first year of Academic School out of the total number of students in the respective cohort.

[23] Standardized score S is computed from original TIMSS scores T as $S = (T - T_{2007})/\sigma_{2007}$, where T_{2007} is mean score and σ_{2007} is standard error in 2007 (both countries).

[24] The ordinal nature of TIMSS test scores allows that means that a one point difference in the test score at the lower tail of the skills distribution does is not necessarily measure equivalent to the same increase in skills in the upper tail. This problem was highlighted by in a recent study of by Bond and Lang (2013) showing the sensitiveness of estimated test score gaps on the choice of a scale. By applying various order-preserving scale transformations of to the TIMSS test scores, similar to those used by Bond and Lang (2013), we find that the estimated parameters from obtained from the transformed scores are not effectively significantly different from those of the baseline model.

We also control for the intensity of competition in admission by relative excess demand *ED* in Column (3). The significant positive coefficient on the interaction term for excess demand and treatment in the girls' case is indicative of the admission tests having a stronger impact on the girls' attainment in districts where competition for places at Academic School is higher. However, this higher competition is not found to have any impact on the effort boys put in to their studies. Regardless of the excess demand, the average treatment effect on boys is equal to 0.2 standard deviations. For girls, the average treatment effect rises to 0.2 standard deviations in districts with relative excess demand equal to 1, i.e. with two applicants per place. It appears that in districts with more competitive admissions, girls devote greater effort to preparing for the tests, resulting in greater achievement. The positive significant coefficient of excess demand refers to the effect of the change in excess demand between 2007 and 2011 on the math test scores. Districts in which excess demand increased by 1 report a rise in the mean math test score by 0.09 standard deviations for girls, however with no change for boys.

In the last column (Column 4) in Table 2.5, we exclude 4 districts which have very high excess demand, including both capital cities. While the average treatment effect does not change, the effect of excess demand in treatment rises, for the girls, to 0.28 standard deviations.

Quantile regression estimates for the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} quantiles are reported in Table 2.6. The estimated impact of the treatment at the 90^{th} quantile (i.e. the upper end of the test score distribution) is higher than the impact of the treatment in lower quantiles and the base-line model estimates in Table 2.5. For boys, as well as for boys, the treatment effect increases steadily, from zero in the lowest decile to significant and positive effects in higher quartiles. The treatment boosts boys' test scores on the 25^{th} quartile by 0.13 standard deviations, rising to 0.28 standard deviations for students on the 75^{th} quartile. A lower treatment effect is observed in the top decile, equal to 0.22 standard deviations. As in the base-line model in Table 2.5, boys are not affected by the more competitive admission process measured in regressions by the excess demand. On the other hand, the average treatment effect is significant only for girls at median test score and in the top decile, and it is equal to 0.13 and 0.18 standard deviations.

Furthermore for girls, the treatment effect notably increases with the intensity of the selection process, across the math test score distribution. Whereas the more competitive admission process have no effect on the lowest decile, the effect becomes significant and further rises from 0.17 standard deviations in the lower quartile to 0.25 standard deviations in the top decile. Thus, in the highly competitive districts, with excess demand equal to 1, the presence of the admission exams raises the girls test score by 0.29 and 0.43 standard deviations, at median and in the top decile, respectively. Hence, in the districts with zero excess demand, the presence of admission exams affects boys' effort in the top decile more than girls' effort (by 0.04 standard deviations) whereas in the highly competitive districts, the treatment effect on achievement in the top decile is higher for girls than boys, by around 0.2 standard deviations.

Table 2.7 presents estimates of models (1) and (2) using the students' estimated absolute distances from district-specific admission thresholds. The estimated treatment effect is equal to 0.21 standard deviations at the admission threshold for girls and 0.19 for boys, and declines significantly with increasing distance from the threshold.

The augmented specification is presented in the 2nd column. Controlling for the excess demand, the average treatment effect becomes insignificant for both girls and boys. This means that there is zero treatment effect at the admission threshold in districts with no selectivity, i.e. with zero excess demand. Also, the effect of distance from the admission margin does not differ with or without the treatment. However, we observe very high positive estimates for excess demand in interaction with treatment for both genders. Here, a unit change in relative excess demand significantly raises the treatment effect by 0.52 standard deviations for girls at the admission threshold and 0.33 for boys. This effect declines further with increasing distance from the admission threshold. A percentile change in distance decreases the treatment effect on girls by 0.005 standard deviations. In other words, the relatively low, or for boys zero, effects more competitive admission procedures have on the average student’s educational achievement are higher or students at the admission margin by around 0.3 standard deviations, that is for students whose future schooling is most decided by the admissions exams.

Focusing on the subset of students who are within 5 percentage points above and below the admission margin, and estimating the model (2), we obtain similar results to those found in the previous specifications. The results are presented in Column (3) of Table 2.7. These students exhibit zero average treatment effects in districts with zero excess demand, but treatment effect increases with excess demand and reaches 0.6 and 0.4 standard deviations in districts with unit excess demand, for girls and boys respectively.

2.7 Summary and conclusion

We address important policy questions concerning the impact of students' motivation on their achievement. Educational psychology studies commonly consider the role of external factors such as rewards, deadlines, competition or evaluation in influencing students' educational progress. A plethora of studies in this area have shown that such incentives can increase students' academic performance, but that they can also undermine it. Our analysis focuses on the particular, very common yet rarely investigated, study incentives created by high-stakes admission exams for selective schools, and their consequent impacts on student achievement. High-stakes admission exams at early ages differ from other high-stakes testing such as exit exams, in that they do not affect in-school teaching intensity. If students and their parents perceive admission exams to selective schools to contribute positively to their future study and labor market outcomes, then these exams create additional study incentives based on personal interest.

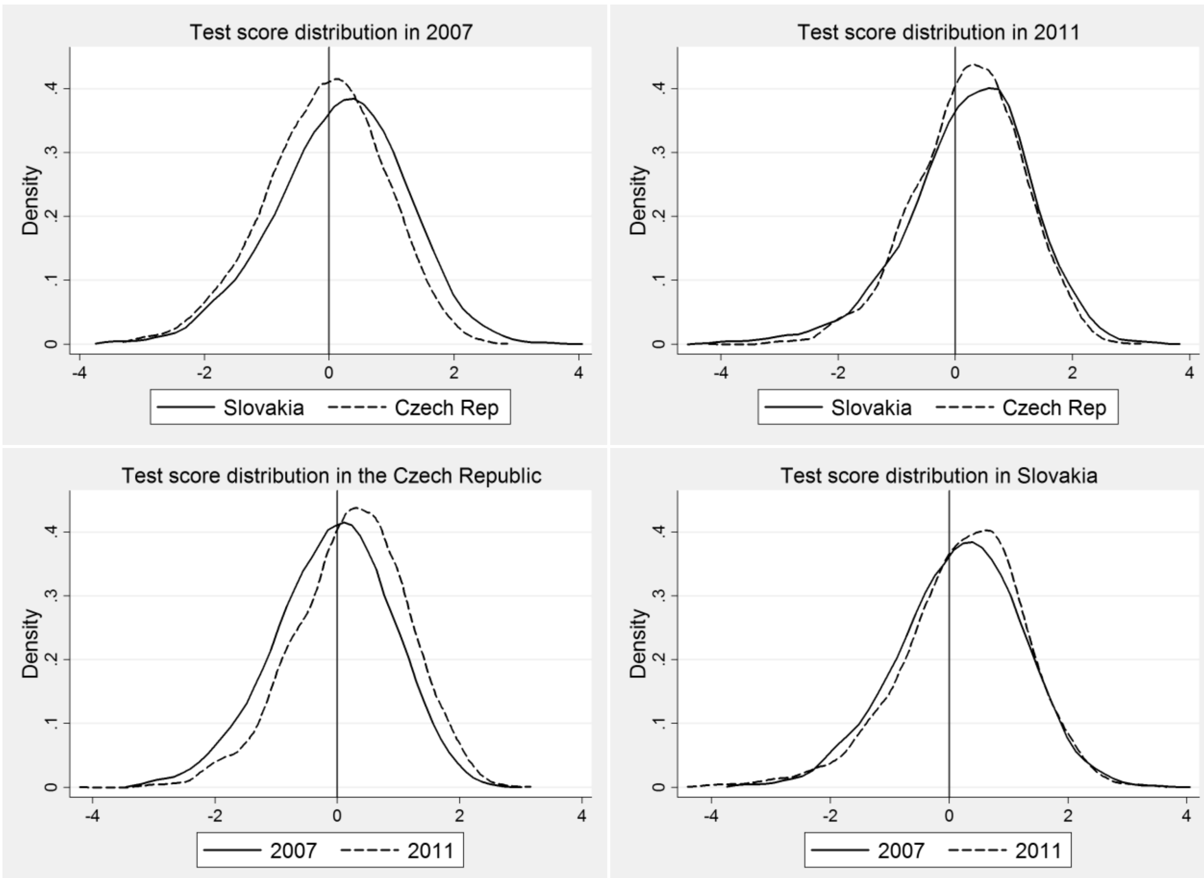
Our empirical findings indicate that high-stakes admission exams to Academic Schools at the end of primary education motivate some 10 to 11years old students to intensify their effort to learn, leading to greater achievement in math. Our identification strategy, which made use of an exogenous quasi-experimental policy intervention, estimates treatment in the range of 20% of the standard deviation of test scores. The size of the treatment effect is similar to the findings in the previous empirical studies, which had reported the effects of increased student effort and teacher effort combined. The effect is of the same nature for both genders, although girls seem to be more sensitive to the degree of competition and peer pressure involved. Although girls on average lag behind boys in mathematics, they are not deterred from admission exams and are willing to put greater effort into succeeding, especially when they face pressure from a more competitive selection process. Focusing on students with the highest probability to apply, i.e. students in the top decile, we find that in districts with the most competitive admission process, i.e. where there are usually two applicants per place available, girls increase their achievement by 0.4 standard deviations when they are about to sit the admission exams. In the case of boys, this is equal to 0.22 standard deviations in all districts regardless of the competitive admission process. Possible gender-related differences in motivation and their effect on academic performance may contribute to gender inequality in educational outcomes, such as participation in academic education or test achievement. However, further research is necessary in order to better understand the origins and consequences of such gender differences in motivation to learn.

These findings about the effects of admission exams to selective schools might be generalized to high-stakes tests of all kinds. In both cases, an important issue would be further to examine the long-term effect of these exams on the student achievement. Regarding the early selection studied in this paper, the important issue is also the response of students not at the skill level for entrance to selective schools on this selection process. The data available for this paper, however, does not allow us to address these questions. So, we rather leave these issues for further analysis.

Our findings should nevertheless be used cautiously in support of early school selection. It is because early selection brings with it also many other effects that are much less desirable from a normative point of view. Selective schooling systems strengthen the role of factors such as family wealth, personal pressure from parents and perseverance, ambitions, and rigor in the students' approach to studying. The role of these factors in access to education should have been diminished by the primary schooling system. Our findings relating to the Slovak reform are relevant for systems in other countries with early selection, including Austria, Germany, and Hungary. They also relate to the current debate in the UK about Grammar Schools accepting students who are well trained to pass their demanding admission exams, but lack sufficient study skills to make longer-term academic progress afterwards.

