

IZA DP No. 8018

**Gender Differences in Strategic Behaviour under
Competitive Pressure: Evidence on Omission Patterns
in University Entrance Examinations**

Tuomas Pekkarinen

March 2014

Gender Differences in Strategic Behaviour under Competitive Pressure: Evidence on Omission Patterns in University Entrance Examinations

Tuomas Pekkarinen

*VATT, Helsinki
and IZA*

Discussion Paper No. 8018
March 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Gender Differences in Strategic Behaviour under Competitive Pressure: Evidence on Omission Patterns in University Entrance Examinations^{*}

This paper studies gender differences in performance in university entrance examinations. We exploit data from the exams that the nine Finnish universities providing education in economics and business use to choose their students. These exams are multiple choice tests where wrong answers are penalized by minus points and omissions yield zero points. This scoring rule means that the number of omitted items will affect the probability of entry. The strategic setting of the applicants varies depending on the university where she is applying to and on the amount of starting points that she is rewarded based on her high school success. The results show that, controlling for starting points, women perform worse than men in the entrance exam and are less likely to gain entry. Women also omit more items in the exam. Using the Rasch Model to derive the predicted probabilities of answering items correctly for each applicant, we show that women deviate more from the number of answered items that would maximise the predicted probability of entry than men and that they do so because they answer to too few items.

JEL Classification: J16, I21, I23

Keywords: gender differences, competition, multiple choice exams

Corresponding author:

Tuomas Pekkarinen
Government Institute for Economic Research (VATT)
Arkadiankatu 7
P.O.Box 1279
00101 Helsinki
Finland
E-mail: tuomas.pekkarinen@vatt.fi

^{*} The author would like to thank Kristian Koerselman, Daniele Paserman, and Roope Uusitalo for useful comments on the earlier drafts of this paper. Also the comments by audiences at various seminars and conferences are acknowledged as well as the financial support by the Academy of Finland. For much of the time when this manuscript was written the author was employed at the Aalto University School of Business in Helsinki, Finland.

1 Introduction

Female educational attainment in industrialised countries has increased to such an extent over the recent decades that it is now a widely-accepted stylised fact that women are on average more educated than men. Yet, this improvement in female educational attainment has not translated into reduced gender gaps in labour market outcomes. Women are still vastly underrepresented at the upper end of professional hierarchies. This contradictory pattern of decreasing gender gaps in skills and persistent gender gaps in labour market outcomes has increased the popularity of psychological explanations. In particular, gender differences in performance under competitive pressure have attracted a lot of attention in recent years.

The experimental literature has rather convincingly documented that women perform worse than men under competitive pressure in lab experiments. However, how this gender gap in performance emerges is still an open question. One possibility suggested by tournament theory is that women and men adopt different strategies in competitive settings. In a tight competition, risk taking associated with overconfidence may increase the probability of winning. If women adopt more conservative strategies than men in highly competitive settings, we would expect to see lower variation of test scores among women and over-representation of men at the tails.

In this paper, we examine gender differences in performance and choice of strategy in a real-life high stakes setting: the university entrance examinations. More specifically, we exploit data from the joint entrance examinations of all the Finnish universities that provide education in business administration and economics. These exams are multiple choice tests where wrong answers are penalized with negative points whereas omissions yield zero points. This means that guessing strategies matter in these tests. Using data on item-by-item answers by applicants, we test whether men and women choose different strategies in competitive settings and whether these strategies result in gender differences in performance measured by the probability of gaining entry to a university.

According to experimental research there are consistent gender differences in attitudes towards risk and competition and in performance in competitive environments. Gneezy, Niederle, and Rustichini (2003) showed that women perform worse than men in competitive environments and, consistent with this result, the study by Niederle and Vesterlund (2007) provided evidence that women also seek to avoid competitive situations. There is no consensus on why these gender differences emerge. Niederle and Vesterlund (2007) attribute the reluctance of women to take part in competitive settings to lower female expectations about performance. On the other hand, the results in Dohmen and Falk (2011) suggest that gender differences in the attitudes towards risk could explain a large part of the differences.

Results from field data come from very specific settings such as sports competitions where the importance of competitive pressure is obvious. However, there is disagreement how important gender differences in performance under competitive pressure are for labour market outcomes. For example, Manning and Saidi (2010) argue that at least in the United Kingdom gender differences in the likelihood of working under pay for performance contracts are so small that it would be very difficult to reconcile them with important gender differences in the attitudes towards competition. Yet, there are situations in the labor market where we would expect attitudes towards competition to matter. Typical examples include promotions that are usually based on the relative performance of workers with respect to their peers. Perhaps even more clearly attitudes towards competition should matter in education where evaluation is in most cases based on relative

performance.¹ Typical examples of competitive settings with high stakes in education are university entrance examinations that are used in many European countries to ration higher education. Indeed, there is evidence that women perform worse than men in these kinds of exams. Both Ors, Palomino, and Peyrache (2013) as well as Jurajda and Munich (2011) show that men and women perform differently in university entrance examinations and that gender difference increase in favor of men when the competitiveness of the examination increases.

Tournament theory suggests that risk taking may increase chances of winning in tight competitions. For example, Bronars (1987) and Hvide (2002) show that when competitors are otherwise identical increasing the spread of output increases the probability of winning. Goel and Thakor (2008) have combined the tournament setting with overconfidence. They show that if overconfidence causes some participants to underestimate the variance of output, they will choose actions that result in a higher spread of output and are hence more likely to win. These theoretical results suggest that if there are gender differences in the choice of strategy, they may result in the kind of performance differences observed in the experimental literature.

In this paper, we contribute to this literature by examining the link between gender differences in strategic choices and performance in university entrance examinations. We have access to the full answer sheets of all the applicants that took part in the joint entrance examination of the Finnish universities that provide higher education in business administration and economics. The applicants apply to a specific university but they all take the same entrance examination. The entrance examination is a multiple choice test where wrong answers are penalized with negative points. The entry into the university is based on test performance and starting points that the applicants get from their high school matriculation exam results. The universities differ in the entry thresholds due to differences in the number of study slots as well as in demand for these slots. Hence, there is variation in the competitive setting across individuals with different starting points and across universities with different entry thresholds.

Our aim is to analyze whether the gender differences in performance under competitive pressure arise because women resort to more conservative strategies. We believe that the data from the university entrance examinations provide a promising setting to study this question. The objective of each applicant is very clear: to maximise the probability of gaining entry. However, as negative answers are penalized in the entrance examination the guessing strategy that maximises the probability of getting accepted depends crucially on the amount of points that the applicants need to gain in the exam. Applicants who answer items correctly with identical probabilities and hence have the same expected score from the exam, should omit different number of items in the exam if they are facing different thresholds of entry. The further away from the threshold the applicant is, the less items he or she should omit. By exploiting the information on the full answer sheets of the applicants as well as their starting points and the average threshold of the university that they are applying to, we can model the strategic setting of each applicant in a detailed way. In what follows, we use the Item Response Theory to derive the predicted probabilities of answering items correctly for each applicants. With these predicted probabilities we calculate the number of items that would maximise the probability of entry for each applicant. With this information, we study whether women deviate more from the predicted answering strategy than men by omitting more or less items in the exam.

Our results show that while there are no large gender differences in performance in

¹For gender differences in performance in typical exam settings see Nekby, Thoursie, and Vahtrik (2008).

these entrance examinations on average, once we control for the higher starting points of women clear differences emerge. Women clearly score lower number of points in the entrance exam and are less likely to gain entry to university than men with similar starting points. Women tend to lose the advantage that they have based on starting points in the entrance examination. Examination of the answering patterns reveals that women omit more items in the exam than men. Moreover, the analysis based on the predicted probabilities that are derived from the Rasch Model of item by item answers, reveals that women deviate more from the optimal number of answered items than men do and that they do so because they answer too few items.

