GENDER GAP IN APPLICATION TO SELECTIVE SCHOOLS: ARE GRADES A GOOD SIGNAL?

Miroslava Federičová

CERGE EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

550

Working Paper Series 550
(ISSN 1211-3298)

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

Miroslava Federičová

CERGE-EI
Prague, September 2015

ISBN 978-80-7343-357-4 (Univerzita Karlova v Praze, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-352-8 (Národohospodářský ústav AV ČR, v. v. i.)

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

Miroslava Federičová[1]

Abstract

Grades are one of the most important factors in the transition between different levels of education. However, conditional on cognitive skills, grades differ substantially between girls and boys. This gender disparity in grade assignments according to cognitive skills may create asymmetrical signals of the probability of admission for girls and boys. This paper examines the role of grades in explaining the gender difference in application rates to selective schools. Using data about transition from primary to selective schools in the Czech Republic, the paper shows that girls apply at significantly higher rates. This difference remains the same after controlling for probability of admission. Test scores collected by an international testing program have no effect on gender differences in applications that are, however, 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.

JEL classification: I21, I24

Keywords: grading, school choice, admissions, gender gap

* This research was supported by a grant P402/12/G130 of the Grant Agency of the Czech Republic. The completion of this project was made possible through the helpful support of the professors at the University of Illinois in Chicago. I would like to especially thank Steven Rivkin, as well as Daniel Münich and Štěpán Jurajda for their helpful comments. All errors remaining in this text are the responsibility of the author.

[1] CERGE-EI is a joint workplace of Charles University in Prague and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic; e-mail: miroslava.federicova@cerge-ei.cz

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' achievements (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[2]. 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. Girls are usually more risk averse (Charness and Gneezy, 2012) and face higher anxiety during testing due to stereotype-threat (Spencer, Steele and Quinn, 1999). 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. 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. 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

[2] 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.

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.

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 pupils' achievements. 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 students' achievements 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's 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 achievements 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 achievements, 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 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.

3 Conceptual Framework

Transition between school levels usually consists of two consequential decisions, i.e. application and admission decision[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

[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.

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 5^{th}, 7^{th} and 9^{th} grade. In the 9^{th} grade, all students have to decide between an academic or vocational track. A different situation occurs at the 5^{th} and 7^{th} 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 9^{th} 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.

3.1 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:

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

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

3.3 The Signalling Model

Suppose that pupils receive a signal (s) about their admission scores and predict their admission score only according to this signal, i.e.

$$T^e = s.$$

Moreover, suppose that this signal is not a perfect proxy of the real admission score and maps the real admission score with an error (*ε*) such that the real admission score is defined as:

$$T = s + \varepsilon.$$

Using the signalling approach, the difference between the perceived and the real probability of admission is:

$$P(admission) - P(admission)^e = a\varepsilon.$$

In this signalling model, the accuracy of predicted admission probability is totally based on the error term from the signal and the weight placed on admission exam score in admission criteria. Hence, the lower the noise in the signal the lower the difference between the real and perceived probability of admission. Moreover, the higher the weight placed on admission exam score in admission criteria, the higher the difference in probabilities. If the signal is a perfect proxy for admission score, i.e. the error in the signal is equal to zero, the perceived probability of admission unambiguously corresponds to the real probability of admission. In this case, pupils make their educational choice in accordance with their real admission chances. Otherwise, the noisy signal leads to imprecise prediction of admission chances, and consequently, to ineffective allocation of talents into educational tracks. Moreover, if this noisy signal differs considerably for specific groups of pupils (i.e. according to the socio-economic status of the family, gender, age, minority group, etc.), it can lead to biased beliefs about admission probability and thus to inequality in educational opportunities. 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.

3.4 Gender disparity

Assume that, according to the observed signal of abilities, girls and boys form their predicted admission scores equally[5]. However, girls and boys may differ in bias produced by this signal according to the real admission score. In other words, the bias can be decomposed to 2 parts: the noise (η) that is common for girls and boys, and the second

[5] Here, I assume that pupils and their parents do not observe the gender bias in the signal and thus they do not adjust their beliefs according to this bias. In other words, the observed signal leads to the same predicted probability of admission for girls and boys. This assumption leans on the fact that pupils make this decision for the first time and hence do not learn about the bias from previous attempts. Moreover, the lack of other possible signals of admission scores make this gender adjustment more difficult.

part (*g*), which is gender specific. Using the gender decomposition of bias, the real test score is now defined as:

$$T^e = s$$

$$T = s + \eta + g, \qquad where\ g = \begin{cases} x_g, & for\ girls \\ x_b, & for\ boys \end{cases} \ and\ x_g, x_b \in \mathbb{R}$$

and the gender difference in perceived probability of admission formed by the predicted admission scores is then:

$$P(admission)_g^e - P(admission)_b^e = a(x_b - x_g).$$

Thus, the higher the difference in gender specific bias between boys and girls the higher the gender difference in perceived probability of admission. In other words, even if girls and boys have the same real admission chances, those with a less accurate signal form a lower perceived probability of admission than others, and hence, apply less often.