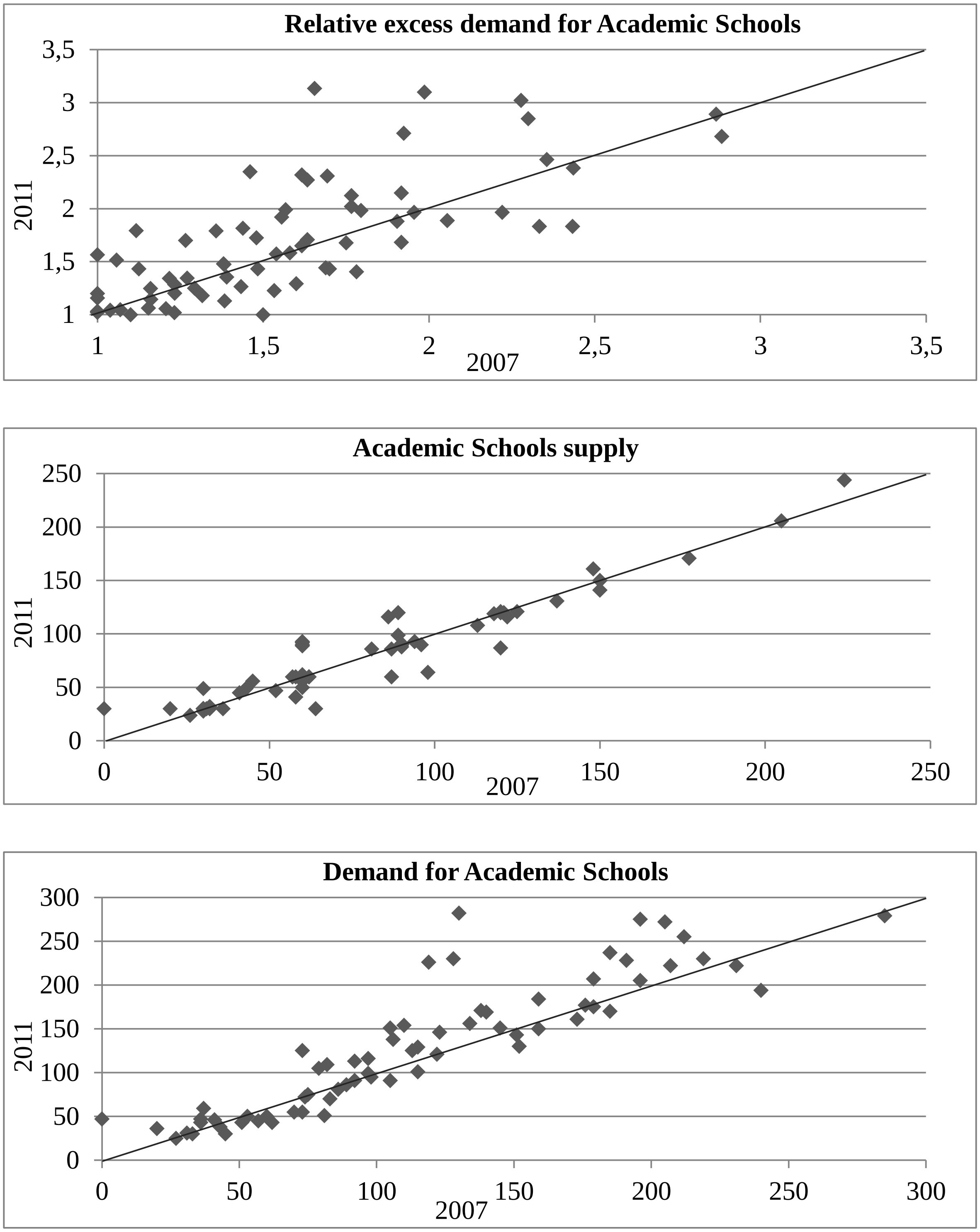
2.A Appendix 2

Figure 2.3: Math test score distributions in years 2007 and 2011



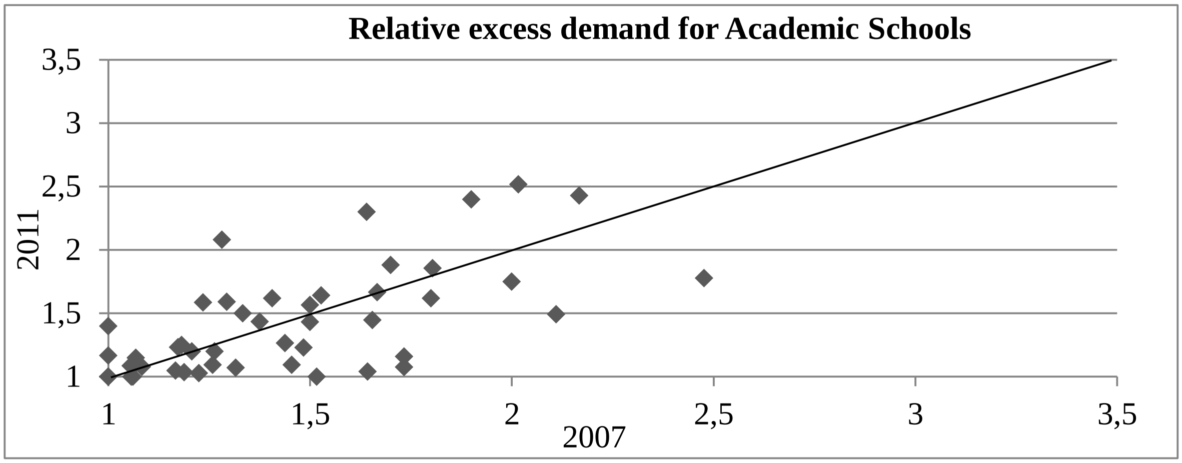
Source: TIMSS 2007 and 2011.

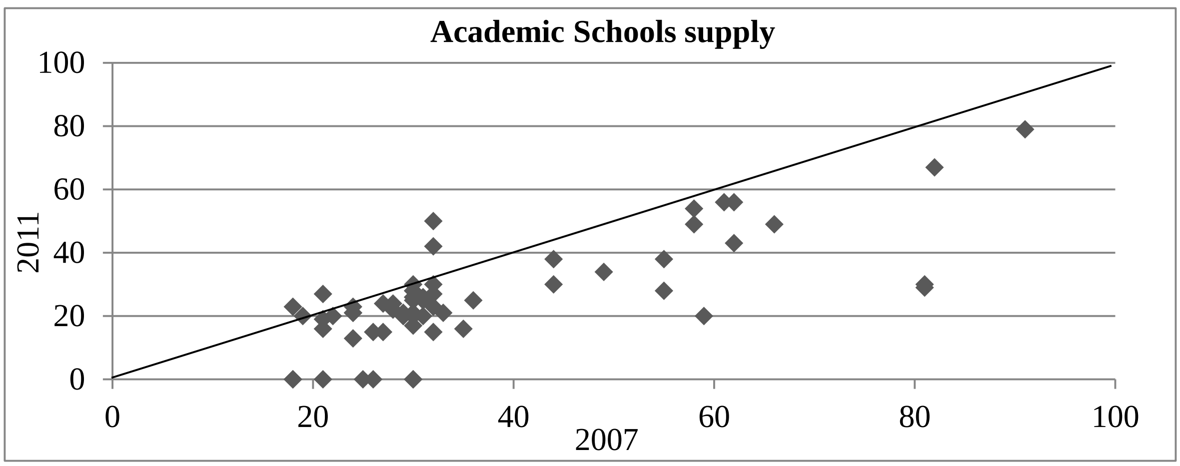
Figure 2.4: Academic Schools' demand, supply and relative excess demand in years 2007 and 2011, Czech districts

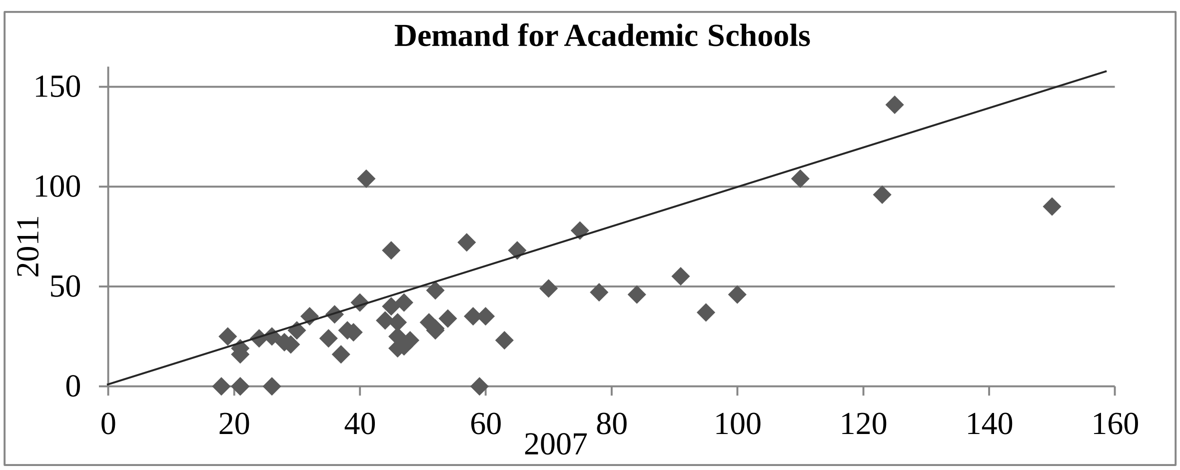


Source: Ministry of Schooling and Youth of the Czech Republic

Figure 2.5: Academic Schools' demand, supply and relative excess demand in years 2007 and 2011, Slovak districts







Source: The Institute of Information and Prognoses of Education

Table 2.3: Students' background characteristics

	Slovakia		Czech Republic	
	2007	2011	2007	2011
Gender (% of boys)	51.2	50.1	52.7	51.6
Age (years)	10.40	10.45	10.26	10.38
Books at home (%)				
0-10	11.4	11.4	6.5	6.2
11-25	32.5	23.9	26.4	19.6
26-100	36.2	38.1	40.1	40.6
101-200	11.8	15.9	15.6	18.4
200+	8.1	10.6	11.5	15.2
Municipality size of school location (%)				
more than 500 000 people	7.4	5.9	8.8	10.3
100 001 to 500 000 people	4.0	6.4	9.7	8.2
50 001 to 100 000 people	11.7	13.7	12.2	12.0
15 001 to 15 000 people	23.2	22.4	20.3	19.7
3 001 to 15 000 people	18.5	22.3	24.2	23.2
3 000 people or fewer	35.2	29.4	24.7	26.6

Source: TIMSS 2007 and TIMSS 2011.