The rest of the paper proceeds as follows. In the second section, we describe the joint entrance examinations for business studies at Finnish universities and the institutional setting in more detail. In the third section, we go over the theoretical setting and derive the optimal answering strategy for these kinds of entrance examinations and show how the strategy varies as a function of the threshold and the probability of answering items correctly. This section relies heavily on recent work by Espinosa and Gardeazabal (2010). The fourth section presents the data and in the fifth section we discuss the results on gender differences in performance in the entrance examinations as well in answering strategies. The sixth section concludes.

2 Institutional setting

University education in Finland is free in the sense that no tuition is charged. However, the amount of study slots is limited and the most popular universities are heavily over-subscribed. The universities typically choose their students based on some combination of high school matriculation examination and university entrance examination results. In the field of economics and business administration, the nine universities that provide education in these fields have co-ordinated the allocation of study slots so that all the universities take part in the same entrance examinations and give starting points from matriculation exam grades based on same rules.

2.1 Starting points

The applicants are given starting points based on their high school matriculation examination results. This is a national examination that the students take at the end of their last year in secondary school. Students are required to take the matriculation examination in at least four subjects and the examination is graded centrally and grades are standardized so that grade is determined by the students position at the national score distribution. Approximately 5-6 % of the participants fail the matriculation exam every year. The matriculation grades vary from 7 (known as *Laudatur* grade) which is assigned to top 5 % of the scores to 1 (*Approbatum*) which is assigned to the bottom 5% of the students who passed.

The admission system credits starting points based on the grading scheme in table 1. The applicant gets points from his or her Finnish exam and either from long or short option of the mathematics exam. In addition the applicant can choose two subjects (long or short options) from the matriculation exam that will be credited in the starting points. These subjects are either humanities or science subjects or additional languages. In total, the maximum amount of points that the applicant can gain from the matriculation exam is 40 points.

In addition to the matriculation exam points, the applicants are awarded starting

points if they have graduated from the secondary school during the same year (2 points for the applicants first choice and 1 for the applicants second) and 2 additional points for the first ranked option and 1 additional point for the second ranked option. This means that in total the maximum number of starting points is 44.

2.2 Entrance examination

Entrance examination is a multiple choice test where applicants are asked to answer questions in economics and business administration related topics. To prepare for the entrance examinations the applicants are advised to study five books in following subjects: organization and management, marketing, accounting, economics, and quantitative methods. These books cover material that is not part of the standard high school curriculum.

In the entrance examination the applicant has to give answers to standard multiple choice questions where one of four alternatives is correct. Correct answers give +1 points and false answers -0.5 points. Applicants are free to give blank answers which are not penalized but in order to be admitted to any of the universities, the applicants need to score the minimum of ten points in the entrance examination. The maximum number of points from the entrance examination is 40.

2.3 The application and admission procedures

The applicants take part in the entrance examination first by filling a centralised application form that is common to all universities that take part in the joint entrance examination. In the application form the applicants can apply up to three universities and they must rank these universities in preference order. Some of the universities admit students directly into specific major subjects and in the case of these universities the applicants also have to rank up to three subjects within the university. The applicants do not fill in their matriculation exam grades themselves since the universities receive them directly from the grade registers. After handing in the application form the students take part in the entrance examinations that are organised at the same time in all 9 university locations.

Each university has a fixed number study slots and they are filled so that 60% of the slots are filled based on the ranking of sum of entrance examination and starting points and 40% based on the ranking of only the entrance examination points. The admission procedure starts by checking whether the joint starting and entrance examination points of the applicant are high enough to place her in the 60% that are admitted to her first choice university based on the joint sum of points. If this is not the case, the system checks whether the entrance examination points of the applicant are high enough to place her in the 40% that are admitted to her first choice university based on the entrance examination points alone. If the applicant is not admitted to her first choice university, the same procedure is repeated for the second and possibly for the third choice university. The applicants whose points are not high enough to any of the universities that she has applied to will not get a study slot. If the university that the applicant has applied to admits students directly to major subjects, the system also checks whether the joint or the entrance examination points of the applicant are high enough to get her to one the subjects that the applicant has ranked in her application. If the points are not high enough to admit her to one of her subject choices, the applicant is not admitted to that university even if her points exceed the admission threshold of the university.

Naturally, the acceptance threshold of each university is endogenous and will be determined by the number and quality of the applicants as well as the number of slots

available. The difficulty of obtaining a place to study varies considerably between these universities. Table 2 provides some descriptive statistics on the entry thresholds of different universities. Although the actual thresholds vary between years the ranking of the universities is similar from year to year. Helsinki School of Economics has historically been the most popular and hence the most selective of the Finnish schools providing education in economics and business administration. On the other hand, the University of Lapland had exam points thresholds that barely exceeded the minimum 10 points that was required to be admissible.

Table 2 also reports the means and variances of both thresholds for each university. As can be seen from the table, there is considerable variation also in the average thresholds and their variances. Variation across years tends to be larger in lower thresholds. Especially the mean and variance of the joint thresholds are strongly negatively correlated with a correlation coefficient of -0.59.

3 Theoretical background

Although multiple choice examinations are a widely studied subject in psychometrics, the economics literature analyzing optimal behaviour in such settings is very scarce. We are aware of only three attempts. Bernardo (1998) models the examinees as maximizing either the score or minimizing the probability of failing the exam. Burgos (2004) applies prospect theory and uses a utility function that assigns different values to losses and gains. In this paper, we follow Espinosa and Gardeazabal (2010) who combine decision and item response theory into a model of subject's behavior in multiple choice tests where risk attitudes play a role.

Espinosa and Gardeazabal (2010) assume that examinees are expected utility maximizers and that utility depends positively on the exam score. The exam consists of N items (or questions) each with M alternatives. For each item i there is a unique correct alternative and $M - 1$ incorrect ones. Items vary in their difficulty which is measured with scalar b_i . They assume that a correct answer yields a point while an incorrect carries a penalty $r > 0$. In our application r will be equal to 0.5. Omission is assumed to yield zero points.

Total of T applicants take part in the exam. Each of them has a fixed belief about their own ability θ_t . Applicants have partial knowledge, that is, they may be able to rule out some of the M alternatives but may not be sure which of the remaining alternatives is the correct one. This is captured by a latent variable $y_{ti} = \theta_t - b_i - v_{ti}$, where v_{ti} is a zero mean random variable with a distribution function $F(v_{ti})$. The applicant t knows the answer to the item i if $y_{ti} \geq 0$. Hence, the probability of knowing the correct answer is $P(y_{ti} \geq 0) = F(\theta_t - b_i)$. Denoting the event of a correct answer to item i by student t with $z_{ti} = 1$ and the probability of a pure guess being right for item i with c_i , $P(z_{ti} = 1)$ can be written as:

$$P(z_{ti} = 1) = c_i + (1 - c_i)P(y_{ti} \geq 0) \quad (1)$$

Espinosa and Gardeazabal (2010) are interested in the implications of risk aversion for optimal penalty for guessing in a context where students care about their final score s_t in the exam. In that case, risk aversion will affect the optimal guessing strategy. However, in the context of entrance examination, it is more realistic to assume that the applicants are not interested in the score s_t itself, but rather whether they gain entry to the university

or not. In this case, the utility function of the applicant t can then be written as:

$$u(s_t) = \begin{cases} u_e, & \text{if } s_t \geq s^* \\ u_n, & \text{if } s_t < s^* \end{cases} \quad (2)$$

where u_e and u_n are fixed values and $u_e > u_n$. In this context, the applicants will answer the items in the order of increasing difficulty, $P(y_{t1} \geq 0) \geq P(y_{t2} \geq 0) \geq \dots \geq P(y_{tN} \geq 0)$, to maximise the expected utility that can be written as:

$$E[u(s_t)] = P(s_t < s^*)u_n + P(s_t \geq s^*)u_e = u_n + P(s_t \geq s^*)(u_e - u_n) \quad (3)$$