At the early stage of school, grades are usually the only signal of abilities that pupils observe. Moreover, several studies[6] show that girls' grades usually overestimate their true cognitive skills, whereas boys' grades are undervalued according to their cognitive skills[7]. Thus, using the assumption that the signal of grades overvalued the true admission scores of girls (i.e. $\eta + x_g < 0$) and undervalued those of boys (i.e. $\eta + x_b > 0$) the difference between the real and predicted admission probability can be expressed separately for girls and boys as: :

$$P(admission)_g - P(admission)_g^e = a(\eta + x_g) < 0$$

$$P(admission)_b - P(admission)_b^e = a(\eta + x_b) > 0$$

where subscripts *g* and *b* refers to girls and boys, respectively. Hence, if admission exams to selective schools are focused mainly on cognitive skills, using grades as a signal for

[6] See e.g. Falch and Naper (2013), Matějů and Smith (2014), and Seligman and Duckworth (2006).
[7] 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). The model presented in the Appendix uses the decomposition of grades to cognitive and noncognitive skills to approach the possible sources of gender difference in grades and its transition on gender difference in application to selective schools.

admission scores, girls overestimate their true probability of admission whereas boys underestimate that probability.

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 noisier the signal of grades is perceived to be, the smaller the gender difference in application given the actual skill distribution. Hence, the gender difference in application to selective schools is affected not only by the gender difference in accuracy of grades as a signal of admission scores, but 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.

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[8] 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 the by international testing programmes TIMSS and PIRLS[9] 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[10]

[8] 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.

[9] 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.

[10] 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.

participate in the subsequent round in the 5th grade. Hence, from 4578 students tested in TIMSS 2011, 3681 were again questioning the following year.

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 accessible in Table 1. 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[11].

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[12] 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

[11] 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.
[12] Math and reading test scores are normalized to mean 0 and standard deviation 1.

schooling. Descriptive statistics for the two assessment measures are presented in Table 1.

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. 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 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. 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.

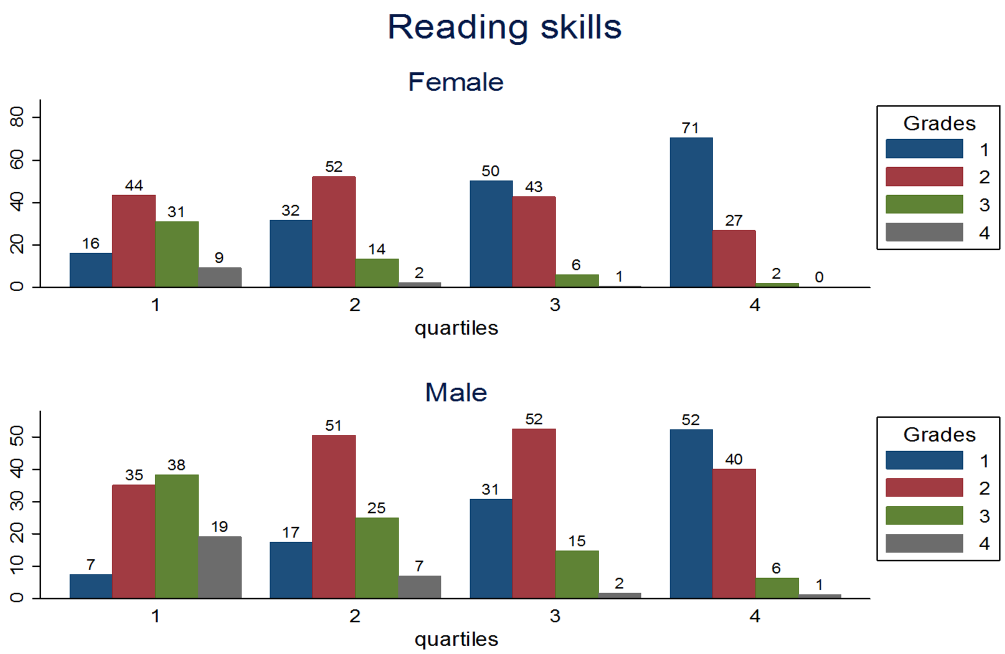
Table 2: Representation of grades and gender difference in test scores for particular grades

	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[13]	1%	1%	0.23	(0.38)	1%	0%	-0.26	(0.55)

Source: Own calculations based on data from CLoSE.
Note: Standard errors are in parenthesis.

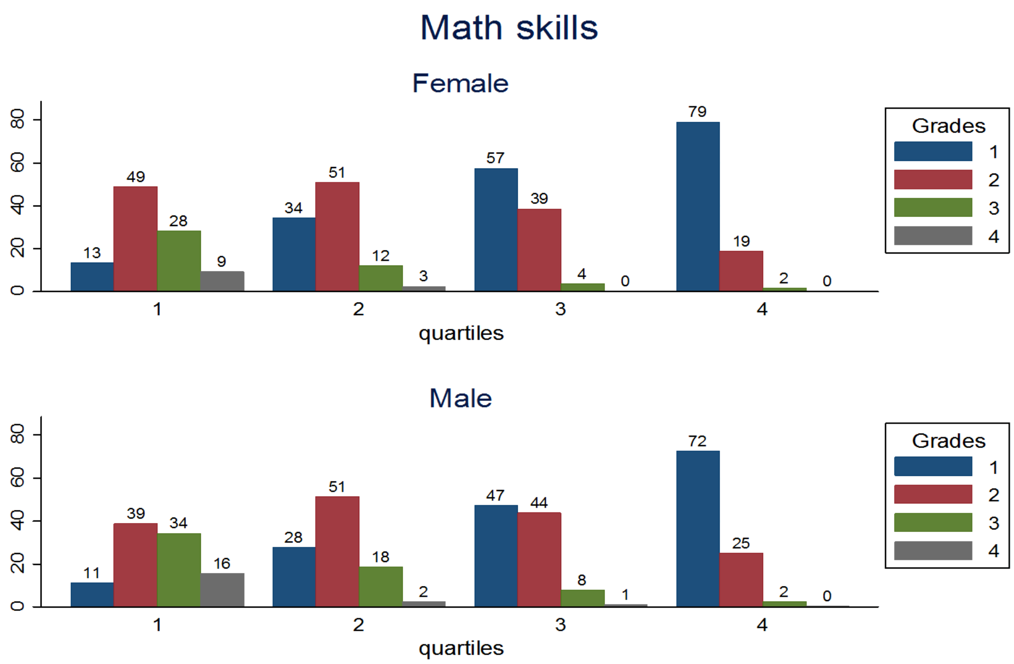
[13] The insignificant difference in gender gap in test scores for grades 4 and 5, i.e. the worst grades, is probably caused by the small number of pupils within these grades.