Table 2.4: Probability of student's admission to Academic School (marginal effects after Probit model)

Dependent variable: Probability of being admitted	(1)
Math score from TIMSS	0.067***
	(0.005)
Boy	-0.015**
	(0.007)
Age	-0.007
	(0.009)
Books at home:	
10-25 books	0.035
	(0.036)
26-100 books	0.054*
	(0.030)
101-200 books	0.100**
	(0.051)
more than 200 books	0.157**
	(0.065)
District controls	Yes
N	3,655
Pseudo R^2	0.180

Note: *p<0.1 **p<0.05 ***p<0.01

Table 2.5: Impact of admission exams on student achievement (treatment effect under various specifications)

Dependent variable: standardized math test scores	(1)		(2)		(3)		(4) without 4 districts[1]	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Treatment	0.110*	0.143**	0.187***	0.188***	0.073	0.189**	0.085	0.188**
	(0.065)	(0.058)	(0.058)	(0.049)	(0.087)	(0.091)	(0.090)	(0.095)
Treatment*Excess demand					0.202*	0.024	0.277*	0.037
					(0.117)	(0.124)	(0.149)	(0.144)
Excess demand					0.091***	0.045	0.074	0.114
					(0.029)	(0.032)	(0.075)	(0.081)
Excess demand*Slovakia					0.058	0.081	-0.199	-0.056
					(0.117)	(0.101)	(0.174)	(0.151)
Excess demand*2007					-0.042	0.030	-0.031	-0.005
					(0.071)	(0.063)	(0.080)	(0.073)
Controls			Yes	Yes	Yes	Yes	Yes	Yes
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,223	9,625	9,218	9,620	8,636	8,994	7,869	8,197
Adj. R^2	0.107	0.110	0.233	0.228	0.229	0.225	0.215	0.216

Note: Standard errors robust to clustering at the year and district level in parentheses. (*p<0.1 **p<0.05 ***p<0.01)

[1] We exclude 4 outliers in excess demand: Praha, Brno, Bratislava, and Usti nad Labem.

Table 2.6: Impact of admission exams on student achievement (Quantile regression)

Dependent variable:	Girls					Boys				
Standard. math test scores	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Treatment	0.024	0.069	0.125**	0.119	0.177**	0.028	0.127***	0.256***	0.283***	0.223**
	(0.128)	(0.050)	(0.063)	(0.083)	(0.065)	(0.134)	(0.035)	(0.053)	(0.069)	(0.078)
Treatment*Excess d.	0.171	0.173***	0.162*	0.160**	0.253*	0.125	0.084	-0.080	-0.028	0.038
	(0.209)	(0.039)	(0.090)	(0.068)	(0.137)	(0.180)	(0.074)	(0.086)	(0.151)	(0.069)
Excess demand	0.070	0.085**	0.094**	0.108**	0.087*	0.096*	0.039	0.044	0.028	0.076
	(0.073)	(0.031)	(0.045)	(0.037)	(0.046)	(0.059)	(0.022)	(0.034)	(0.039)	(0.049)
Excess demand*Slovakia	0.126	0.106	0.040	-0.089	-0.002	0.064	0.142	0.035	0.027	0.032
	(0.215)	(0.078)	(0.052)	(0.089)	(0.072)	(0.041)	(0.075)	(0.069)	(0.045)	(0.083)
Excess demand*2007	-0.027	-0.045	-0.003	-0.015	-0.027	0.202**	0.056	0.018	-0.009	-0.016
	(0.162)	(0.076)	(0.070)	(0.048)	(0.068)	(0.068)	(0.048)	(0.033)	(0.067)	(0.057)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8,636	8,636	8,636	8,636	8,636	8,994	8,994	8,994	8,994	8,994

Note: Standard errors robust to clustering at the year and country level in parentheses. Standard errors are bootstrapped using 50 replications. (*p<0.1 **p<0.05 **p<0.01)

Table 2.7: Impact of admission exams on student achievement, controlling for the distance from the admission threshold

Dependent variable: Standard. math test scores	(1)		(2)		(3) Obs. around admission margin[1]	
	Girls	Boys	Girls	Boys	Girls	Boys
Treatment	0.212***	0.187***	-0.065	-0.025	-0.122	-0.062
	(0.082)	(0.068)	(0.114)	(0.099)	(0.077)	(0.074)
Treatment*Distance	-0.002**	-0.002**	0.001	-0.001		
	(0.001)	(0.001)	(0.002)	(0.002)		
Treatment*Distance* Excess d.			-0.005***	-0.002		
			(0.002)	(0.002)		
Treatment*Excess d.			0.517***	0.334**	0.577***	0.271*
			(0.171)	(0.149)	(0.165)	(0.146)
Distance	-0.029***	-0.031***	-0.030***	-0.032***		
	(0.003)	(0.002)	(0.003)	(0.003)		
Distance *Excess d.			0.001	0.001		
			(0.001)	(0.001)		
Excess demand			0.041	0.097	0.158*	0.148
			(0.077)	(0.068)	(0.076)	(0.081)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes
District dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	7,698	7,927	7,205	7,413	1,025	963
Adj. R^2	0.710	0.728	0.711	0.729	0.345	0.362

Note: Standard errors robust to clustering at the year and district level in parentheses. (*p<0.1 **p<0.05 ***p<0.01)
[1] Included only students 5 percentage points above and below admission margin. Standard errors clustered at the year and country level.

3 Chapter

Children Left Behind: Self-confidence of Pupils in Competitive Early Tracking Environments

Co-authored by Filip Pertold and Michael Smith

3.1 Introduction

Economic literature often relates gender gaps in labor market outcomes to higher male competitiveness driven by their higher confidence (Niederle and Vesterlund 2011). Educational research has also shown that one's self-confidence can play an important role in the choice of educational fields and performance during exams (Kinzie, Delcourt and Powers 1994; Dean and Fleckenstein 2007; Marra et al. 2009). In particular, the evidence suggests that gender differences in self-confidence contribute to girls' underperformance in math and science exams (Louis and Mistele 2012) and shy them away from competitive educational fields (Niederle and Vesterlund 2007). This is in line with social cognitive theory suggesting that self-confidence is essential for students' motivations to achieve (Zimmerman 2000; Bandura 2001). Although the importance of self-confidence is well recognized in the social sciences, previous literature has not addressed whether individuals' self-confidence is an inherent trait, or may be also a product of environments induced, for example, by educational systems.[1]

[1] Bandura (1997) posited that self-efficacy, or self-confidence in one's own capacities, was the result of observed performance and other social-psychological factors (Usher and Pajares 2009). Previous literature also discuss problem

Early tracking processes are high stakes tournaments that produce unequal outcomes in educational attainment because they sort pupils into better or worse schools, purportedly by ability (Gamoran and Mare 1989; Betts 2011). In this context, track placement into a 'better' or 'selective' school means that such schools achieve higher levels of academic performance (e.g. Werfhorst and Mijs 2010), higher rates of college entry (Shavit, Arum and Gamoran 2007), and lead to greater job prospects. While there has been conducted a substantive research on the effects of track placement on subsequent achievement, the literature has ignored how the competitive allocation process itself may impact children left behind, i.e. children who do not apply to selective schools but are nonetheless observers of the process and outcome of the competition.

In the paper, we present evidence using longitudinal data that the process of competitive environment induced by early tracking negatively affects pupils' self-confidence in mathematics. In particular, we show that having classmates who are applying to a selective school negatively impacts the self-confidence of those who did not apply. This effect is particularly strong for girls and in classes where classmates are successful in the admission process. Our results show that the effect persists over time, and is robust to controlling for GPA, number of slots in selective schools, parental involvement in application process, class level of self-confidence or objectively measured math abilities. The negative effect is also slightly stronger when parents do not want their children to apply, which is in line with the fact that parental background and decision of parents play crucial role in application process. Our findings add into the discussion about gender differences in responses to competitive environments (Gneezy, Niederle and Rustichini 2003; Niederle and Vesterlund 2010), as well as gender differences in inequality aversion, which are particularly acute in early adolescence (Fehr and Schmidt 1999; Fehr, Bernard and Rochenbach 2008).

In our set-up, children and their parents have the opportunity to decide whether to apply during the 5th grade to a prestigious academic junior high school, attendance at which brings social prestige, higher quality academic instruction, and better educational opportunities. In fact, 90 percent of applicants in our sample report that their application decision was driven by the decision

---

of nature vs. nurture in gender differences in competitiveness (Niederle and Vesterlund 2010), but not for the self-confidence as such.

of parents.[2] If children do not apply to such schools, they by default continue their studies at an elementary school until the end of compulsory education. The set-up is similar to situations where children apply for placement into classrooms for the "gifted" or other advanced placement schemes, while children who choose to not apply or who are not accepted maintain their status quo.[3] The children left behind, who do not apply for academic track constitute in our case 87% of the school population.

Our empirical strategy takes advantage of a longitudinal dataset based on the participants in the TIMSS (Trends in International Mathematics and Science Study) survey in the 4th grade and its follow-up study CLoSE (Czech Longitudinal Study in Education) in the 5th grade and 6th grade. Based on this panel, we can take advantage of specific questions about pupils' self-confidence, as well as of detailed measures of academic achievement, social and parental background, and classroom characteristics. All these variables are used to control for the potential sorting of students into classes and explain a large portion of variation in pre-application self-confidence across classes. All pupils in our sample are situated in standardized, non-differentiated classrooms, with school placement based mainly on catchment area. Early in the 5th grade, these pupils then make decisions on whether to apply to academic junior high schools, each of which designs its own entrance exam specific to the school. Our follow-up panel of TIMSS participants was fielded later on in the 5th grade, after application decisions were made and after they received notice of acceptance. Our main outcome variable is the change in self-confidence in mathematics between 4th and 5th (6th) grades, that is, before and after the application process is complete.

We focus on the competitive allocation process because such settings may reveal different preferences in the distribution of payoffs – maximin, efficiency or fairness preferences, which has

---

[2] Previous evidence and our data suggest that parental background and 4th grade GPA are the main predictors for the application decisions. It is therefore reasonable to assume that 11 years old have rather passive role in the decision to apply and no decision to apply would be made without support of their parents.

[3] Our institutional set up is the application process for entering 8-year gymnasia (academic junior high schools) in the Czech Republic, which is a similar process as in a number of other countries with early tracking. These junior high schools are widely regarded as the most prestigious public schools in the country. Acceptance into these schools is highly predictive of future university attendance and high occupational and income status, and is strongly determined by family background (Greger 2015). The Czech system is based on the German model of education, where early tracking processes often take place in the 4th or 5th grade (around 9-11 years of age). These early tracking systems have been a recent focal point of research on educational efficiency and inequality (Gamoran and Mare 1998; Heubert and Hauser 1998; Hanushek 2006; Woessmann 2007).

been the subject of intensive debate in economics (Engelmann and Strobel 2004; Fehr, Naef, Schmidt 2006; Bolton and Ockenfels 2006). In the related literature, child psychologists have shown that the competitive environments induce social comparisons among children that lead to negative emotional states, such as envy, gloating or taking delight of someone else misfortune (Smith and Kim 2007). At the age of 10, children tend to have high inequality aversion and are not willing to take spiteful decisions, which put other children into disadvantaged positions. When children lose, they feel worse about it if others win than when others also lose (Steinbeis and Singer 2013). This is analogical to the situation in early tracking environments when children, who do not apply to selective schools, for example due to the insistence of their parents for the status quo, watch their peers apply and be admitted to selective schools that offer superior educational experience and prospects. In this case, inequality aversion is not revealed as an ex ante preference, but as an emotional response to the observation of a competitive process with unequal payoffs, which is measured in terms of the reduction in self-confidence of children who are left behind from that competition.