The maximization of expected utility then boils down to maximization of the probability of getting in $P(s_t \geq s^*)$. This probability will depend on the probability with which the applicant can achieve the score s_t and, given that there is uncertainty about where the actual threshold is, on the probability that s_t exceeds the threshold.²

We denote the probability of answering k out of l items correctly and obtaining the score $s_{k,l} = k - r(l - k)$ with $p_k(l)$. Assuming that the entry thresholds are distributed normally with mean μ and variance σ^2 , the probability that $s_{k,l}$ exceeds the threshold is $\Phi(\frac{s_{k,l} - \mu}{\sigma})$. Hence, the probability of getting in by answering l items is:

$$P_l = \sum_{k=0}^l p_k(l) \Phi\left(\frac{s_{k,l} - \mu}{\sigma}\right) \quad (4)$$

In Figure 1 we illustrate numerically how the number of items that maximizes the probability of exceeding the threshold varies with the probability of answering items correctly and with the position of the threshold. In the left-hand side figure A, we have depicted the probability of exceeding the threshold that is normally distributed with mean 20 and variance 4 for applicants with different probabilities of answering the items correctly. The figure uses three probability profiles $p_i = 1 - d(i - 1)$ for $i = 1, 2, \dots, 40$ where $d = 0.02$ for high probability, $d = 0.0225$ for medium probability, and $d = 0.025$ for low probability profiles. The probability of each possible combination of answering k out of l items correctly, $p_k(l)$, and obtaining $s_{k,l} = k - 0.5(l - k)$ points can be calculated using the recursive algorithms in Kuo and Zuo (2003) that rely on optimal reliability modeling.³ Figure A clearly shows that the number of items that the individual should answer in order to maximise the probability of exceeding the threshold is increasing with the probability with which she answers the items correctly. The individual with the highest probabilities in figure A maximises the probability of gaining entry by answering 35 out of 40 items whereas the individuals with medium or low probabilities maximise their entry probabilities with 31 and 29 items, respectively.

In the right-hand side figure B of figure 1 provides a similar kind of illustration for applicants that are facing different thresholds. We apply three thresholds with equal standard deviations of 4: low with mean 16, medium with mean 20, and high with mean 24. The applicant is assumed to have the high probabilities $p_i = 1 - 0.02(i - 1)$ of answering items $i = 1, 2, \dots, 40$ correctly. As can be seen from figure B, the number of items that maximises the probability of entry is now increasing with the mean of the threshold

²We assume that the applicants answering behavior does not affect the threshold. This seems like a realistic assumption in an entrance examination setting where the number of applicants is large.

³We apply the recursive algorithm that relies on Markov Chain imbeddable structures. These algorithms greatly diminish the number of combinations that have to be calculated. See chapter 7 in Kuo and Zuo (2003) for details.

distribution. The individual facing the low threshold will maximise her probability of entry by answering 33 items whereas the individuals facing the medium or high threshold will maximise their entry probabilities by answering 35 and 36 items, respectively. These examples illustrate that maximising the probability of entry has implications for the omission behavior of individuals who have the same probabilities of answering items correctly but differ in the entry thresholds that they are facing.

4 Data and descriptive statistics

We have access to the full entrance examination data for the years 2005 to 2008. These data include the full entrance examination answers of all the applicants as well as their grades from the matriculation examination. The background information includes the gender of the applicant, birth year, municipality, mother tongue as well as the applicants ranking of the universities.

Table 3 presents descriptive statistics on the applicants at the different stages of the application process. The data contain information on 16,405 applicants of whom 45% are women who are clearly underrepresented among the applicants. The share of women among the accepted students is almost identical to their share among the applicants. Hence, on average women seem to be just as likely to gain entry to university as men are.

In the first nine rows of table 3 we have classified the applicants by the university which they ranked as their first choice in their applications. Looking across the nine different universities the share of female applicants varies between 40% and 49%. In the case of all but two universities we are able to reject the hypothesis that the share of male and female applicants is the same. However, the share of women among the accepted students varies considerably more from the low of 36% in the University of Joensuu to 48% in Oulu in which case we cannot reject the hypothesis that the share of male and female students is the same.

The last column of table 3 reports p-values of tests of the hypothesis that the female shares of applicants and accepted students are the same. These tests do not provide strong support for female underperformance in the admission process. In most cases we are unable to reject the hypotheses that the female share of applicants and accepted students is the same. Furthermore, in the cases where these shares differ, that is in the case of Helsinki School of Economics and University of Vaasa, the female share of accepted students is actually significantly higher than the female share of applicants. Only in the case of University of Turku the female share of accepted students is significantly lower.

However, it is important to condition the performance in the admission process on the starting points of the applicants. Table 4 provides a first glimpse on the gender differences in starting points and entrance examination performance. Gender differences in the starting points that are based on the high school matriculation examination are clear. Women have on average approximately starting points 19 points whereas men enter the application procedure with 17 points on average. These differences are robust also across the nine first ranked universities. Furthermore, examining the percentiles of the starting points distribution at the bottom rows of table 4 reveals that women almost first order stochastically dominate men in starting points. Hence, it seems very clear that women start the application process from a much better position than men.

The columns 4 to 6 of table 4 provide similar statistics on the entrance examination points. Here, the absence of clear gender differences is the most striking feature. On average, both men and women gain 10 points in the entrance examination. The amount of points gained varies across first ranked universities but gender differences still remain small

and do not have a consistent sign. However, the distributions of entrance examination points differ between men and women in interesting ways. The last 9 rows of table 4 reveal that male entrance examination point distribution seems to have larger tails than the female distribution. In particular at the 1st and 5th percentile men score lower points than women. At the same time, men score slightly higher at the 99th percentile although there the difference is not statistically significant.

The numbers in table 4 suggest that women have higher starting points than men and, on average, similar entrance examination points. Therefore we would expect women to be more likely to gain entry to a university in the admission process. Table 5 provides average acceptance rates by gender and by ranking order as well as across first ranked universities. The acceptance rates are almost identical across gender. On average 27% of male and female applicants get accepted and 17% to their first ranked university. Also across universities the gender differences in acceptance rates are small and rarely statistically significant. Hence, the evidence in table 5 imply that despite their advantageous starting points women do not do better than men in the application process. This pattern can be explained either by differences in the application behaviour or in the performance in the entrance examination. It could be that women apply to more demanding universities than men and therefore are not more likely to gain entry despite their more advantageous starting position or it could be that women simply under-perform in the entrance examination. In the next section we study these questions in detail.

5 Results

In this section, we will try to answer the following questions. First, using the information on the rankings of universities that the applicants state in their application forms, we study whether there are gender differences in the type of universities that the individuals choose to apply to. This analysis will address the question of whether women shy away from competition in the sense that they apply to universities that have lower thresholds than men with similar characteristics. Second, we will analyze the performance of students in the entrance examination. We examine whether there are gender differences in the entrance exam scores and in the probability of gaining entry to the university that the applicants apply to. Finally, we analyze the answering strategies of the applicants. In particular, we try to test whether women are more likely to omit items in the entrance exam than men and whether differences in the tendency to omit gives rise to gender differences in deviations from predicted optimal answering strategies which are estimated by a simple Rasch model that exploits the item by item information in the answer sheets

5.1 Do women shy away from competition?

As was noted in the introduction, one of the findings in the experimental literature has been that women are more reluctant to take part in high-stake competitive settings. Niederle and Vesterlund (2007), for example, found that when given the choice of payment scheme, women were more likely to choose a scheme that is based on absolute performance level rather than relative payment schemes. Field evidence for this kind of behavioural difference is scarce, however. Jurajda and Munich (2011), for example, do not find any gender difference in the degree of competitiveness of the universities for which the Czech students apply to and even though the gender difference in the likelihood of working under pay for performance payment scheme is the only significant gender difference that Manning and Saidi (2010) find, they argue that it is too small to explain any important

gender differences in labour market outcomes.