Figure 1: Distribution of grades in Czech according to quartiles of reading test scores, by gender



Source: Own calculations based on CLoSE.

Figure 2: Distribution of grades in Math according to quartiles of math test scores, by gender



Source: Own calculations based on CLoSE.

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 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[14]. 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 and Figure 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[15], 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 these are specific for gender[16] (Matějů and Smith, 2014; Seligman and Duckworth, 2006). The different amount of cognitive skills needed for particular grades

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

[15] 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).

[16] 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.

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 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.

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.

To examine the application decision conditional on the probability of admission the analysis is firstly focused on 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 are 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:

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

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 4th grade. Science and math test scores are highly correlated and probably both exposed to the same measurement error. 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[17]. 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:

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

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, parents education, etc. 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, I further estimate the extent

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

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[18]. 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.

6 Results

Firstly, 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.

[18] 99% of pupils who applied to selective schools got grade 1 or 2 from mathematics and 98% from Czech.

6.1 Application Decision

The analysis is firstly focused on gender difference in application rate and the ability measures that affect it. Table 3 reflects successively the impact of grades and test scores on the decision to apply. All specifications presented in separate Columns of Table 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[19].

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.

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

[19] 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.

the pupils who applied and further predicted for the whole sample. Table 4 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.

The distribution of predicted probability of admission is depicted separately for girls and boys in Figure 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[20]. The gender difference in admission probability is also not confirmed within admission deciles (Figure 4). The dotted line in Figure 4 shows that girls and boys are almost evenly distributed in admission deciles. These results indicate no gender difference in the probability of admission. Thus, the lower application rate of boys is not a response to their lower admission chances as is suggested in the first possible explanation of the gender gap in application rate.

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 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.

[20] The gender difference in probability of admission on average is also not confirmed for pupils who decided to apply (the graph in Figure 3, on the right).

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[21] of the admission probability in Figure 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.

The gender difference in application conditional on admission observed in Figure 5 and 6 also persists on average (see Table 5). 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. However, the estimates of interest are now the interactions between admission deciles and the gender variable. Controlling only for individual characteristics in Column 1, 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 2 and Column 3, respectively. Conditional on grades, the gender difference in the 9th decile drops and becomes insignificant, while conditional on test scores it even rises to 10%. 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.

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

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 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[22]. 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. 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.

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 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

[22] 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.

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 7). 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, 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 causes the gender gap in application decision.

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 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.

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 9 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[23] of weights placed on grades in each district, I find no difference in the influence of grades on the application decision (Table 8). In other words, pupils in districts with very low weight on grade in the admission 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[24].

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

[23] 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.

[24] 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.

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[25]. The results are presented in Table 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' 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[26].

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

[25] 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 10 depicts the variation in Scio exams across Czech districts.

[26] 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.

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 ineffective allocation of pupils to school tracks if grades are the only signal about pupils' academic performance that pupils observe. 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 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, in this paper the gender difference in application rate is explained 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.

References

Bennett, R. E., Gottesman, R. L., Rock, D. A. and Cerullo, F. (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. and Morell, D. (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.

Charness, G. and Gneezy, U. (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.

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

Feinberg, A. B. and Shapiro, E. S. (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.

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

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

Jurajda, S. and Münich, D. (2011). Gender gap in performance under competitive pressure: Admissions to Czech universities. *American Economic Review, 101*(3), 514-518.

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 Smith, M. (2014). Are boys that bad? Gender gaps in measured skills, grades and aspirations in Czech elementary schools. *British Journal of Sociology of Education,*2014.

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.

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

Niederle, M. and Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Prespectives, 24*(2), 129-144.

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 Tiefenthaler, J. (2008). The role of grade sensitivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review*, 27, 676-687.

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

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

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

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

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

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

Appendix

In the conceptual framework, it is shown how grades and admission exam scores enter into both the application and admission decision. The following model decomposes grades and admission scores into cognitive and noncognitive skills. This simplified approach allows the description of the possible streams through which relying on grades in the application decision may cause gender differences in application even after controlling for admission. Firstly, the general model of cognitive and noncognitive skills is presented, followed by the implementation of gender disparities in these two skills.