Recent experimental research in economics suggests that there may be important gender differences in inequality aversion, but that these vary greatly by the type and specific parameters of the competitive environment (Feingold 1994; Eagly and Wood 1999; Eckel and Grossman 1998; Andreoni and Vesterlund 2001). The literature seems consistent in finding that, already at an early age, boys respond more positively than girls to winner-take-all tournaments, such as the desire to compete, even if there may not be gender differences in actual performance (Croson and Gneezy 2009). The literature has also found that while preferences for egalitarian distributions increase with age among both girls and boys and have similar propensities at age 7-8, boys are less averse against disadvantageous inequality to in-group members, such as classmates, compared to out-group members, while girls' degree of inequality aversion does not vary by in-group or out-group status (Fehr, Bernhard, and Rochenbach 2008).

Lastly, our study on changes in self-confidence as a responsive form of inequality aversion is also motivated by the role of self-confidence as a non-cognitive skill that can significantly affect future academic achievement and employability (Norman and Hyland 2003; Pool and Sewell 2007; Andrews and Higson 2008). There are also gender differences in the development of soft skills at an early age, with girls exhibiting more responsibility than boys (Montgomery 2005).

Heckman and Kautz (2012) have argued strongly in favour of the long-term benefits of soft skills, but they operationalize soft skills uni-dimentionally (Big 5 personality traits), without taking into account self-confidence, self-efficacy and related concepts. Our study contributes to this literature by examining how competitive environments in a common institutional setting can contribute to changes in self-confidence, and thus the formation of one type of soft skill during the developmental stage of early adolescence.

In the following, we overview the institutional context of our study and the key variables. The empirical strategy specifies the baseline model tested, the results of which are then reported along with empirical extensions that enable us to differentiate the roles of the competitive process and outcome of competition on self-confidence, as well as the persistence of the effect in the 6th grade. The conclusion assesses the robustness of the results and points to future directions for follow-up studies.

3.2 Institutional background and data

The Czech Republic as well as many other, especially Central European, countries provide pupils with the possibility to transfer to elite junior high schools after the completion of primary education in the 5^{th} grade, i.e. usually around the age of 11. This early-tracking school system can be contrasted with comprehensive school systems, such as in Finland, in which all pupils are educated in one school track up to the age of 15 or 16, or other standardized school systems, such as in the United States, England, or South Korea, where primary and middle schools (up to the age of 15 or 16) offer relatively standardized curricula for all students. In the early-tracking school system, since the allocation to the academic track is nonrandom, early tracking introduces selection in the schooling process at very early age, particularly by family background (Brunello and Checchi 2007; Betts 2011).

This paper focuses on the Czech school system in which pupils in the 5^{th} grade have an option to apply to academic junior high schools. Typically, slightly less than 20% of pupils apply to these schools and more than half are admitted, such that about a tenth of former fifth graders end up in this academic track in the next school year. Pupils who did not apply or were not admitted continue in the same track up to the end of lower secondary education, i.e. up to the 9^{th} grade. Pupils who apply have to usually pass the high-stake entrance exams that are administered at the end of April.

The admission decision is then based on results from entrance exams and on the primary school grade point average (GPA). Although the application decision is announced before mid-March, there is evidence that pupils start preparing for entrance exams, usually though demanding private courses, already at the beginning of the school year in September. Federičová and Münich (2014) show that half of all pupils who applied to academic junior high schools dedicated some time almost every day to preparing for entrance exams in the semester beforehand. Thus, at least 6 months before the entrance exams, fifth graders are divided into two groups: those who are applying, and thus preparing to take the entrance exams, and those who are not.

To examine the effect of this class division on students' self-confidence, we explore Czech panel data collected before and after the selection process in the 5th grade. The panel data consists of two datasets, the international TIMSS survey held in 2011 and its follow-up study, CLoSE (Czech Longitudinal Study in Education). TIMSS tests a nationally representative sample of pupils in the 4th grade in math and science in four year cycles. Together with these test scores, TIMSS collects school, parental, and student questionnaires that provide detailed information about pupils' school and socio-economic background as well as their perception of their own study aptitudes. The sample of Czech fourth graders tested by TIMSS in 2011 then completed the CLoSE survey in the 5th grade, which focused on the application, preparation and admission process of pupils. Of 4,578 students TIMSS 2011, 3,681 students were followed up by the CLoSE survey in the 5th grade[4].

3.2.1 Self-confidence measure

The challenging process of preparing for and dealing with the prospects of success or failure in high-stakes entrance exams affects not only those who applied but all students in the class. Different degrees of competition, indicated by the number of students in each class taking part in the entrance exams, can alter self-confidence of those students who are left behind. The surveys in both the 4th and the 5th grades asked pupils to self-assess their own performance in mathematics (i.e. *Do well in math*). In the 4th grade, this is indicated by pupils' responses to the statement "I

---

[4] A sub-sample of the 5th graders – i.e. 2837 students – were further questioned in the subsequent school years. To examine the long-term effects of tracking, we provide also the main descriptive statistics for 6th graders. However, as the sub-sample of the 6th graders is non-random, we do not include this data in our results.

usually do well in mathematics" and in the 5th grade to the statement "I was always good at math." As these statements of self-assessment differ between the 4th and 5th grade, we also analyse pupils' responses to similar statements about math and science[5]. The response categories to these statements in both surveys are based on the same 4-point scale: *Agree a lot*=1, *Agree a little*=2, *Disagree a little*=3, and *Disagree a lot*=4.

Table 3.1 presents the descriptive statistics from both TIMSS and CLoSE of pupils' self-confidence in the 4th, 5th, and 6th grades by gender and the competitive level of the class. Since we are interested in how the competitive environment impacts pupils left behind, these statistics focus on the self-confidence only of those pupils who did not apply. Between the 4th and 5th grades, the averages of the self-confidence index increase, indicating a decline in pupils' self-confidence. This drop in self-confidence also continues in the 6th grade. Table 3.1 further indicates that girls not only have lower self-confidence in math than boys in the 4th grade (a trend common in the majority of countries participating in TIMSS), but they also experience a larger decline between the 4th and 5th grades that leads to even greater gender differences in self-confidence in the 5th grade.

Table 3.1: Descriptive characteristics of the index of self-confidence by gender and the number of applicants in the classroom, only for pupils who did not apply (using individual weights)

Index	Girls						Boys					
	0	1	2	3	4	>5	0	1	2	3	4	>5
4th grade	1.79 (.810)	1.57 (.747)	1.83 (.905)	1.88 (.790)	1.89 (.881)	1.92 (.822)	1.57 (.747)	1.61 (.760)	1.66 (.808)	1.66 (.877)	1.58 (.692)	1.71 (.827)
5th grade	2.27 (.883)	2.29 (.824)	2.39 (.858)	2.39 (.839)	2.38 (.798)	2.46 (.866)	1.97 (.836)	1.95 (.850)	1.87 (.765)	2.04 (.982)	1.84 (.736)	2.10 (.901)
6th grade	2.26 (.901)	2.22 (.843)	2.34 (.860)	2.27 (.844)	2.33 (.912)	2.33 (.885)	1.89 (.875)	1.81 (.775)	1.94 (.864)	1.85 (.921)	1.80 (.733)	2.13 (.994)
N	284	235	168	186	156	369	330	270	187	197	159	345

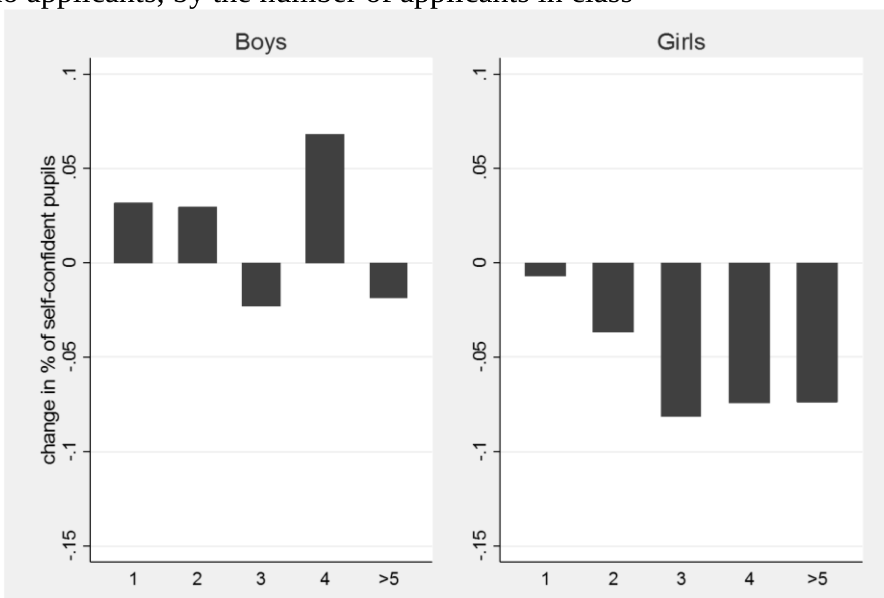
Note: Standard deviations are in parenthesis.

To facilitate the interpretation of results, in the empirical analysis we converted the index into a dichotomous variable indicating low or high levels of self-confidence. This transformation enables us to classify those who changed their self-assessment from high to low between the 4th

[5] TIMSS survey includes pupils' responses to 3 other self-assessments: 1) I usually do well in science; 2) I learn things quickly in mathematics; 3) I learn things quickly in science. The fourth index refers to pupils' relationship towards mathematics and is based on pupils' response to statement: I enjoy learning mathematics. Pupils' responses to these statements are again based on a 4-point scale.

and 5th grades as negative switchers of self-confidence. Figure 3.1 depicts the change in pupils' self-confidence between the 4th and 5th grades with respect to the degree of change in classes with no applicants. For boys, the drop in self-confidence does not differ by the competitive level of the class. However, girls in classes with more than two applicants experience a decline in their self-confidence by about 8 p.p. relative to girls in classes with no applicants.