In this paper, we benefit from a fairly homogeneous population that has addressed a desire to study economics, finance, or business administration and has the possibility of choosing whether to apply for a more or less competitive university. The information in table 2 revealed that the Finnish universities differ substantially in their entry thresholds. The degree of difficulty of gaining entry to study depends crucially on the choice of university. Hence, shying away from competition in this context would mean applying to less demanding universities than other applicants with similar characteristics.

One way to study gender differences in the application behaviour is to examine how the previous year's threshold correlates with female dummy. In table 6, we present results from regressions of the previous years university acceptance threshold, based on the sum of starting points and entrance examination points, on the gender dummy and controls.⁴ The first column shows that the raw gender difference in the threshold is actually positive but not statistically different from zero. However, once we control for the starting points of the applicants as well as other observable controls, the gap becomes negative, -0.28, and significant. This means that women apply to less demanding universities than men with similar starting points. The difference of -0.28 points is still relatively small compared to the observed variation in the thresholds in table 2. Moreover, once we repeat the regression on the subsample of applicants who actually show up in the examinations, in the last column of table 6, we do not find any gender difference in the amount of points that the applicants has to gain in the entrance examination. Thus, these results seem to be in line with the findings of Jurajda and Munich (2011) as well as Manning and Saidi (2010) and do not point to any large gender differences in the tendency to avoid competitive settings.

5.2 Do women perform worse under competitive pressure?

The experimental literature provides even stronger evidence for gender differences in performance under competitive pressure and the field evidence, at least from educational settings, seems to concur with the experimental literature. In our setting, female under-performance would imply gaining less points in the entrance exam and being less likely to gain entry to a university than men with similar characteristics.

As we saw in tables 4 and 5, men and women score almost identical amounts of points in the entrance examination are equally likely to gain entry to the university that they apply to. However, these summary statistics mask relatively large gender differences starting points that were also reported in table 4. Since we do not find any large differences in the type of universities that men and women apply to above, we would expect the combination of no gender differences in performance in the entrance exam and larger female starting points to mean that women under-perform in the entrance exam relative to men with similar starting points.

This question is examined in tables 7 and 8 where we regress the points scored in the entrance exam and the entry probability, respectively, on the female dummy and starting points as well as on controls for the first choice university and the same set of observables as in table 6. The first column of table 7 confirms the result in table 4 according to which there are no significant gender differences in the average exam scores. However, once we control for starting points in the column (2) of table 7 the gender difference in starting points becomes clear. Women score over half a point, -0.57 points, less than men

⁴Using the threshold based on entrance examination points alone as the dependent variable yields qualitatively similar results as the ones in table 6.

who are applying to the same universities with similar starting points. This result is very robust to the inclusion of additional controls in the last column of table 7. This should mean that women have a lower probability of gaining entry to the university than men who start the entrance examination in a similar setting.

This intuition is confirmed in table 8. While there are no significant gender differences in the probability of gaining entry to the university on average, controlling for the applicant's starting points as well as for the university where he or she is applying to reveals that women are almost three percentage points less likely of getting accepted than men who come to the entrance examination in a similar setting. This translates to probability difference of approximately 8% which can be considered to be a large difference in the probability of entry. Again, this gender difference is very robust to the inclusion of additional controls.

5.3 Do women omit too many items?

The evidence presented above suggests that women under-perform in the entrance examinations. Despite their higher starting points women are less likely to get accepted. This evidence is in line with the previous results in the numerous lab experiments and in the limited number of field studies that have examined this question. However, the analysis that only limits to the examination of total scores or probabilities of entry does not really reveal why these gender differences in performance arise. They could naturally reflect genuine performance differences in the exam in the sense that men simply are better at answering these kinds of multiple choice entrance exam questions. However, the theoretical discussion in section 3 revealed that the choice of how many items to respond to has crucial implications for the entry probabilities in this setting. The question that naturally arises is whether this under-performance of women is due to strategic choices in the entrance examination.

Table 9 presents descriptive statistics on the number of omitted items by gender.⁵ Women do omit more items (12) than men (11) in these entrance examinations on average. Examining differences in the tendency to omit by first choice universities reveals that this gender difference in the tendency to omit is reasonably constant across first choices. In four out of nine universities female omission rates are significantly higher and they are never significantly lower than the males rates.

These gender differences in the number of omitted items could simply reflect the fact that women need to score less points in the exam and may therefore omit more items. After all the results presented above showed that women apply to slightly less demanding universities and have, on average, higher starting points. Therefore, women should have less distance to cover in the exam and could omit more items.

Table 10 presents results from similar regressions as the table 7 but now the exam scores are replaced with the number of omitted items as the dependent variable. Columns (2) and (3) of table 10 clearly show that the gender difference in the number of omitted items is robust to controlling for both the starting points of the applicant as well as the first choice university. These results rather convincingly show that women omit one item more than men with similar starting points and applications.

However, the results in table 10 do not really show that women who omit more items than men are behaving sub-optimally. After all, gender differences in omissions may reflect either under-responding by women or over-responding by men. Furthermore, the

⁵In the analysis of omissions we only use data on 2005 and 2006 exams. After 2006 scoring and the number of items was changed so that the analysis of the omission patterns with full data would be difficult.

differences in omission rates may arise from real differences in the knowledge about the exam material. In order to determine whether the differences in the number of omitted items can explain the performance differences presented above one would need to know whether women deviate more from their optimal answering strategies than men.

It is of course difficult to determine the optimal answering strategies with observational data where we do not have any additional information about the reasons for why individuals omit items other than what is contained in the application forms and in the answer sheets. However, the information provided by the applicants' answers can be used to derive the predicted probabilities of answering each item correctly. This procedure, also known as the Rasch Model, is commonly used in Item Response Theory.

To remain consistent with the notation used in section 3, denote the probability that the applicant t answers the item i correctly with $Pr(Z_{ti} = 1)$. In the Rasch model this probability is defined as:

$$Pr(Z_{ti} = 1) = \frac{\exp(z_{ti}(\theta_t - \delta_i))}{1 + \exp(\theta_t - \delta_i)} \quad (5)$$

for items $i = 1, \dots, N$ where θ_t is the fixed effect for the applicant t and δ_i is "the difficulty parameter" for item i . Under the assumption that answers to different items are independent after controlling for fixed effects, the likelihood for the applicant t is:

$$L_t(\boldsymbol{\delta}, \theta_t | \mathbf{z}_t) = \prod_{i=1}^N Pr(z_{ti} = 1 | \theta_t, \delta_i) \quad (6)$$

with $\mathbf{z}_t = (z_{ti})_{i=1, \dots, N}$ and $\boldsymbol{\delta} = (\delta_i)_{i=1, \dots, N}$. The Rasch model has the useful property that the number of correct answers is a sufficient statistic for the latent trait θ_t . Therefore the estimation of the Rasch model simplifies to a logit model where the probability of answering each item correctly is regressed on the full set of 40 item dummies and a set of 40 dummies for the counts of correct answers.

The estimated coefficients of the Rasch model can be used to derive the predicted probabilities of answering each item correctly for each individual. The average predicted probabilities are plotted by gender and item in figure 2. As can be seen from this figure, the predicted probabilities are almost linearly decreasing in item difficulty and are virtually identical for men and women.

The predicted probabilities can be used in the same way as the examples of probability profiles that were used in section 3 to derive the optimal number of item responses in the entrance exam. As the mean of the entry threshold we use the past year's thresholds, both the ones based on the entrance examination points alone and the ones based on the sum of entrance examination points and starting points, and as the standard deviation the sample standard deviation of that particular threshold in the data. Similarly as in section 3, we assume that the thresholds are distributed normally. With this information, the vector of predicted probabilities of answering items correctly for each individual and the probability that each score exceeds the threshold of the university that the applicant is applying to, we can calculate the number of items that maximises the probability of entry for each individual in the data following the same recursive algorithms as in section 3.