Cognitive vs. Noncognitive Skills

Cognitive and noncognitive skills enter into both the grades and admission scores. Generally speaking, admission scores and grades can be viewed as a weighted average of cognitive and noncognitive skills, with the restriction that the weight on cognitive skills is higher for admission scores than for grades. Using this simplification with the assumption that cognitive (C) and noncognitive skills (NC) are equally distributed, the exam scores (T) and grades (G) are formulated as:

$$T = \alpha_1 C + \alpha_2 NC \text{ , where } \alpha_1 + \alpha_2 = 1 \quad (1)$$

$$G = \beta_1 C + \beta_2 NC \text{ , where } \beta_1 + \beta_2 = 1 \quad (2)$$

where α_1, resp. β_1 are the weights for cognitive and α_2, resp. β_2 for noncognitive skills in admission exam scores resp. grades. If the weights for cognitive and noncognitive skills are equally distributed in admission scores and grades, these yield an identical measure of pupils' achievement. This makes grades a perfect proxy for admission scores.

Suppose now that pupils observe no other information about their skills except grades, and hence, they use only grades to form their expectations about admission scores:

$$T^e = i(G) \quad (3)$$

In this case, grades enter into the perceived probability of admission not only directly through known admission criteria but also indirectly through their role in expected admission scores. Substituting grades in equation (3) with equation (2) and using the assumption that cognitive and noncognitive skills totally explain both grades and admission scores yields:

$$T^e = \big(\alpha_1 + (\alpha_2 - \beta_2)\big)C + \big(\alpha_2 + (\alpha_1 - \beta_1)\big)NC$$

And thus,

$$T^e \ = \ T \ + \ (\alpha_2 - \beta_2)C + (\alpha_1 - \beta_1)NC \tag{4}$$

Again, if cognitive and noncognitive skills are represented in grades and admission scores equally, i.e. $\alpha_1 - \beta_1 = 0$ and $\alpha_2 - \beta_2 = 0$, the expected admission score is equal to the real admission score. Recall that in this case, grades serve as a perfect proxy for expected admission score, and hence, for the expected admission probability. If however the admission exam puts higher weight on cognitive skills than grades, i.e. $\alpha_1 - \beta_1 > 0$ and $\alpha_2 - \beta_2 < 0$, noncognitive skills are overrepresented and cognitive skills underrepresented in pupils' expectations of their admission exam score. This is caused by the imperfect proxy of grades. In other words, grades as a proxy for admission exam score underweight cognitive skills, and as a consequence, pupils with lower grades but higher cognitive skills are likely to underapply.

If pupils know the weights on cognitive and noncognitive skills in grades and admission scores, i.e. coefficients α and β are known, relying on grades would not cause any difference between the expected and the real admission scores. Pupils can assume that admission exams are largely aimed at cognitive skills, and thus, equal α_1 to one and α_2 to zero. But, the composition of cognitive and noncognitive skills in grades is not easy to construct. Teachers use diverse rules to assign grades as no national framework or guidelines for grading are set. This enables the prediction of weights on cognitive and noncognitive skills in grades. Moreover, during primary education, pupils and parents usually lack some additional information on pupils' study aptitudes, e.g. from national testing, that can be compared with assigned grades.

Gender Disparity

The equation (4) shows that the real and expected admission exam scores differ if pupils use an imperfect proxy for the real admission exam scores. The equation (4) thus yields:

$$T^e = T + \varepsilon$$

, where the error term ε caused by an imperfect proxy for the admission score is equal to $(\alpha_2 - \beta_2)C + (\alpha_1 - \beta_1)NC$. Hence, the gender disparity in application conditional on probability of admission occurs if error term ε is specific for gender, i.e. $\varepsilon \perp gender$ does not hold. This is the case if girls and boys have different distribution of cognitive and nonocgnitive skills. As admission exams to academically selective schools test primarily cognitive skills, the gender gap in the application decision conditional on probability of admission is mainly triggered by the gender difference in noncognitive skills distribution.

The gender disparity in the application decision conditional on admission is composed of two factors: the gender difference in distribution of noncognitive skills and the size of the error term. The first is based on the fact that girls outperform boys in noncognitive skills, i.e. their noncognitive skill distribution lies to the right of that of boys (Seligman and Duckworth, 2006). For the error term it holds that the higher the difference in representation of noncognitive skills in the real and the expected admission score, the higher the error term. Thus, even with no gender difference in cognitive skills – which are primarily tested in admission exams – the distribution of perceived probability of being admitted is shown to be higher for girls.

Tables

Table 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: 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 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

Standard errors in parentheses clustered on class level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 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
[1] Instrument for reading and math test score
* p < 0.1, ** p < 0.05, *** p < 0.01

Table 5: Application decision to selective school conditional on the deciles of predicted probability of being admitted, LPM