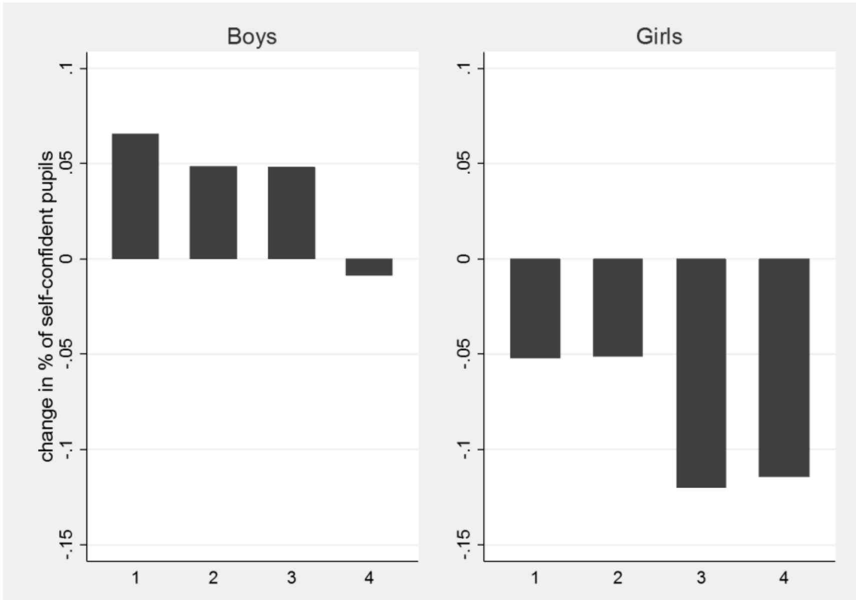
Figure 3.1: Change in self-confidence of pupils who did not apply with respect to the change in classes with no applicants, by the number of applicants in class



Regarding the number of admitted pupils in class, Figure 3.2 depicts the change in self-confidence across classes with different numbers of admitted pupils in comparison to those classes in which no one was admitted, and hence, no one is leaving the class[6]. Although boys in classes with a higher number of admitted pupils do not differ in self-confidence with respect to boys in classes with no admitted classmates, girls seem to be prone to declines in their self-confidence according to the number of admitted pupils in class.

[6] Here, we refer to the loss of classmates to selective academic junior high schools. Pupils can also experience the loss of classmates e.g. when they move to a different city, but because such changes are few and heterogeneous, we do not involve them in our analysis.

Figure 3.2: Relative change in the self-confidence of pupils who did not apply (normalized by the change in classes with no admitted pupils), by the number of admitted pupils in the class



3.2.2 Assessment data

Two academic assessment measures are available in the data: the math test score[7] collected in the 4th grade by TIMSS, and the GPA[8] that pupils received in the 4th and 5th grades. Descriptive statistics of these measures – as well as of other individual and class characteristics – are presented in Table 3.2, which shows that boys outperform girls in the math test score by nearly 0.2 standard deviations, while girls receive better grades than boys in math and in the overall GPA in both the 4th and 5th grades. Pupils' socio-economic background – measured by the share of pupils with parents with a university education – is evenly distributed for boys and girls. Regarding the selection process, on average 3 pupils in each class apply to academic junior high schools and almost 2 pupils are admitted. Non-applying girls experience a slightly more competitive environment than boys in terms of the number of applicants and the success rate of their

[7] The math test score from TIMSS is normalized to a mean of 0 and a standard deviation of 1.
[8] The measure of GPA is constructed as an average of final grades from Mathematics, the Czech language and other foreign language. In the Czech Republic, grades are distributed on a 5-point scale from 1 - the best grade - to 5.

classmates. Lastly, a greater share of girls report that their parents did not want them to apply, possibly due to cultural norms, financial costs, risk of failure, prevention of stress, etc.

3.3 Empirical strategy

In our estimation, we employ the level of self-confidence in the 5th grade as our main left hand side variable. The variable is dichotomous, referring to high or low levels of self-confidence in mathematics. The key right hand side variables are set of dichotomous dummies indicating the number of classmates applying to junior academic high schools. For subsequent analyses, we also constructed similar dummies for whether the number of admitted pupils was 0, 1, 2, … 5. The main control variables are the level of self-confidence in the 4th grade, GPA in the 4th grade, the TIMSS math score indicating math ability, parental education (two dummies for whether each completed tertiary education), gender, and age. We also control for commensurate classroom characteristics (class averages in 4th grade TIMSS math scores, 4th grade GPA, and 4th grade self-confidence, and the number of pupils in the class). Lastly, we also include as controls additional attitudinal variables ("other indexes" in Table 3.3) relating to math and science: whether the pupil reports doing well in science, learning things quickly in science, learning quickly in math, and enjoying math.

One identification issue of our setup is the degree to which our control variables can explain the potential sorting of students into classes and schools, where change in the number of applied students and change in the self-confidence can be driven by common unobserved factors. Although we cannot directly test for this sorting, we analysed to what extent our control variables explain variation in the number of applied students in the class. As shown in Table 3.8 in the Appendix, we explain 40% of the variation and more importantly, the unexplained part of the number of applicants in a given class is not correlated with class level self-confidence. This suggests that our control variables are able to capture most of *ex-ante* association between the pre-treatment level of self-confidence and the number of applicants across classes. The second identification issue is what mechanism stands behind the effect of the number of applicants in a class on the change in self-confidence. We discuss several potential mechanisms in the results, such as the role of parents in application decision or a change in GPA.

For our analysis, we employ both probit and linear probability models, and compare the coefficients as a matter of robustness. Because we find that the results for both models are substantively identical, we mainly report the LPM results due to their ease of interpretation and the computational tractability of the model. The model can thus be expressed in the simple form:

$$y_{i,t,c} = \alpha + \beta y_{i,t-1,c} + \#applied_{t,c} + X_{i,c} + X_c + \varepsilon_{i,t,c}$$

where the self-confidence *y* of pupils *i* in classrooms *c* in grade *t* (i.e. the 5th grade) is predicted by those pupils' self-confidence in the 4th grade (*t-1*), the number of pupils applying to the selective schools in the 5th grade *(#applied*), the vectors of both individual and family characteristics of the pupil as well as classroom characteristics *X*, and the error term. We also run the model separately for boys and girls, as well as together, in which case gender is included in the model.

3.4 Results

Our main results are summarized in Table 3.3. In order to measure the change in self-confidence for those who did not apply to academic junior high schools, we apply firstly the linear probability model presented in the first three columns. Controlling for the individual and average class abilities in math, as well as other individual characteristics described above, Column 1 reveals that self-confidence declines with the increasing number of applicants in class. Thus, those pupils who did not apply and experience high level of competition in their class (i.e. class with 5 and more pupils who applied) have on average a 7% higher probability of experiencing a decline in self-confidence compared to pupils in classes with no applicants. Dividing pupils by gender in Columns 2 and 3, the results indicate that this decline in self-confidence is experienced by girls in all modelled situations where there are at least three applicants in the classroom, whereas the competitive level of class seems to have no measurable effect on the self-confidence of boys. For girls, the probability of experiencing a decline in self-confidence between classes with 3 and more applicants and no applicants is about 10%.

The individual characteristics used as control variables in the model have the expected signs in all regressions (see Table 3.9 in Appendix): self-confidence and math score in the 4th grade positively affect the self-confidence in the following grades, whereas higher grade average in the

4th grade decreases the self-confidence; boys are more likely to report self-confidence than girls, and parents education seems to have no impact on pupils' self-confidence.

Table 3.3: Change in self-confidence of pupils who did not apply, LPM and marginal effects after Probit

Self-confidence (5th grade)	LPM			Probit		
	All	Girls	Boys	All	Girls	Boys
Number of applicants (0 as control group)						
1	-0.014	-0.024	-0.006	-0.014	-0.030	-0.003
	(0.027)	(0.042)	(0.035)	(0.027)	(0.041)	(0.034)
2	-0.016	-0.064	0.036	-0.016	-0.069	0.034
	(0.030)	(0.046)	(0.039)	(0.029)	(0.045)	(0.037)
3	-0.060**	-0.090**	-0.034	-0.058**	-0.087**	-0.039
	(0.029)	(0.044)	(0.038)	(0.029)	(0.043)	(0.038)
4	-0.032	-0.115**	0.043	-0.033	-0.115**	0.040
	(0.032)	(0.049)	(0.041)	(0.031)	(0.048)	(0.039)
5 or more	-0.069***	-0.085**	-0.056	-0.071***	-0.088**	-0.060*
	(0.026)	(0.039)	(0.035)	(0.026)	(0.039)	(0.035)
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Class characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Other Indexes (4th grade)	Yes	Yes	Yes	Yes	Yes	Yes
N	2479	1216	1263	2479	1216	1263
R^2	0.265	0.255	0.248			

Note: Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Using the same identification strategy, the second half of Table 3.3 summarizes the results of marginal effects according to the Probit model. In the substantively same way as the linear probability model, the results show a significant decline in the self-confidence of pupils with the increasing level of competition in the class. Again, girls seem to be particularly affected by this class environment. This decline in self-confidence may be triggered by several different mechanisms, such as the natural increase in competition in class, the leaving of classmates/friends from class, a more demanding teacher policy due to entrance exams to academic junior high schools (especially in classrooms where more pupils are preparing to apply for those exams), etc. Our subsequent models are designed to tease out which of these mechanisms are driving our results, and to detect the qualities of pupils that are mostly likely to be affected by this early-tracking school policy. As mentioned above, because LPM and probit models yield to similar results, for simplicity we further present the results only for LPM.

We continue by differentiating the effect of the class environment in which peers are preparing for the demanding entrance exams, and the effect of payoffs, caused by the loss of classmates at the end of the school year due to their success in the entrance exams. The coefficients for payoffs can be potentially seen as expressions of inequality aversion, as it is a direct response to the future sorting of classmates into selective and status quo schools.

Table 3.4: Change in self-confidence of pupils who did not apply, controlling for the number of applicants and admitted pupils in the classroom.

Self-confidence (5th grade)	(1) All	(2) Girls	(3) Boys
Number of applicants (0 as control group)			
1	-0.017	0.049	-0.080*
	(0.036)	(0.056)	(0.046)
2	-0.034	0.021	-0.076
	(0.043)	(0.068)	(0.055)
3	-0.056	0.013	-0.134**
	(0.043)	(0.066)	(0.057)
4	-0.030	-0.004	-0.059
	(0.047)	(0.073)	(0.060)
5	-0.044	0.039	-0.124**
	(0.046)	(0.071)	(0.061)
Number of admitted pupils (0 as control group)			
1	0.003	-0.108**	0.104**
	(0.034)	(0.053)	(0.044)
2	0.034	-0.084	0.149***
	(0.041)	(0.063)	(0.054)
3	-0.023	-0.139**	0.096*
	(0.042)	(0.065)	(0.055)
4	-0.034	-0.136*	0.064
	(0.051)	(0.079)	(0.066)
5	-0.050	-0.139*	0.026
	(0.051)	(0.076)	(0.068)
Individual characteristics	Yes	Yes	Yes
Class characteristics	Yes	Yes	Yes
Other Indexes (4th grade)	Yes	Yes	Yes
N	2479	1216	1263
R^2	0.267	0.259	0.255

Note: Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

In the regression framework, we add dummy variables representing the number of successful applicants to the model summarized in Table 3.3. This approach enables us to differentiate two mechanisms of change in self-confidence due to the competitive process (i.e. classmates preparing for exams) and the effect of classmates leaving the class. Table 3.4 summarizes these results. After controlling for both the number of admitted pupils and pupils in the classroom who applied, the overall effects for boys and girls together in the first column are insignificant. This, however, masks key gender differences expressed in Columns 2 and 3. Girls seem to be unaffected by the number of applicants in their class, but their self-confidence does significantly decline with the increasing number of classmates that are leaving the class. That is, girls are more inequality adverse: they are more likely to experience declines in self-confidence due to the loss of peers. For boys, the estimated coefficients go in the opposite direction: they are negative for the number of applicants and positive for admitted classmates. However, in many predicted scenarios (e.g. one applicant in the class, who was accepted), the coefficients cancel each other out, such that the net effect in most scenarios is negligible, or zero.