Table 11 presents the descriptive statistics on the observed answering behaviour as well as on the optimal behaviour by gender. The first row of the table repeats the fact that is already familiar from table 10. Women answer fewer items in the entrance examinations than men. As can be seen from the second row of table 11, this answering pattern

nevertheless yields a higher probability of entry for women than for men. The difference in the predicted probabilities of entry based on actual behaviour is approximately one percentage point in favour of women and significant. However, the third row shows that the optimal number of items, the one that maximizes the predicted probability of entry, is higher than the actual number of answered items both for men and for women. This optimal number of answers is slightly, but significantly, higher for men, at 30.9 items, than for women, at 30.7 items. The optimal number of items would naturally yield higher average predicted probabilities of entry for both men and women so that the difference in the probabilities would remain approximately at the same level as with the probabilities based on the actual behaviour. The last two rows of table 11 provides summary statistics on how much men and women deviate, on average, from the behaviour that is optimal based on the Rasch model. By answering 2.8 items more women could increase their probability of entry by 1.1 percentage points. Men would have to answer only 1.9 items more to reach the maximum predicted probability of entry which is 0.9 percentage points higher than the predicted probability based on actual behaviour.

These summary statistics clearly imply that both men and women are answering, on average, too few items in the entrance examinations. The deviation from optimal answering strategy also seems to be larger for women who would need to answer more additional items to arrive at the optimum and by doing so would increase their probability of entry more than men would by obtaining the optimum. This interpretation of the summary statistics is confirmed when we plot the Kernel distributions of the deviations from the predicted optimal number of items in figure 3. This figure reveals that the female distribution clearly lies on the left of the male distribution. The fraction of women who answer too few items is substantially larger than the fraction of men. At the same time, there are fewer women that answer too many questions than men. Hence, women deviate more from the predicted optimal behaviour and they do that by answering too few items.

The analysis based on the Rasch model seems to confirm that women are less likely to gain entry to university because they are too timid in their answering strategies. If we take the results literally, women are omitting items that would increase their probability of entry. This would be consistent with the argument that gender differences in performance in competitive settings such as entrance examinations may arise because of differences in strategic choices. However, we should also be careful when interpreting these results. These kind of answering patterns could also arise if the applicants have some information on their probabilities of answering correctly that is not observable to us. In that case the differences in performance would reflect genuine differences in ability to answer the items rather than strategic choices. Based on the data at hand, it is not possible to fully discriminate between these explanations.

6 Conclusions

The extensive lab and much more limited field literature has rather convincingly documented that women tend to perform worse in competitive settings than men. However, very little is known about the reasons for why this gender gap in performance emerges. While the theoretical literature has highlighted the importance of strategic choices in competitive settings, the empirical evidence, and especially the field evidence, on gender differences in strategic choices and their implications has been limited.

In this paper we analyze performance differences in university entrance examinations. We exploit data from the joint exams that the nine Finnish universities providing education in economics and business use to choose their students. Entry to the universities is

based on the joint sum of starting points that the applicants are credited based on their high school matriculation exam results and the points from the entrance examination. The entrance examination is a multiple choice test where wrong answers are penalized and omissions yield zero points. This feature of the entrance examinations means that guessing and omission strategies of the applicants matter for their probabilities of gaining entry to the university that they apply to.

Our evidence show that there are no large differences in how men and women perform in the entrance examinations on average. However, once we control for the starting points of the applicants, clear gender differences arise. Women have clearly higher starting points than men on average which reflects their superior performance in secondary education. Once these differences in the starting points are controlled for, women clearly score lower in the entrance exam and are less likely to gain entry to university than men.

We also analyze whether these gender difference in performance can potentially be explained by differences in answering strategies. Women clearly omit more items than men in the entrance examinations. We use the Rasch Model to derive predictions of the individual probabilities of answering items correctly in the entrance exam. Using these predicted probabilities we can derive the number of items that each individual should answer to maximise their probability of gaining entry. Our results show that women deviate more from the optimal number of items than men and that they do so because they tend to answer too few items.

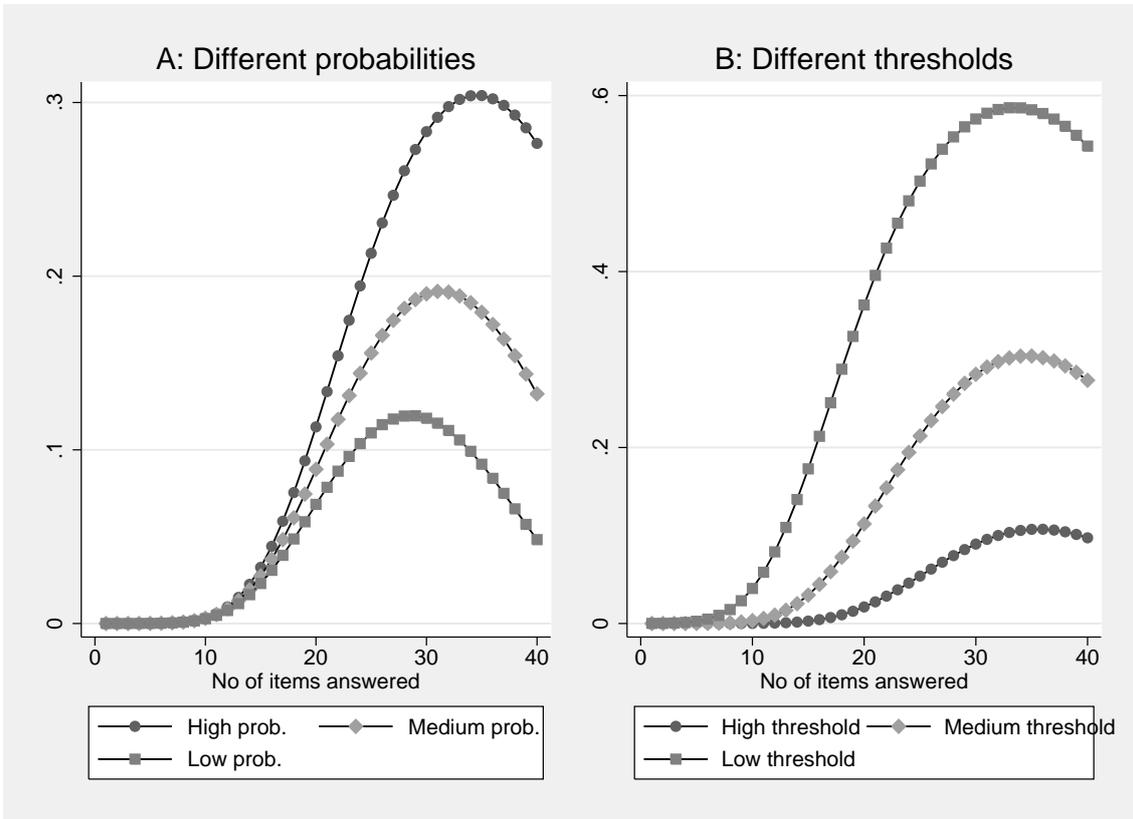
We think that these results are consistent with the argument that differences in strategies may explain, at least partly, the observed gender differences performance under competitive pressure. The women in these entrance examinations seem to be more conservative and timid test takers than men. While these tendencies could benefit women if the objective was simply to maximise the expected score, the failure to answer items that the respondent is less sure about may hurt the chances of exceeding the entry threshold. However, we should be cautious in interpreting these results. Since responses of the applicants may be driven by information about their own abilities that is unobservable to us, these answering patterns could also reflect genuine differences in the ability to answer the items. Future field and lab work should nevertheless devote effort to disentangle the role of strategic and behavioral factors in explaining the gender differences in performance under competitive pressure.