Application	(1)	(2)	(3)
Boy	-0.001	-0.001	0.027***
	(0.011)	(0.012)	(0.013)
Predict. admission=3	0.013	0.009	0.044**
	(0.015)	(0.015)	(0.018)
Predict. admission=4	0.107***	0.084***	0.146***
	(0.029)	(0.029)	(0.032)
Predict. admission=5	0.113***	0.074***	0.166***
	(0.028)	(0.028)	(0.037)
Predict. admission=6	0.218***	0.142***	0.274***
	(0.036)	(0.038)	(0.043)
Predict. admission=7	0.249***	0.154***	0.312***
	(0.034)	(0.037)	(0.047)
Predict. admission=8	0.339***	0.224***	0.413***
	(0.040)	(0.045)	(0.054)
Predict. admission=9	0.467***	0.324***	0.548***
	(0.038)	(0.047)	(0.053)
Predict. admission=10	0.564***	0.410***	0.681***
	(0.046)	(0.050)	(0.073)
Boy * Predict. ad.=3	0.004	0.008	-0.008
	(0.020)	(0.021)	(0.021)
Boy * Predict. ad.=4	-0.021	-0.006	-0.031
	(0.038)	(0.037)	(0.038)
Boy * Predict. ad.=5	-0.055*	-0.036	-0.073**
	(0.031)	(0.031)	(0.033)
Boy * Predict. ad.=6	-0.044	-0.012	-0.057
	(0.046)	(0.046)	(0.046)
Boy * Predict. ad.=7	0.013	0.041	0.003
	(0.051)	(0.051)	(0.051)
Boy * Predict. ad.=8	-0.076	-0.043	-0.085
	(0.053)	(0.053)	(0.054)

Boy * Predict. ad.=9	-0.090*	-0.060	-0.099*
	(0.050)	(0.051)	(0.051)
Boy * Predict. ad.=10	-0.013	-0.001	-0.033
	(0.051)	(0.051)	(0.051)
Grades	No	Yes	No
Scores	No	No	Yes
N	3017	3005	3017
R^2	0.178	0.192	0.187

Standard errors in parentheses clustered on class level.
* p < 0.1, ** p < 0.05, *** p < 0.01

Table 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 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

Table 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 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 3: Distribution of predicted probability of being admitted, by gender and application

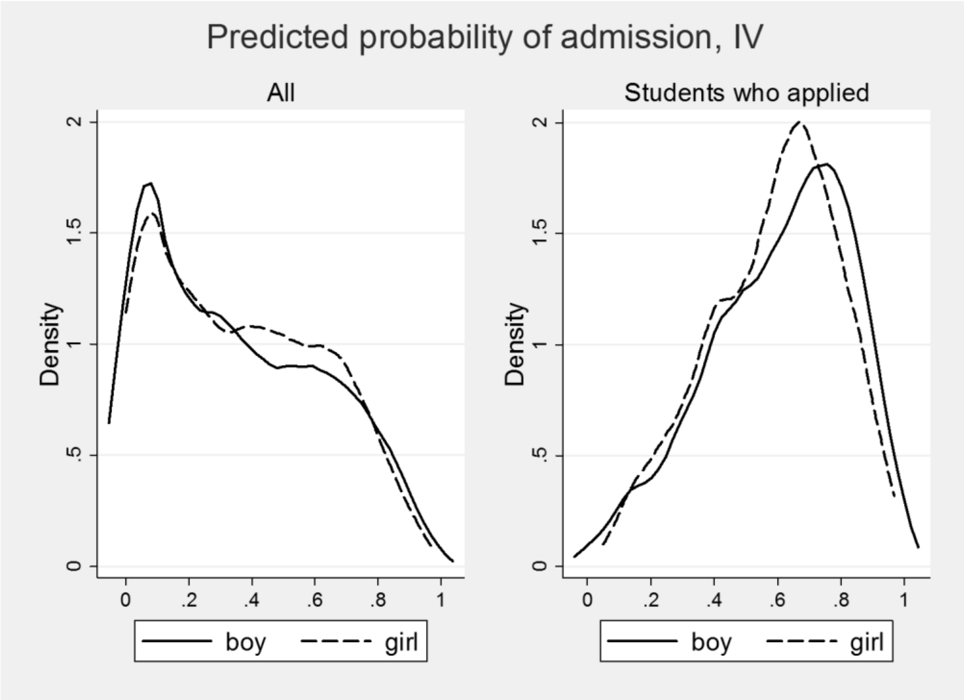
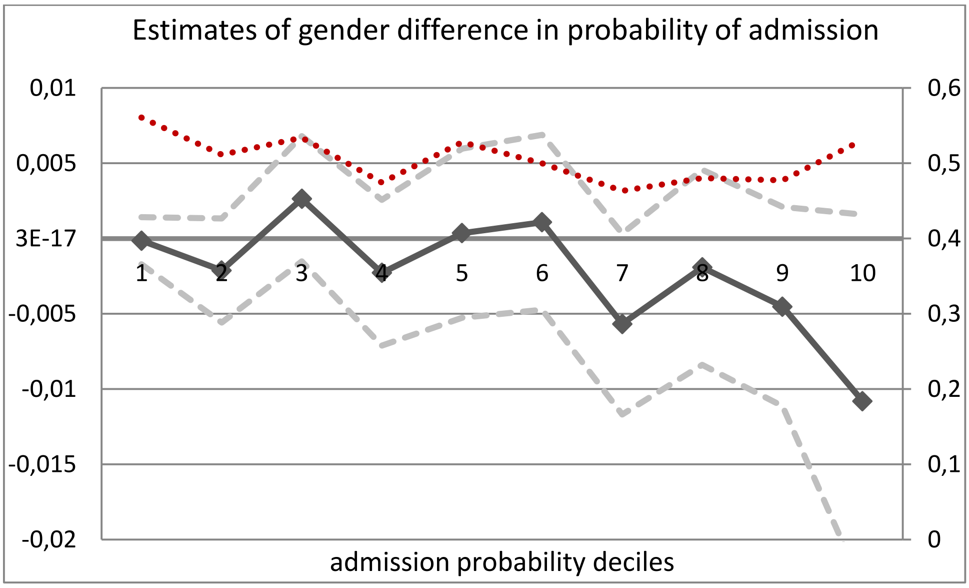


Figure 4: The estimates and 95% confidence intervals of gender difference in probability of admission, by probability admission deciles (all pupils)



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.