3.4.1 Change in GPA

Another mechanism possibly contributing to reduced self-confidence among pupils in classes with a higher number of applicants may be the change in grades given to students between the 4^{th} and 5^{th} grades. For example, the more demanding class environment imposed by teachers in classes with a higher number of applicants – and hence a higher number of pupils preparing for high-stakes entrance exams – may naturally cause a drop in grades of those who did not apply as they are excluded from the preparation process for these exams. Pupils usually dedicate effort and extra time to prepare for the exams, which would likely boost their grades. This change in grades may also negatively impact the self-confidence of those left behind. Alternatively, however, pupils who are disappointed that they were discouraged from applying might change their effort, and their GPA can drop. In both cases, GPA can potentially explain a part of the decline in self-confidence.

To test for that, we include in our main empirical specification the change in math grades between those years. The reference group is constituted by pupils who improved their grades from the 4^{th} to the 5^{th} year. The results are presented in Table 3.5. Although the drop in math grades between the 4^{th} and 5^{th} year induces the significant drop in self-confidence, controlling for this change in GPA does not diminish the effect of competition in class. In other words, the negative effect of the number of applicants in the classroom on the level of self-confidence is not influenced by the change in GPA, representing here the more demanding teacher policy or other discouragement leading to the drop in GPA.

Table 3.5: Change in self-confidence of pupils who did not apply, the change in GPA.

Self-confidence (5th grade)	(1) All	(2) Girls	(3) Boys
Number of applicants (0 as control group)			
1	-0.019	-0.034	-0.007
	(0.027)	(0.043)	(0.034)
2	-0.011	-0.062	0.045
	(0.030)	(0.046)	(0.039)
3	-0.064**	-0.092**	-0.040
	(0.029)	(0.044)	(0.037)
4	-0.038	-0.129***	0.046
	(0.031)	(0.049)	(0.041)
5	-0.072***	-0.083**	-0.065*
	(0.026)	(0.040)	(0.035)
Change in math grade (controlling for better grade in the 5th grade)			
No change in math grade	-0.126***	-0.094*	-0.162***
	(0.033)	(0.050)	(0.043)
Worse math grade in the 5th gr	-0.247***	-0.228***	-0.277***
	(0.035)	(0.053)	(0.045)
Individual characteristics	Yes	Yes	Yes
Class characteristics	Yes	Yes	Yes
Other Indexes (4th grade)	Yes	Yes	Yes
N	2431	1193	1238
R^2	0.289	0.274	0.282

Note: Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

3.4.2 Intervention of parents in application decision

Another potential factor that could explain the role of early tracking on student self-confidence is the intervention of parents. In general we assume that parents play an important role for decision to apply and there is an evidence that support this decision. While it is somewhat obvious why parents encourage their children to apply (enhanced educational and labor market opportunities derived from placement in the best schools, and with no material application cost), it is less obvious why other parents take the opposite approach, as it may appear to be irrational. Parental education per se does not play a role for the change in self-confidence either boys or girls (see Table 3.3).

Parents may not want their children to go to selective schools due to their preference for value transmission (parents prefer that their children have a similar education and experiences as themselves), the cost of stress of applying and then attending demanding schools, their unwillingness or inability to invest in private preparatory courses to train pupils in the exam process, or their belief that their children are not good enough or would not succeed in selective schools.

Table 3.6: Intervention of parents in interaction with the number of applicants in the class

Self-confidence (5th grade)	(1) All	(2) Girls	(3) Boys
Parents didn't want me apply=1	0.045	0.061	0.020
	(0.038)	(0.060)	(0.049)
Number of applicants (0 as control group)			
1	0.019	-0.013	0.039
	(0.042)	(0.068)	(0.054)
2	0.022	-0.034	0.076
	(0.045)	(0.072)	(0.057)
3	-0.044	-0.033	-0.072
	(0.047)	(0.074)	(0.059)
4	0.015	-0.072	0.082
	(0.049)	(0.078)	(0.062)
5	-0.015	-0.040	0.006
	(0.040)	(0.062)	(0.051)
Parents=1 * number of applicants			
1	-0.044	0.007	-0.080
	(0.057)	(0.088)	(0.074)
2	-0.064	-0.025	-0.091
	(0.062)	(0.095)	(0.081)
3	-0.020	-0.069	0.055
	(0.061)	(0.093)	(0.080)
4	-0.081	-0.063	-0.073
	(0.064)	(0.098)	(0.083)
5	-0.115^{**}	-0.094	-0.126^{*}
	(0.052)	(0.079)	(0.068)
Individual characteristics	Yes	Yes	Yes
Class characteristics	Yes	Yes	Yes
Other Indexes (4th grade)	Yes	Yes	Yes
N	2244	1101	1143
R^2	0.264	0.258	0.247

Note: Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

To take the explicit intervention of parents into account for pupils who did not apply (the great majority of the panel), we created a dummy variable indicating that their parents did not want them to apply, and then interacted the variable with the number of applicants in the class. The results, reported in Table 3.6, indicate that parents do influence their children in the face of a competitive environment, particularly boys in the most competitive classrooms. For example, comparing non-applying boys in the most competitive classes (5 applicants or more), those with discouraging parents have by 13% higher probability of experiencing a decline in self-confidence than boys without this parents' intervention. However, the intervention of parents seems to have no significant impact on girls' self-confidence.

Table 3.7: Persistence of the decline in self-confidence until the 6th grade

Self-confidence (6th grade)	(1) All	(2) Girls	(3) Boys
Number of applicants (0 as control group)			
1	0.029	-0.023	0.072*
	(0.030)	(0.048)	(0.038)
2	-0.086**	-0.089*	-0.083*
	(0.034)	(0.052)	(0.044)
3	-0.047	-0.080*	-0.016
	(0.031)	(0.048)	(0.041)
4	-0.032	-0.116**	0.044
	(0.035)	(0.054)	(0.044)
5	-0.051*	-0.069	-0.034
	(0.029)	(0.044)	(0.038)
Individual characteristics	Yes	Yes	Yes
Class characteristics	Yes	Yes	Yes
Other Indexes (4th grade)	Yes	Yes	Yes
N	2050	1013	1037
R^2	0.224	0.201	0.222

Note: Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

3.4.3 Persistence of the effect in the next grade

Using the same specification as for the main model in Table 3.3 (i.e. controlling for the number of applicants in class, 4th grade self-confidence and other 4th grade indexes as well as individual and class characteristics), we further test the persistence of the effect on pupils´ self-confidence until the 6th grade. The persistence of declines in self-confidence is policy relevant, as it would reveal possibly long-term effects on academic achievement. Our approach can eliminate the possibility of only a momentary change in pupils' self-assessment. The results of the persistence effect presented in Table 3.7 confirm the negative impact of the number of applicants in the classroom on the self-confidence of those who did not apply. Moreover, the magnitude of the effect on pupils' self-confidence in the 6^{th} grade is similar to the one in the previous grade, and again with significant impact only in the case of girls.

These results should be, however, taken with caution as not all fifth graders in our sample were followed in the 6^{th} grade. Hence, our subsample in the 6^{th} grade may not be representative and may induce sample selection bias in the results. Nevertheless, as the results are strongly significant and in line with the effects found in the main model, they can be taken as an indication of persistence of the effect to higher grades and of its possible impact on further schooling outcomes.

3.5 Conclusion

Early tracking is a feature of many educational systems around the globe, the typical aim of which is to sort pupils into homogeneous ability groups that may make it more efficient for students to learn and teachers to teach. Our goal is to not assess early tracking as such, but to examine the effect of the competitive application process inherent in early tracking systems on the self-confidence of students who do not apply to such selective schools. Our interest in changes in self-confidence derives from the role of self-confidence as a soft skill that can potentially impact future educational and occupational attainment, as well as a form of aversion to unequal outcomes of competition. To do this, we operationalize the competitive process in terms of the number of pupils who apply, and the number of admitted, in each classroom, while controlling for a variety of individual and contextual variables, including changes in GPA.

We first found that girls, but not boys, are inequality adverse, in terms of experiencing a decline in self-confidence between the beginning and the end of their observation of the competition to selective schools. When we differentiate the number of applicants from the number of successful ones, we find that girls are particularly averse to the latter, i.e. by the unequal payoffs realized in the tracking process, rather than by the number of applicants per se. While boys seem to be affected by the competitive process net of the realized payoffs, in most predicted scenarios the net effect of competition on their self-confidence is rather zero. We also find that the gender difference in the decline in self-confidence, as well as the degree of that decline, is persistent between 5th and 6th grades, indicating that there could be longer term effects of early selection on students' self-confidence.

We believe that these results are relevant to the academic literature in a number of ways. As research increasingly examines the role of soft and hard skills on educational and occupational attainment, it will become increasingly important to understand the role of features of educational systems in shaping key traits like self-confidence in mathematics. Second, gender gaps in the attainment of academic degrees in STEM fields, and the policy goal of increasing women's participation in those fields, suggests it is crucially important to understand why girls might become less self-confident in STEM fields during the critical years of their academic development. And lastly, as scholars continue to debate the degree of inequality and efficiency in educational systems with varying types and degrees of tracking, we show that the role of tracking on the 'socioemotional' development of children warrants scholarly attention.

The results of our analysis come with a number of limitations and caveats. First, besides the grades given by teachers to students, we have little information about changes in teacher policy between the 4th and 5th grades. It is possible that changes in teacher policy, or changes in study materials, would influence students' self-confidence. That, however, would not explain why self-confidence would be so strongly influenced by the intensity of the competitive process. But if it could be shown that there are some unobserved teacher behaviors correlated to the intensity of competition, those unobservables could potentially explain some of the estimated effects we have found in this paper. That would not detract from our overall findings on the impact of early tracking on self-confidence, but would enable us to better differentiate the role of teachers during the competitive process from the competitive process itself.

Second, our study would benefit from an additional follow-up analysis. To what degree does early tracking impact self-confidence in later adolescent years, and how does that self-confidence impact latter educational aspirations and attainment? Given that we have a randomized panel of pre-treatment pupils along with their cognitive and non-cognitive skills, the results of this study would be enriched by a closer examination of the long-term effects of early-tracking competition.

3.A Appendix 3

Table 3.2: Descriptive statistics of individual characteristics by gender, whole sample

	All	Girls	Boys
% of boys in own class	51.6	47.5	55.4
Age	10.37 (.413)	10.44 (.426)	10.30 (.386)
Math score (see footnote 5)	0 (1)	-.090 (.984)	.085 (1.008)
GPA – 4th grade (see footnote 6)	1.63 (.653)	1.54 (.599)	1.71 (.689)
GPA – 5th grade (see footnote 6)	1.86 (.774)	1.74 (.731)	1.97 (.798)
Math grade-4th	1.59 (.732)	1.57 (.705)	1.61 (.756)
Math grade-5th	1.81 (.855)	1.78 (.820)	1.84 (.886)
Father - university education	18.4 % (38.7)	17.7 (38.2)	18.9 (39.2)
Mother - university education	14.8 (35.5)	14.6 (35.3)	15.0 (35.7)
Learn quickly math – 4th grade	1.88 (.918)	2.01 (.931)	1.76 (.888)
Learn quickly science – 4th grade	1.92 (.898)	1.87 (.854)	1.96 (.936)
Do well science – 4th grade	1.72 (.809)	1.69 (.765)	1.74 (.848)
Enjoy math - 4th grade	1.84 (.956)	1.91 (.948)	1.77 (.958)
Class size	17.6 (5.4)	17.7 (5.3)	17.4 (5.4)
Number of applicants in own class	3.0	3.2	2.9
Number of admitted pupils in own class	1.9	2.0	1.8
Parents didn’t want me to apply (% of non-applicants)	56.9	58.1	55.8
N	3681	1817	1864

Note: Standard deviations are in parenthesis.

Table 3.8: Variation in the number of pupils who applied

Number of applicants (on class level)	(1)
Number of free slots in district	1.048***
	(0.271)
Math score (4th grade)	0.077***
	(0.027)
GPA (4th grade)	-0.125***
	(0.047)
University education - father	0.006
	(0.106)
University education - mother	0.301**
	(0.129)
Boy=1	-0.110
	(0.070)
Age	0.065
	(0.072)
Other Indexes (4th grade)	Yes
N	216
R^2	0.432
Adj. R^2	0.402

Note: Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 3.9: Change in self-confidence of pupils who did not apply, LPM and marginal effects after Probit, full specification.

Self-confidence (5th grade)	LPM			Probit		
	All	Girls	Boys	All	Girls	Boys
Number of applicants						
1	-0.014	-0.024	-0.006	-0.014	-0.030	-0.003
	(0.027)	(0.042)	(0.035)	(0.027)	(0.041)	(0.034)
2	-0.016	-0.064	0.036	-0.016	-0.069	0.034
	(0.030)	(0.046)	(0.039)	(0.029)	(0.045)	(0.037)
3	-0.060**	-0.090**	-0.034	-0.058**	-0.087**	-0.039
	(0.029)	(0.044)	(0.038)	(0.029)	(0.043)	(0.038)
4	-0.032	-0.115**	0.043	-0.033	-0.115**	0.040
	(0.032)	(0.049)	(0.041)	(0.031)	(0.048)	(0.039)
5 or more	-0.069***	-0.085**	-0.056	-0.071***	-0.088**	-0.060*
	(0.026)	(0.039)	(0.035)	(0.026)	(0.039)	(0.035)
Other indexes – 4th grade						
Self-confidence	0.113***	0.129***	0.074*	0.076***	0.114***	0.025
	(0.027)	(0.038)	(0.040)	(0.025)	(0.036)	(0.035)
Learn quickly math	-0.074***	-0.099***	-0.048***	-0.064***	-0.089***	-0.042***
	(0.012)	(0.018)	(0.016)	(0.011)	(0.017)	(0.014)
Learn quickly science	0.016	0.018	0.015	0.013	0.015	0.011
	(0.012)	(0.018)	(0.015)	(0.011)	(0.018)	(0.015)
Do well science	-0.004	-0.009	-0.000	-0.004	-0.010	0.001
	(0.012)	(0.019)	(0.016)	(0.012)	(0.019)	(0.016)
Enjoy math	-0.102***	-0.098***	-0.107***	-0.093***	-0.096***	-0.092***
	(0.011)	(0.016)	(0.014)	(0.010)	(0.015)	(0.012)
Individual characteristics						
4th grade math score	0.061***	0.053***	0.069***	0.063***	0.053***	0.072***
	(0.011)	(0.017)	(0.014)	(0.011)	(0.016)	(0.014)
4th grade GPA	-0.127***	-0.126***	-0.126***	-0.118***	-0.125***	-0.104***
	(0.016)	(0.025)	(0.020)	(0.015)	(0.024)	(0.018)
University education - father	0.003	-0.031	0.041	0.000	-0.028	0.032
	(0.026)	(0.040)	(0.033)	(0.026)	(0.039)	(0.033)
University education - mother	-0.015	0.008	-0.048	-0.013	0.002	-0.040
	(0.027)	(0.041)	(0.037)	(0.027)	(0.040)	(0.037)
Boy=1	0.115***	-	-	0.115***	-	-
	(0.017)			(0.017)		
Age	-0.029	-0.042	-0.021	-0.026	-0.041	-0.018
	(0.020)	(0.033)	(0.026)	(0.020)	(0.032)	(0.025)
Class characteristics (4th grade)						
Class GPA	-0.023	0.037	-0.090**	-0.019	0.044	-0.086*
	(0.035)	(0.054)	(0.046)	(0.035)	(0.053)	(0.044)
Class score – math	-0.080***	-0.061*	-0.097***	-0.081***	-0.060*	-0.097***
	(0.023)	(0.035)	(0.031)	(0.023)	(0.035)	(0.031)
Class self-confidence	0.010	0.046	-0.038	0.018	0.039	-0.022
	(0.086)	(0.126)	(0.117)	(0.083)	(0.117)	(0.117)
Class count	0.000	0.002	-0.002	0.000	0.002	-0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
N	2479	1216	1263	2479	1216	1263
R^2	0.265	0.255	0.248			

Note: Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Bibliography

Andreoni, J., and L. Vesterlund. 2001. "Which is the fair sex? Gender differences in altruism." *Quarterly Journal of Economics*: 293-312.

Andrews, J., and H. Higson. 2008. "Graduate employability, 'soft skills' versus 'hard' business knowledge: A European study." *Higher education in Europe* 33 (4): 411-422.

Anger, S., and G. Heineck. 2010. "Do smart parents raise smart children? The intergenerational transmission of cognitive abilities." *Journal of Population Economics* 23 (3): 1105-1132.

Bandura, A. 1997. *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. 2001. "Social cognitive theory: An agentic perspective." *Annual Review of Psychology* 52 (1): 1-26.

Bennett, R. E., R. L. Gottesman, D. A. Rock, and F. Cerullo. 1993. "Influence of behavior perceptions and gender on teachers' judgements of students' academic skills." *Journal of Educational Psychology* 85 (2): 347-356.

Betts, J. R. 2011. "The economics of tracking in education." In *Handbook of the Economics of Education, Vol. 3*, edited by E. Hanushek, S. Machin, and L. Woessmann, 341-381. Amsterdam: North Holland.

Betts, J. R., and D. Morell. 1999. "The determinants of undergraduate grade point average. The relative importance of family background, high school resources, and peer group effects." *The Journal of Human Resources* 34 (2): 268-293.

Bishop, J. H. 1997. "The effect of national standards and curriculum-based exams on achievement." *The American Economic Review* 87 (2): 260-264.

Bishop, J. H. 1999. "Are national exit examinations important for educational efficiency?" *Swedish Economic Policy Review* 6 (2): 349-401.

Bolton, G. E., and A. Ockenfels. 2006. "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment." *American Economic Review*: 1906-1911.

Bond, T. N., and K. Lang. 2013. “The evolution of the black-white test score gap in grades K-3: The fragility of results.” *The Review of Economics and Statistics* 95 (5): 1468-1479.

Brunello, G., and D. Checchi. 2007. “Does school tracking affect equality of opportunity? New international evidence.” *Economic Policy* 22 (52): 781-861.

Charness, G., and U. Gneezy. 2012. “Strong Evidence for Gender Differences in Risk Taking.” *Journal of Economic Behavior & Organization* 83 (1): 50–58.

Coladarci, T. 1986. “Accuracy of teacher judgments of student responses to standardized test items.” *Journal of Educational Psychology* 78 (2): 141-146.

Croson, R., and U. Gneezy. 2009. “Gender differences in preferences.” *Journal of Economic literature*: 448-474.

Crumpton, H. E., and A. Gregory. 2011. “"I'm Not Learning": The role of academic relevancy for low-achieving students.” *The Journal of Educational Research* 104 (1): 42-53.

Cunha, F., and J. J. Heckman. 2007. “The technology of skill formation.” *The American Economic Review* 97 (2): 31-47.

Dean, D. J., and A. Fleckenstein. 2007. “Keys to success for women in science.” In *Women and Minorities in Science, Technology, Engineering, and Mathematics: Upping the Numbers,* edited by R. J. Burke and M. C. Mattis, 28-46. Cheltenham: Edward Elgar Publishing.

Dearden, L., J. Ferri, and C. Meghir. 2002. “The effect of school quality on educational attainment and wages.” *Review of Economics and Statistics* 84 (1): 1-20.

Dunifon, R., and G. J. Duncan. 1998. “Long-run effects of motivation on labor-market success.” *Social Psychology Quarterly* 61 (1): 33-48.

Eagly, A. H., and W. Wood. 1999. “The origins of sex differences in human behavior: Evolved dispositions versus social roles.” *American Psychologist* 54 (6): 408-423.

Eckel, C. C., and P. J. Grossman. 1998. “Are women less selfish than men? Evidence from dictator experiments.” *The Economic Journal* 108 (448): 726-735.

Elliot, A. J., H. A. McGregor, and S. Gable. 1999. “Achievement goals, study strategies, and exam performance: A mediational analysis.” *Journal of Educational Psychology* 91(3): 549-563.

Engelmann, D., and M. Strobel. 2004. “Inequality aversion, efficiency, and maximin preferences in simple distribution experiments.” *American Economic Review* 94 (4): 857-869.

Falch, T., and L. R. Naper. 2013. “Educational evaluation schemes and gender gaps in student achievement.” *Economics of Education Review* 36: 12-25.

Federičová, M., and D. Münich. 2014. “Preparing for the eight-year gymnasium: The great pupil steeplechase.” *IDEA Study 2/2014*. Prague: CERGE-EI.

Fehr, E., and K. M. Schmidt. 1999. “A theory of fairness, competition, and cooperation.” *Quarterly Journal of Economics*: 817-868.

Fehr, E., M. Naef, M., and K. M. Schmidt. 2006. “Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment.” *American Economic Review* 96 (5): 1912-1917.

Fehr, E., H. Bernhard, and B. Rockenbach. 2008. "Egalitarianism in young children." *Nature* 454: 1079-1083.

Feinberg, A. B. and E. S. Shapiro. 2010. “Teacher accuracy: An examination of teacher-based judgments of students' reading with differing achievement levels.” *The Journal of Educational Research* 102 (6): 453-462.

Feingold, A. 1994. “Gender differences in personality: A Meta-analysis.” *Psychological Bulletin* 116 (3): 429-456.

Feinstein, L. 2003. “Inequality in the early cognitive development of British children in the 1970 cohort.” *Economica* 70 (277): 73-97.

Foy, P., J. Galia, and I. Li. 2008. “Scaling the data from the TIMSS 2007 mathematics and science assessments.” In *TIMSS 2007 Technical report*, edited by J. F. Olson, M. O. Martin, and I. V. S. Mullis. Chestnut Hill, MA: Boston College.

Friedel, J. M., K. S. Cortina, J. C. Turner, and C. Midgley. 2007. “Achievement goals, efficacy beliefs and coping strategies in mathematics: The roles of perceived parent and teacher goal emphases.” *Contemporary Educational Psychology* 32: 434-458.

Gamoran, A., and R. D. Mare. 1989. “Secondary school tracking and educational equality: Compensation, reinforcement, or neutrality?” *American Journal of Sociology* 94: 1146-83.

Gneezy, U., and A. Rustichini. 2004. “Gender and competition at a young age.” *American Economic Review* 94 (2): 377-381.

Gneezy, U., M. Niederle, and A. Rustichini. 2003. “Performance in competitive environments: Gender differences.” *Quarterly Journal of Economics* 118 (3): 1049-1074.

Greger, D. 2015. “The Czech Republic: 25 years of educational transformation and growing inequalities 1989–2014.” In *Education in the European Union: Post-2003 Member States*, edited by T. Corner, 75-92. London: Bloomsbury.

Grolnick, W. S., and R.M. Ryan. 1989. “Parent styles associated with children´s self-regulation and competence in school.” *Journal of Educational Psychology* 81 (2): 143-154.

Guyon, N., E. Maurin, and S. McNally. 2012. “The effect of tracking students by ability into different schools: A natural experiment.” *Journal of Human Resources* 47 (3): 684-721.

Hanushek, E. A. 2006. “Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries.” *The Economic Journal* 116: C63-C76.

Hao, L., and D. Q. Naiman. 2007. *Quantile regression*. Thousand Oaks, CA: Sage publications.

Hastings, J. S., and J. M. Weinstein. 2008. "Information, school choice, and academic achievement: evidence from two experiments." *The Quarterly Journal of Economics* 123 (4): 1373-1414.

Heckman, J. J., and T. Kautz. 2012. "Hard evidence on soft skills." *Labour Economics* 19 (4): 451-464.

Heubert, J. P., and R. M. Hauser. 1998. *High Stakes: Testing for Tracking, Promotion, and Graduation*. Washington DC: National Academies Press.

Hidi, S., and J. M. Harackiewicz. 2000. "Motivating the academically unmotivated: A critical issue for the 21st century." *Review of Educational Research* 70 (2): 151-179.

Hoge, R. D., and T. Coladarci. 1989. "Teacher-based judgments of academic achievement: A review of literature." *Review of Educational Research* 59 (3): 297-313.

Jacob, B. A. 2005. "Accountability, incentives and behaviour: The impact of high-stakes testing in the Chicago Public Schools." *The Journal of Public Economics* 89: 761-796.

Jurajda, Š., and D. Münich. 2011. "Gender gap in admission performance under competitive pressure." *The American Economic Review* 101 (3): 514-518.

Jürges, H., K. Schneider, and F. Büchel. 2005. "The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany." *Journal of the European Economic Association* 3 (5): 1134-1155.

Jürges, H., K. Schneider, M. Senkbeil, and C. H. Carstensen. 2012. "Assessment drives learning: The effect of central exit exams on curricular knowledge and mathematical literacy." *Economics of Education Review* 31 (1): 56-65.

Jürges, H., and K. Schneider. 2010. "Central exit examinations increase performance... but take the fun out of mathematics." *Journal of Population Economics* 23 (2): 497-517.

Kinzie, M. B., M. A. B. Delcourt, and S. M. Powers. 1994. "Computer technologies: Attitudes and self-efficacy across undergraduate disciplines." *Research in Higher Education* 35 (6): 745-768.

Koenker, R. 2005. *Quantile regression*. New York: Cambridge University Press.

Lenroot, R. K., N. Gogtay, D. K. Greenstein, E. Molloy Wells, G. L. Wallace, L. S. Clasen, et al. 2007. "Sexual dimorphism of brain developmental trajectories during childhood and adolescence." *NeuroImage* 36: 1065-1073.

Louis, R. A., and J. M. Mistele. 2012. "The differences in scores and self-efficacy by student gender in mathematics and science." *International Journal of Science and Mathematics Education* 10 (5): 1163-1190.

Marra, R. M., K. A. Rodgers, D. Shen, and B. Bogue. 2009. "Women engineering students and self-efficacy: A multi-year, multi-institution study of women engineering student self-efficacy." *Journal of Engineering Education* 98 (1): 27-38.

Marsh, H. W. 1984. “Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility.” *Journal of Educational Psychology* 76 (5): 707-754.

Marsh, H. W. 1987. “The big-fish-little-pond effect on academic self-concept.” *Journal of Educational Psychology* 79 (3): 280-295.

Matějů, P., and M. Smith. 2014. “Are boys that bad? Gender gaps in measured skills, grades and aspirations in Czech elementary schools.” *British Journal of Sociology of Education* 36: 871-895.

Mechtenberg, L. 2009. “Cheap talk in the classroom: How biased grading at school explains gender differences in achievements, career choices and wages.” *Review of Economic Studies* 76: 1431–1459.

Ministry of Education Youth and Sports, Czech Republic (MŠMT). (n.d.) Retrieved August 22, 2013, from: http://www.msmt.cz/

Montgomery, H. 2005. “Gendered childhoods: A cross disciplinary overview.” *Gender and Education* 17 (5): 471-482.

Mullis, I. V. S., M. O. Martin, and P. Foy. 2008. *TIMSS 2007 International mathematics report*. Chestnut Hill, MA: Boston College.

Mullis, I. V. S., M. O. Martin, A. M. Kennedy, K. L. Trong, and M. Sainsbury. 2009. *PIRLS 2011 assessment framework.* Chestnut Hill, MA: Boston College.

Niederle, M., and L. Vesterlund. 2007. “Do women shy away from competition? Do men compete too much?” *The Quarterly Journal of Economics* 122 (3): 1067-1101.

Niederle, M., and L. Vesterlund. 2010. “Gender and competition.” *Annual Review of Economics* 3: 601-630.

Niederle, M., and L. Vesterlund. 2011. “Explaining the gender gap in math test scores: The role of competition.” *Journal of Economic Perspectives* 24 (2): 129-144.

Norman, M., and T. Hyland. 2003. “The role of confidence in lifelong learning.” *Educational Studies* 29 (2-3): 261-272.

OECD. 2013. *Education at a Glance 2013: OECD Indicators*. OECD Publishing.

Olson, J. F., M. O. Martin, and I. V. S. Mullis. 2008. *TIMSS 2007 Technical report*. Chestnut Hill, MA: Boston College.

Piopiunik, M. 2014. “The effects of early tracking on student performance: Evidence from a school reform in Bavaria.” *Economics of Education Review* 42: 12-33.

Pool, L. D., and P. Sewell. 2007. “The key to employability: Developing a practical model of graduate employability.” *Education + Training* 49 (4): 277-289.

Rask, K. 2010. “Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences” [Electronic version]. Retrieved 24.4.2015, from Cornell

University, School of Industrial and Labor Relations site: http://digitalcommons.ilr.cornell.edu/workingpapers/118/

Rask, K., and J. Tiefenthaler. 2008. “The role of grade sensitivity in explaining the gender imbalance in undergraduate economics.” *Economics of Education Review* 27: 676-687.

Rosenthal, R., and L. Jacobson. 1968. Pygmalion in the classroom. *The Urban Review* 3 (1): 16-20.

Rubin, D. B. 1987. *Multiple imputations for nonresponse in surveys*. New York: John Wiley & Sons, Inc.

Ryan, R. M., and E. L. Deci. 2000. “Intrinsic and extrinsic motivations: Classic definitions and new directions.” *Contemporary Educational Psychology* 25: 54-67.

Sabot, R., and J. Wakeman-Linn. 1991. “Grade inflation and course choice.” *Journal of Economic Persepctives* 5 (1): 159-170.

Seligman, M., and A. L. Duckworth. 2006. “Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores.” *Journal of Educational Psychology* 98 (1): 198-208.

Shavit, Y., R. Arum, and A. Gamoran. 2007. *Stratification in higher education: A comparative study*. Stanford: Stanford University Press.

Smith, R. H., and S. H. Kim. 2007. “Comprehending envy.” *Psychological Bulletin* 133 (1): 46-64.

Spencer, S. J., C. M. Steele, and D. M. Quinn. 1999. “Stereotype threat and women’s math performance.” *Journal of Experimental Social Psychology* 35 (1): 4–28.

Steinbeis, N., and T. Singer. 2013. “The effects of social comparison on social emotions and behavior during childhood: The ontogeny of envy and Schadenfreude predicts developmental changes in equity-related decisions.” *Journal of Experimental Child Psychology* 115 (1): 198-209.

Straková, J., and D. Greger. 2013. “Faktory ovlivňující přechod žáků 5. ročníku na osmileté gymnasium” [Transition of 5-graders to multi-year grammar school]. *Orbis Scholae* 7 (3): 73-85.

Südkamp, A., J. Kaiser, and J. Möller. 2012. “Accuracy of teachers' judgments of students' academic achievement: A meta-analysis.” *Journal of Educational Psychology* 104 (3): 743-762.

The Institute of Information and Prognoses of Education (ÚIPŠ). (n.d.) Retrieved August 23, 2013, from: http://www.uips.sk/

UIV. 2009. Genderová ročenka školství 2006/07 – 2009/10.

Usher, E.L., and F. Pajares. 2009. “Sources of self-efficacy in mathematics: A validation study.” *Contemporary Educational Psychology* 34 (1): 89-101.

Vansteenkiste, M., J. Simons, W. Lens, K. M. Sheldon, and E. L. Deci. 2004. “Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts.” *Journal of Personality and Social Psychology* 87 (2): 246-260.

Werfhorst, H. G. V., and J. J. B. Mijs. 2010. “Achievement inequality and the institutional structure of educational systems: A comparative perspective.” *Annual Review of Sociology* 36: 407-428.

Wilder, G. Z., and K. Powell. 1989. *Sex differences in test performance: A survey of the literature*. College Board Report No. 89(3). New York: College Entrance Examination Board.

Woessmann, L. 2002. “Central exams improve educational performance: International evidence.” *Kiel Discussion Papers* 397.

Woessmann, L. 2007. "Fundamental determinants of school efficiency and equity: German states as a microcosm for OECD countries." *IZA Discussion Paper* 2880.

Wooldridge, J. M. 2002. *Econometric analysis of cross section and panel data*. Cambridge: The MIT Press.

Zimmerman, B. 2000. “Self-efficacy: An essential motive to learn.” *Contemporary Educational Psychology* 25: 82–91.