References

- Bernardo, J. M. (1998). A decision analysis approach to multiple-choice examinations. In *Applied decision analysis* (pp. 195–207). Springer.
- Bronars, S. (1987). Risk taking in tournaments. *Unpublished Working Paper-University of Texas at Austin*.
- Burgos, A. (2004). Guessing and gambling. *Economics Bulletin*, 4(4), 1–10.
- Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review*, 101(2), 556–590.
- Espinosa, M. P., & Gardeazabal, J. (2010). Optimal correction for guessing in multiple-choice tests. *Journal of Mathematical Psychology*, 54(5), 415–425.
- Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3), 1049–1074.
- Goel, A. M., & Thakor, A. V. (2008). Overconfidence, ceo selection, and corporate governance. *The Journal of Finance*, 63(6), 2737–2784.

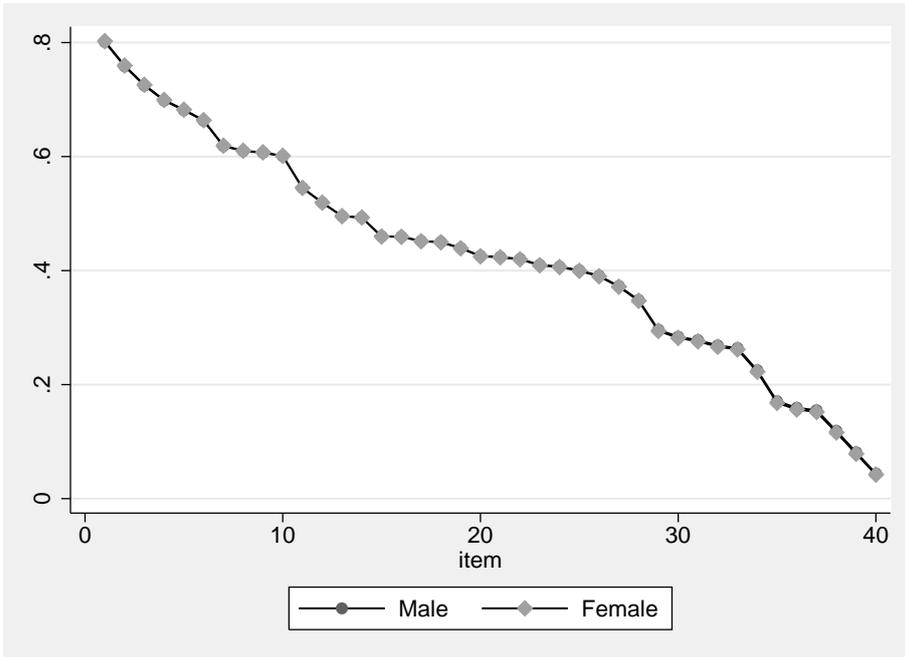
- Hvide, H. K. (2002). Tournament rewards and risk taking. *Journal of Labor Economics*, 20(4), 877–898.
- Jurajda, S., & Munich, D. (2011). Gender gap in performance under competitive pressure: Admissions to czech universities. *The American Economic Review*, 101(3), 514–518.
- Kuo, W., & Zuo, M. J. (2003). *Optimal reliability modeling: principles and applications*. Wiley. com.
- Manning, A., & Saidi, F. (2010). Understanding the gender pay gap: what’s competition got to do with it? *Industrial and Labor Relations Review*, 681–698.
- Nekby, L., Thoursie, P. S., & Vahtrik, L. (2008). Gender and self-selection into a competitive environment: Are women more overconfident than men? *Economics Letters*, 100(3), 405–407.
- Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The Quarterly Journal of Economics*, 122(3), 1067–1101.
- Ors, E., Palomino, F., & Peyrache, E. (2013). Performance gender gap: Does competition matter? *Journal of Labor Economics*, 31(3), 443–499.

Figure 1: Probability of exceeding the threshold by number or items



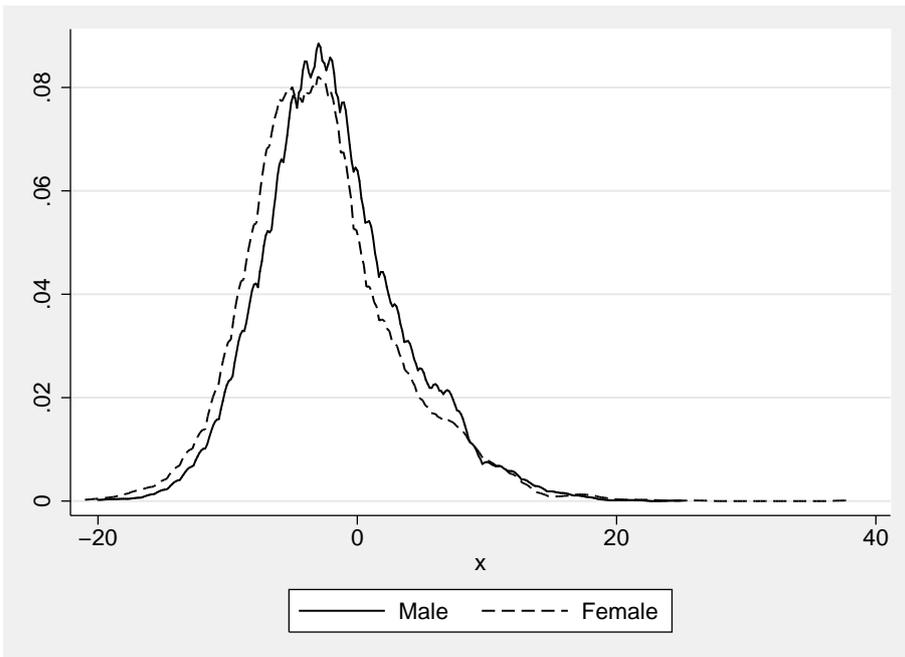
Note: The probabilities of answering each item correctly are determined by equation $p_i = 1 - d(i - 1)$ for $i = 1, 2, \dots, 40$ where $d = 0.02$ for high probability, $d = 0.0225$ for medium probability, and $d = 0.025$ for low probability profiles. The thresholds are distributed normally with standard deviation of 4 so that the medium threshold, which is also used in the left-hand side figure, has a mean of 20, the low threshold has a mean of 15, and the high threshold has a mean of 25.

Figure 2: Average predicted probabilities of answering items correctly by gender



Note: The predicted probabilities are the fitted values from the logit model where the probability of answering correctly to each item is regressed on a full set of item dummies and dummies for each count of correct answers.

Figure 3: Deviation from the optimal number of items by gender



Note: Kernel distributions of the predicted optimal number of items and the actual number of items by gender. 2005-2006 data.

Table 1 Starting points based on matriculation exam results

| <i>Subject</i> | <i>7 (100-95 percentiles)</i> | <i>6 (95-80 percentiles)</i> | <i>5 (80-60 percentiles)</i> | <i>4 (60-36 percentiles)</i> | <i>3 (36-16 percentiles)</i> |
|---------------------------------|---------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| Finnish | 9 | 8 | 6 | 3 | 1 |
| Mathematics, long | 10 | 9 | 7 | 4 | 2 |
| Mathematics, short | 7 | 6 | 4 | 3 | 1 |
| Other subject, long | 7 | 6 | 4 | 3 | 1 |
| Other subject, short | 5 | 4 | 3 | 2 | 1 |

Note: Entries in the table report the amount of points that are credited for the performance in the various high school matriculation exam subjects in the entrance exam.

Table 2 Entry thresholds of the universities

| | | 2005 | 2006 | 2007 | 2008 | Mean | Variance |
|---------------------|--------------|-------------|-------------|-------------|-------------|-------------|-----------------|
| Helsinki | <i>Joint</i> | 57 | 55.5 | 53 | 52.75 | 54.6 | 3.16 |
| | <i>Exam</i> | 28.5 | 27 | 24 | 23.5 | 25.79 | 4.34 |
| Joensuu | <i>Joint</i> | 36 | 32 | 30.5 | 28 | 31.74 | 9.45 |
| | <i>Exam</i> | 14.5 | 12.5 | 12.25 | 10.5 | 12.47 | 1.51 |
| Kuopio | <i>Joint</i> | 38 | 36 | 34.5 | 35.25 | 35.92 | 1.73 |
| | <i>Exam</i> | 16 | 14.5 | 14 | 13 | 14.35 | 1.27 |
| Lapland | <i>Joint</i> | 34 | 28 | 29.75 | | 30.79 | 6.52 |
| | <i>Exam</i> | 13.5 | 10 | 10.25 | | 11.37 | 2.66 |
| Lappeenranta | <i>Joint</i> | 41 | 39 | 37.25 | 36.25 | 38.55 | 3.38 |
| | <i>Exam</i> | 19 | 17 | 15.5 | 14.5 | 16.67 | 2.99 |
| Oulu | <i>Joint</i> | 41.5 | 36.5 | 37.25 | 33.75 | 37.36 | 8.03 |
| | <i>Exam</i> | 17.5 | 14.5 | 12.5 | 10.5 | 13.84 | 6.91 |
| Tampere | <i>Joint</i> | 48 | 48.5 | 45.75 | 45.25 | 46.96 | 1.93 |
| | <i>Exam</i> | 20.5 | 21 | 18.5 | 18.25 | 19.64 | 1.44 |
| Turku | <i>Joint</i> | 47.5 | 44 | 45.25 | 42.5 | 44.85 | 3.41 |
| | <i>Exam</i> | 21 | 17.5 | 17.25 | 15.75 | 17.89 | 3.79 |
| Vaasa | <i>Joint</i> | 39.5 | 38.5 | 35.75 | 34 | 37.15 | 4.85 |
| | <i>Exam</i> | 18 | 16 | 11.25 | 11.5 | 14.5 | 8.74 |

Note: Entries in the table reports the realized entrance thresholds of the universities by year. Joint refers to the thresholds that is based on the sum of starting points and the entrance exam points and based on which 60% of the students are chosen. Exam refers to the threshold that is based on the entrance exam points alone and based on which 40% of the students are chosen.

Table 3 Shares of male and female applicants and accepted students by university

| <i>University</i> | <u>Applicants</u> | | Test: $P[\text{Pr}(\text{Male})=\text{Pr}(\text{Female})]$ | <u>Accepted</u> | | Test: $P[\text{Pr}(\text{Male})=\text{Pr}(\text{Female})]$ | Test: $P[\text{Pr}(\text{Female}_{\text{Applicant}})=\text{Pr}(\text{Female}_{\text{Accepted}})]$ |
|---------------------|-------------------|---------------|--|-----------------|---------------|--|--|
| | <i>Male</i> | <i>Female</i> | | <i>Male</i> | <i>Female</i> | | |
| <i>Helsinki</i> | 57.79 | 42.21 | 0.000 | 54.69 | 45.31 | 0.000 | 0.034 |
| <i>Joensuu</i> | 59.52 | 40.48 | 0.000 | 63.98 | 36.02 | 0.000 | 0.207 |
| <i>Kuopio</i> | 58.52 | 41.48 | 0.000 | 60.45 | 39.55 | 0.000 | 0.604 |
| <i>Lapland</i> | 52.98 | 47.02 | 0.123 | 57.24 | 42.76 | 0.014 | 0.389 |
| <i>Lappeenranta</i> | 55.06 | 44.94 | 0.000 | 56.76 | 43.24 | 0.000 | 0.497 |
| <i>Oulu</i> | 54.90 | 45.10 | 0.000 | 52.00 | 48.00 | 0.149 | 0.190 |
| <i>Tampere</i> | 51.09 | 48.91 | 0.092 | 52.90 | 47.10 | 0.047 | 0.423 |
| <i>Turku</i> | 51.91 | 48.09 | 0.001 | 55.65 | 44.35 | 0.000 | 0.027 |
| <i>Vaasa</i> | 57.27 | 42.73 | 0.000 | 54.88 | 45.12 | 0.000 | 0.027 |
| <i>Total</i> | 55.37 | 44.63 | 0.000 | 55.32 | 44.68 | 0.000 | 0.955 |

Note: The first two columns report the shares of male and female applicants by first choice university. The third column reports the p-value of the test of the hypothesis that these shares are equal. The columns 4 and 5 report the shares of male and female students among the accepted. The sixth column reports the p-value of the test of the hypothesis that these shares are equal. The seventh column reports the p-values from a test of the hypothesis that female shares in columns 2 and 5 are equal.

Table 4 Starting and entrance examination points by gender

| <i>University</i> | Starting points | | | Exam Points | | |
|-----------------------------------|------------------------|---------------|---------------------|--------------------|---------------|---------------------|
| | <i>Male</i> | <i>Female</i> | <i>Diff</i> | <i>Male</i> | <i>Female</i> | <i>Diff.</i> |
| <i>Helsinki</i> | 19.01 | 21.05 | 2.04 ^{***} | 13.00 | 13.61 | 0.61 ^{***} |
| <i>Joensuu</i> | 14.29 | 17.37 | 3.08 ^{***} | 5.43 | 5.18 | -0.25 |
| <i>Kuopio</i> | 13.88 | 16.11 | 2.23 ^{***} | 5.47 | 5.30 | -0.17 |
| <i>Lapland</i> | 12.34 | 15.84 | 3.50 ^{***} | 5.42 | 5.13 | -0.30 |
| <i>Lappeenranta</i> | 14.67 | 16.64 | 1.97 ^{***} | 7.27 | 7.13 | -0.14 |
| <i>Oulu</i> | 15.95 | 17.69 | 1.74 ^{***} | 6.95 | 7.48 | 0.53 |
| <i>Tampere</i> | 17.77 | 19.56 | 1.79 ^{***} | 9.23 | 9.37 | 0.14 |
| <i>Turku</i> | 17.75 | 19.78 | 2.03 ^{***} | 10.24 | 9.53 | -0.72 ^{**} |
| <i>Vaasa</i> | 13.92 | 16.84 | 2.92 ^{***} | 6.75 | 6.55 | -0.20 |
| <i>Total</i> | 17.27 | 19.34 | 2.07 ^{***} | 10.07 | 10.18 | 0.11 |
| <i>1st percentile</i> | 0 | 0 | 0 | -6.25 | -5.5 | 0.75 [*] |
| <i>5th percentile</i> | 2 | 1 | -1 | -2.5 | -2 | 0.5 ^{**} |
| <i>10th percentile</i> | 6 | 7 | 1 ^{***} | -0.5 | -0.25 | 0.25 |
| <i>25th percentile</i> | 11 | 14 | 3 ^{***} | 3.5 | 3.25 | -0.25 |
| <i>Median</i> | 17 | 20 | 3 ^{***} | 9 | 9.25 | 0.25 |
| <i>75th percentile</i> | 23 | 26 | 3 ^{***} | 16.25 | 16.5 | 0.25 |
| <i>90th percentile</i> | 29 | 31 | 2 ^{***} | 22.5 | 22.5 | 0 |
| <i>95th percentile</i> | 32 | 33 | 1 ^{***} | 25.5 | 25.5 | 0 |
| <i>99th percentile</i> | 37 | 37 | 0 | 30 | 29.5 | -0.5 |

Note: The entries in first ten rows report average male and female starting and entrance exam points as well as their differences. ^{***} indicate significance at 1%-level, ^{**} at 5%-level, and ^{*} at 10%-level. The last 10 rows of the table report the percentiles of the starting points and entrance exam points by gender as well as their differences.

Table 5 Acceptance rates by gender and university

| University | Male | Female | Diff. |
|---------------------|-------------|---------------|---------------------|
| <i>Helsinki</i> | 0.13 | 0.15 | 0.01 ^{**} |
| <i>Joensuu</i> | 0.11 | 0.10 | -0.01 |
| <i>Kuopio</i> | 0.09 | 0.09 | 0.00 |
| <i>Lapland</i> | 0.10 | 0.11 | 0.01 |
| <i>Lappeenranta</i> | 0.14 | 0.15 | 0.01 |
| <i>Oulu</i> | 0.15 | 0.19 | 0.03 ^{**} |
| <i>Tampere</i> | 0.08 | 0.08 | 0.00 |
| <i>Turku</i> | 0.15 | 0.13 | -0.02 ^{**} |
| <i>Vaasa</i> | 0.16 | 0.17 | 0.01 |
| <i>Total</i> | 0.27 | 0.27 | 0.00 |
| <i>First rank</i> | 0.17 | 0.17 | 0.00 |
| <i>Second rank</i> | 0.07 | 0.06 | -0.01 ^{**} |
| <i>Third rank</i> | 0.03 | 0.04 | 0.00 |
| <i>Rejected</i> | 0.73 | 0.73 | 0.00 |

Note: The entries in the first nine rows report the average acceptance rates by gender and by first choice universities as well as their differences. ^{***} indicate significance at 1%-level, ^{**} at 5%-level, and ^{*} at 10%-level. The last five rows report the average acceptance rates by gender, the share of applicants that are accepted to their first, second, and third choice university by gender and their differences. The final row reports the shares of students that do not gain entry to any university.

Table 6: Regression of the previous year's acceptance threshold on the female dummy, starting points and observable characteristics

| VARIABLES | (1) | (2) | (3) |
|-------------------------------------|----------------------|---------------------------|---------------------------|
| | All | All | Show up |
| Female = 1 | 0.100 (0.133) | -0.272** (0.128) | -0.0357 (0.140) |
| Starting points | | 0.172*** (0.00727) | 0.167*** (0.00812) |
| Age | | -0.142*** (0.0141) | -0.168*** (0.0171) |
| Distance to the university in 100km | | -0.00827*** (0.000448) | -0.00880*** (0.000490) |
| Second time applying | | 0.0153 (0.146) | -0.188 (0.157) |
| Third time applying | | -0.0144 (0.299) | -0.120 (0.326) |
| Fourth time applying | | -0.749 (0.740) | -1.434* (0.850) |
| Year 2006 | | -1.418*** (0.157) | -1.144*** (0.171) |
| Year 2007 | | -3.259*** (0.157) | -3.246*** (0.170) |
| Constant | 47.79*** (0.0884) | 50.38*** (0.390) | 51.10*** (0.457) |
| Observations | 14,611 | 14,476 | 11,839 |
| R-squared | 0.000 | 0.095 | 0.100 |

Note: The table reports the results from the OLS regressions of the past year's threshold of the university that the applicants apply to on gender dummy, the amount of starting points, and controls. Standard error are reported in parentheses. *** indicate significance at 1%-level, ** at 5%-level, and * at 10%-level.

Table 7 Regression of entrance examination scores on female dummy and control variables

| | (1) | (2) | (3) |
|---------------------------------|----------------------|-----------------------|-----------------------|
| VARIABLES | | | |
| Female = 1 | 0.109 (0.132) | -0.566*** (0.119) | -0.590*** (0.117) |
| Starting points | | 0.315*** (0.00690) | 0.333*** (0.00689) |
| First choice university dummies | | Yes | Yes |
| Controls | | Yes | Yes |
| Constant | 10.07*** (0.0877) | 3.872*** (0.523) | 1.235** (0.629) |
| Observations | 17,355 | 17,354 | 17,354 |
| R-squared | 0.000 | 0.212 | 0.238 |

Note: The table reports the results from the OLS regressions of the entrance examination scores on gender dummy, the amount of starting points, and controls. Controls include age of the applicant, distance to the first choice university in 100km, dummies for applying for the second, third, and fourth times, as well as year dummies. Standard errors are reported in parentheses. *** indicate significance at 1%-level, ** at 5%-level, and * at 10%-level.

Table 8 Regression on entry probability on female dummy and controls

| | (1) | (2) | (3) |
|---------------------------------|-----------------------|-------------------------|-------------------------|
| Female = 1 | 0.000459 (0.00616) | -0.0288*** (0.00590) | -0.0255*** (0.00587) |
| Starting points | | 0.0145*** (0.000336) | 0.0143*** (0.000339) |
| First choice university dummies | | Yes | Yes |
| Controls | | | Yes |
| Constant | 0.274*** (0.00411) | -0.0161 (0.0243) | 0.0877*** (0.0294) |
| Observations | 21,262 | 21,122 | 21,122 |
| R-squared | 0.000 | 0.105 | 0.116 |

Note: The table reports the results from the OLS regressions of the probability of gaining entry to university on gender dummy, the amount of starting points, and controls. Controls include age of the applicant, distance to the first choice university in 100km, dummies for applying for the second, third, and fourth times, as well as year dummies. Standard errors are reported in parentheses. *** indicate significance at 1%-level, ** at 5%-level, and * at 10%-level.

Table 9 Number of omitted items by gender

| University | Male | Female | Diff. |
|---------------------|--------------|---------------|---------------|
| <i>Helsinki</i> | 8.18 | 9.54 | 1.35** |
| <i>Joensuu</i> | 16.04 | 15.65 | -0.38 |
| <i>Kuopio</i> | 13.06 | 13.06 | 0.00 |
| <i>Lapland</i> | 17.41 | 17.58 | 0.17 |
| <i>Lappeenranta</i> | 13.75 | 13.88 | 0.13 |
| <i>Oulu</i> | 15.77 | 16.22 | 0.45 |
| <i>Tampere</i> | 11.08 | 12.36 | 1.28** |
| <i>Turku</i> | 9.40 | 11.34 | 1.94** |
| <i>Vaasa</i> | 14.61 | 15.59 | 0.98** |
| <i>Total</i> | <u>10.95</u> | <u>12.12</u> | <u>1.17**</u> |

Note: The entries in the table report the average number of omitted items by gender and by first choice universities as well as their differences. *** indicate significance at 1%-level, ** at 5%-level, and * at 10%-level.

Table 10 Regression of the number of omitted items on female dummy and control variables

| | (1) | (2) | (3) |
|---------------------------------|----------------------|------------------------|------------------------|
| VARIABLES | | | |
| Female = 1 | 1.172*** (0.143) | 1.035*** (0.131) | 1.082*** (0.130) |
| Starting points | | 0.0349*** (0.00763) | 0.0205*** (0.00769) |
| First choice university dummies | | Yes | Yes |
| Controls | | Yes | Yes |
| Constant | 10.95*** (0.0962) | 7.620*** (0.187) | 10.15*** (0.422) |
| Observations | 8566 | 8566 | 8566 |
| R-squared | 0.008 | 0.189 | 0.204 |

Note: The table reports the results from the OLS regressions of the number of omitted items on gender dummy, the amount of starting points, and controls. Controls include age of the applicant, distance to the first choice university in 100km, dummies for applying for the second, third, and fourth times, as well as year dummies. Standard errors are reported in parentheses. *** indicate significance at 1%-level, ** at 5%-level, and * at 10%-level.

Table 11

| Variable | Male | Female | Diff. |
|---|-------------|---------------|-----------------------|
| <i>Actual no. of items answered</i> | 29.05 | 27.88 | -1.172 ^{***} |
| <i>Actual prob. of entry</i> | 0.082 | 0.094 | 0.012 ^{**} |
| <i>Optimal no. of items</i> | 30.93 | 30.67 | -0.265 ^{**} |
| <i>Max. prob. of entry</i> | 0.091 | 0.105 | 0.014 ^{**} |
| <i>Distance to the optimal no. of items</i> | -1.863 | -2.784 | -0.921 ^{**} |
| <i>Distance to the max. prob. of entry</i> | -0.009 | -0.011 | -0.003 ^{**} |

Note: The table reports the average number of items by gender as well as their difference. The second row reports the predicted probability of entry based in this number of items by gender as well as their difference. The third row reports the optimal number of items based on the analysis of the Rasch Model and the fourth row the predicted probabilities based on these numbers of items. *** indicate significance at 1%-level, ** at 5%-level, and * at 10%-level.