Figure 5: Application rates in admission deciles, by gender

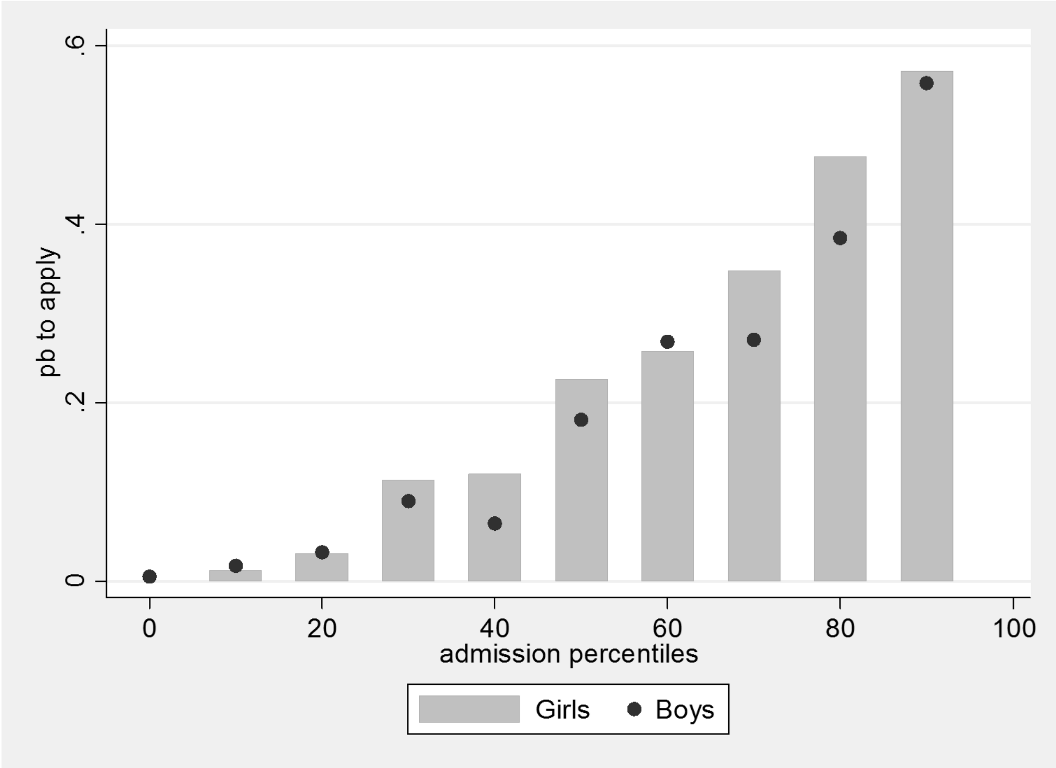


Figure 6: Application rates in admission percentiles, by gender

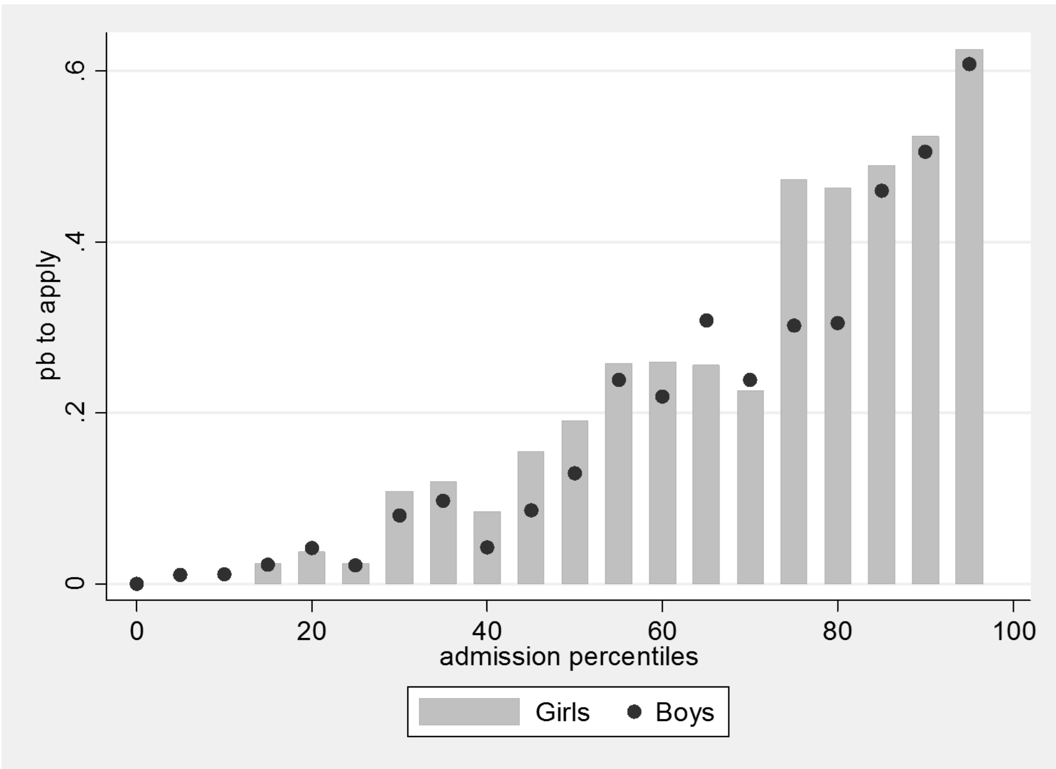
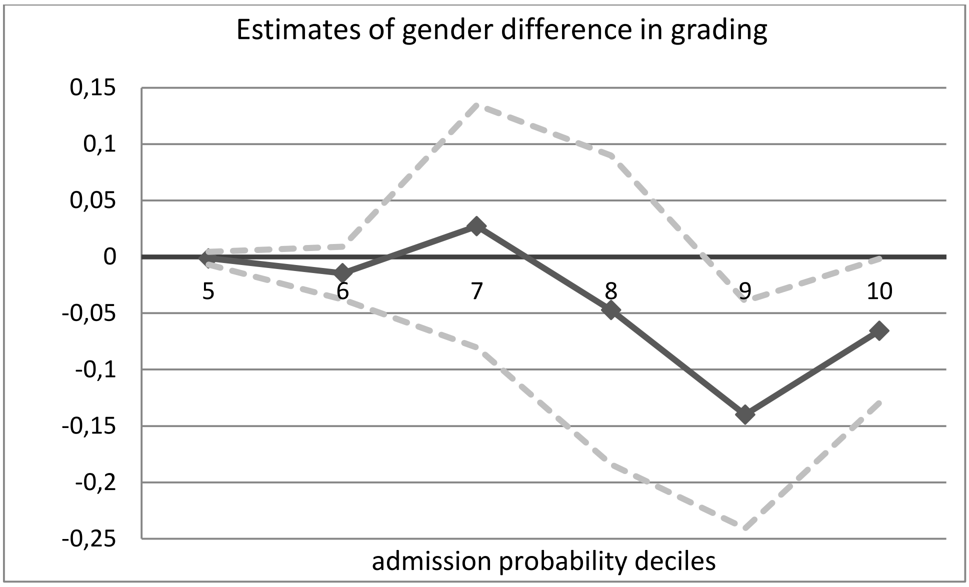


Figure 7: Estimates and 95% confidence interval of gender difference in grading conditional on test scores, by admission deciles



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

Figure 8: Application rates conditional on probability of admission, by GPA

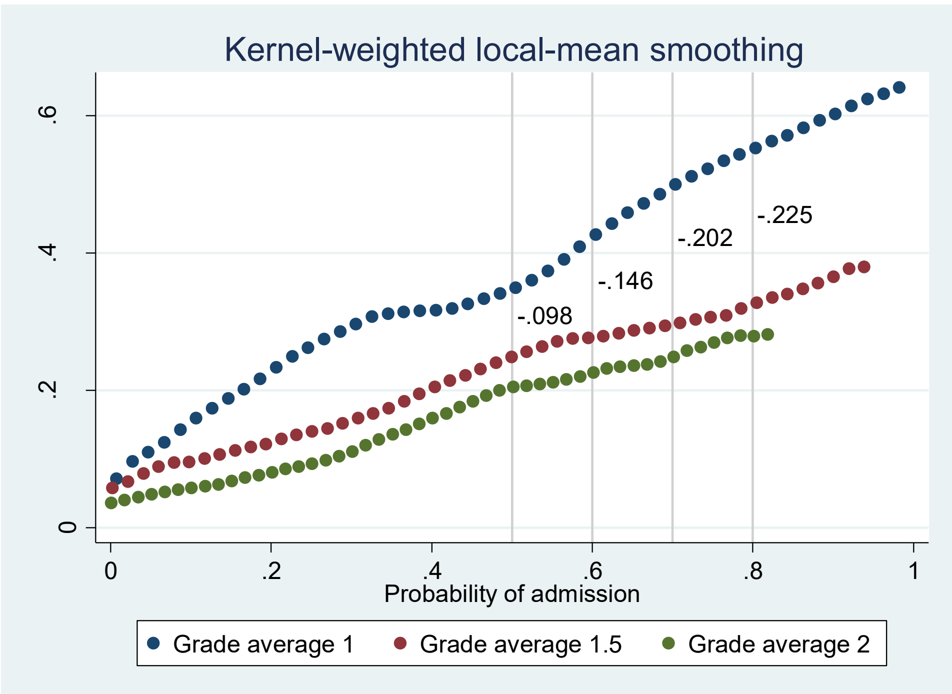
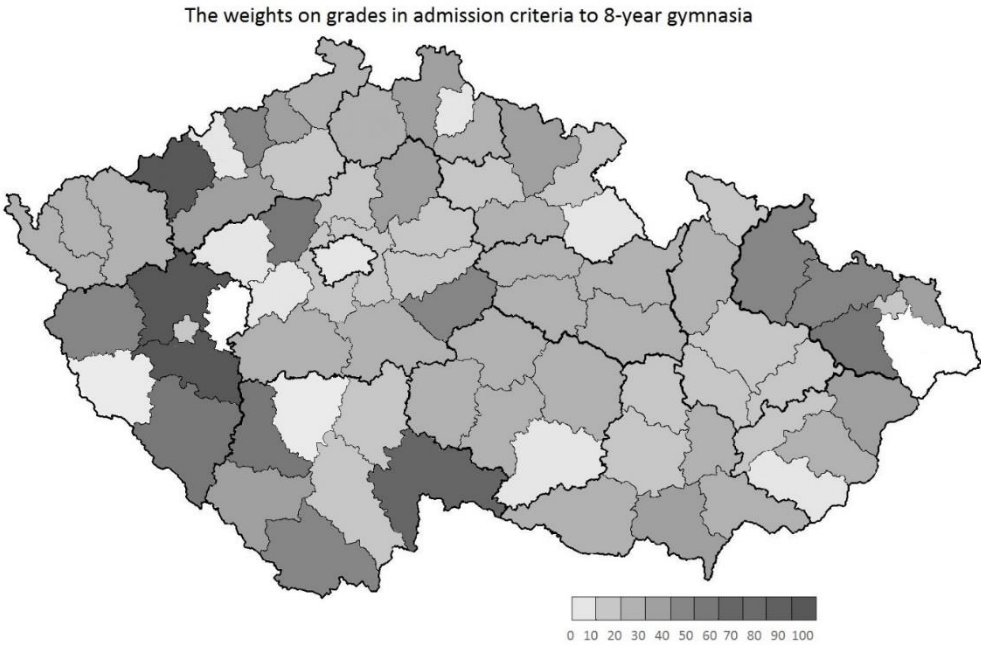
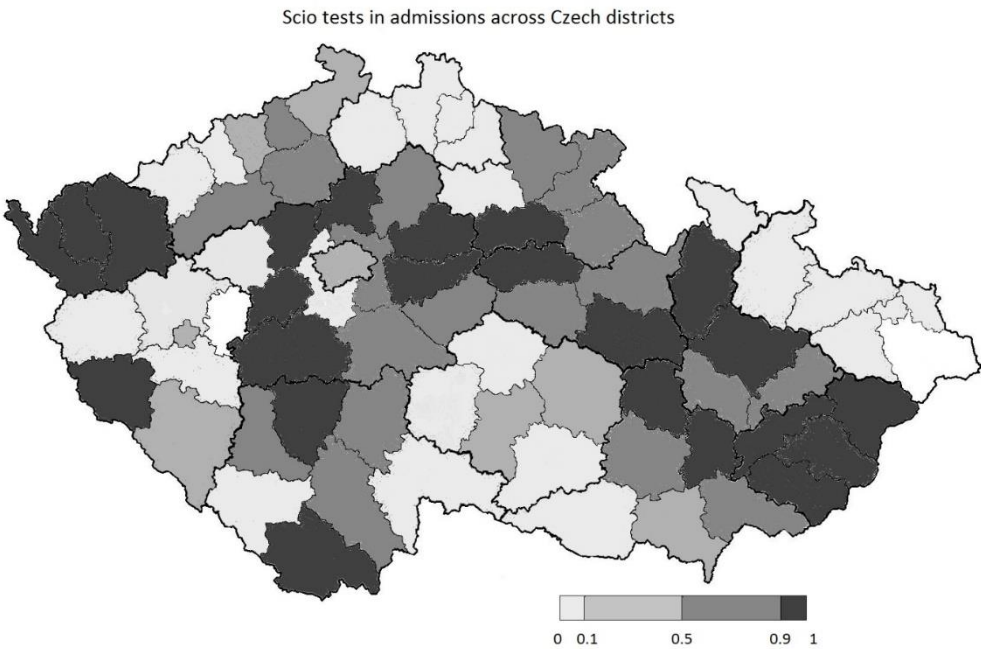


Figure 9: The weight placed on grades in admission criteria to 8-year gymnasium, by Czech districts



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

Figure 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.

Abstrakt

Známky jsou jedním z nejdůležitějších faktorů pro přechod mezi jednotlivými stupni vzdělání. Nicméně známky se značně liší mezi děvčaty a chlapci, a to i pokud mají stejné kognitivní dovednosti. Tato genderová nerovnost ve známkování vzhledem ke kognitivním schopnostem může pro děvčata a chlapce vytvářet asymetrický signál o pravděpodobnosti přijetí na selektivní školu. Tento článek zkoumá vliv známek na genderový rozdíl v míře přihlášek na selektivní školy. Použitím údajů z přechodu mezi primárním stupněm vzdělání a víceletým gymnáziem v České Republice článek ukazuje, že děvčata se hlásí ve významně větší míře. Tento rozdíl se nemění ani po zohlednění šancí jednotlivých žáků na přijetí. Výsledky z mezinárodního šetření nemají žádný vliv na genderový rozdíl v přihlašování se, který je nicméně 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.

Working Paper Series
ISSN 1211-3298
Registration No. (Ministry of Culture): E 19443

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the ASCR, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: http://www.cerge-ei.cz

Editor: Marek Kapička

The paper is available online at http://www.cerge-ei.cz/publications/working_papers/.

ISBN 978-80-7343-357-4 (Univerzita Karlova v Praze, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-352-8 (Národohospodářský ústav AV ČR, v. v. i.)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz