IFAU
Institute for Evaluation of Labour
Market and Education Policy

Examination behavior
– Gender differences in preferences?

Lena Nekby
Peter Skogman Thoursie
Lars Vahtrik

WORKING PAPER 2013:1

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

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

ISSN 1651-1166

Examination behavior – Gender differences in preferences?[§]

by

Lena Nekby[¨], Peter Skogman Thoursie[©] and Lars Vahtrik[a]

January 10, 2013

Abstract

A unique examination strategy in first year microeconomics courses is used to test for gender differences in preferences in examination behavior. Students have the possibility of attaining a seminar bonus on the final exam for near-perfect seminar attendance and are given two voluntary initial quizzes during the semester. At the final exam, the scores received on initial quizzes can either be accepted as is, or students can attempt to improve their marks by answering similar quiz questions on the exam. Results suggest that female students are more likely to take initial quizzes and receive a seminar bonus but are less likely to re-take quiz-questions on the final exam. These results suggest higher risk aversion among female students relative to male students, behavioral differences with tangible implications in terms of final grades on the course.

Keywords: Gender, preferences, risk aversion, overconfidence
JEL codes: J16, A12, A14

[§] The authors are grateful for comments from Antonio Cabrales, Erik Grönqvist, Caroline Hall, Jonas Karlsson, Tuomas Pekkarinen, Arthur Schram, Björn Öckert and seminar participants at the 2009 EAEPE annual conference and the 2010 COSME-FEDEA Annual Workshop on Gender Economics.

[¨] Lena Nekby: Department of Economics, Stockholm University. lena.nekby@ne.su.se.
[©] Peter Skogman Thoursie: Department of Economics, Stockholm University and Institute for Labour Market Policy Evaluation (IFAU); peter.thoursie@ne.su.se.
[a] Lars Vahtrik: Department of Economics, Stockholm University, lv@ne.su.se.

Table of contents

1 Introduction

At present, there is a large and growing research interest on gender differences in preferences (see Croson and Gneezy (2009) and Bertrand (2010) for excellent overviews of this literature). Women, in this largely experimental literature, have been found to be more risk-averse on average, less willing to compete, less overconfident, more altruistic and more inequality averse than likewise men. There is however a "striking" lack of research in real settings establishing the empirical relevance of these factors for actual outcomes (Bertrand, 2010). We hope to fill this gap in at least one context by studying gender differences in behavior in an academic setting, namely among undergraduate students in a first year economics course at Stockholm University. Students in this course face a number of decisions regarding examination strategy, both prior to and during the actual final examination, with potential repercussions for actual educational outcomes. The aim of our study is to examine if there are gender differences in behavior in examination strategies and to estimate to what degree these systematic behavioral differences are costly in terms of final grades on the exam (and the course).

Students in the first year microeconomic courses at Stockholm University are offered two voluntary quizzes during the semester, the scores of which are credited to the final exam. Quiz takers, however, also have the opportunity of re-taking corresponding quiz questions on the final exam to improve (or lower) their earlier established scores. In addition, students can be awarded a seminar bonus on the final examination for near-perfect seminar attendance during the semester. This set-up allows us to investigate gender differences in examination behavior in three ways. First, taking a quiz is one way to increase the probability of higher exam scores since quiz-takers can redo the quiz on the exam (i.e., have a second chance). Note that quiz-takers also have potentially more time to devote to other exam questions given the fixed time constraint for the final exam. Second, near-perfect seminar attendance yields both an examination bonus and, again, more time on the final exam, both of which should increase the probability of attaining higher final exam scores. Third, quiz-takers have the opportunity of improving their scores by redoing one or both of the quizzes on the exam. The question to answer is whether there are gender differences in examination

behavior regarding these three choices and if these potential behavioral differences are costly in terms of final scores on the exam.

Note that there are a number of potential mechanisms behind observed behavioral differences in examination strategies having to do with, for example, gender differences in confidence, risk-aversion, competitiveness and/or time usage (such as procrastination) which are difficult to tease out, especially in the type of real world setting examined here. Bertrand (2010) notes, for example, that gender differences in risk attitudes may be a consequence of male overconfidence in their relative ability. Likewise, women may systematically underperform in competitive environments and shy away from competition, both of which may be due to gender differences in confidence or risk attitudes. A thoughtfully carried out experimental study can at times get close to weeding out the influence of a particular type of preference component on outcomes of interest, for example, by changing a controlled environment and introducing competition without a change of risk-levels and/or by controlling for risk via attitudinal survey questions. The question remains, however, if it is possible to distinguish between, for example, risk-aversion and overconfidence in a real world setting where both may exist simultaneously.

A related issue concerns other factors that may be considered as part of, or correlated to, individual preferences. If underlying skills (ability, motivation etc.) are correlated with components such as risk aversion and (over)confidence, it is unclear how to interpret gender differences in behavior if such differences are not taken into account. Despite the advantages of experimental studies described above, many do not account for how men and women are selected into any given experiment. This implies that observed gender differences could be due to an underlying difference in skill distribution. Our study analyzes gender differences in preferences as manifested in behavioral choices concerning examination strategies among undergraduate university students. Although we may not be able to completely separate between different possible mechanisms behind observed gender differences, we discuss to what degree these mechanisms are consistent with the reported results. We also have rich information on students, including grades in the high school math courses that are a pre-requisite for this undergraduate course and can therefore account for any systematic underlying selection in relevant (for the course) skills by gender.

Based on individual information on 2,121 students who completed the first year microeconomics course during the five semesters from 2006 to 2008, female students are found to be more likely to take one or both of the initial quizzes offered during the semester, all else equal, and are also more likely to receive a seminar bonus. During the actual exam, female quiz-takers are less likely to re-take corresponding quiz questions than male quiz-takers. All three strategies (seminar attendance, quiz taking and quiz re-taking) are correlated with higher final exam scores implying that female students win via their higher seminar attendance and higher propensity to take initial quizzes, but lose due to their lower propensity to re-take quiz questions on the final exam. We find that these results are likely to be due to female risk aversion.

Results presented in this paper complement the existing literature on gender differences in preferences and may shed some light on how pedagogical methods differentially influence education strategies by gender, with potential repercussions for educational outcomes. Note that the decisions students make concerning exam strategies have private consequences only implying that behavioral differences are not driven by gender differences in care and concern for others.

The rest of the paper is organized as follows. In Section 2 we describe the examination procedure and in Section 3 the data and the empirical set-up. Results are reported in Section 4 followed by a discussion in Section 5. The paper is concluded in Section 6.

2 The exam procedure

The exam procedure of the first year microeconomics course at Stockholm University provides a great opportunity to study gender differences in examination behavior. All students have the opportunity to take two voluntary quizzes offered during the first and third quarter of the semester. Quizzes aim to test student knowledge and to motivate students to study for exams at an early stage of the course. Both initial quizzes consist of ten multiple-choice questions with a maximum possible score of ten points per quiz.

Students who take one or both quizzes during the semester have the option of skipping corresponding quiz-like questions on the final exam. Original quiz results are then awarded to the corresponding question on the exam. However, quiz-takers also have the option to re-take the corresponding quiz question on the exam in an attempt to improve the initial score. If eligible students chose to re-take a quiz question on the

exam, the scores received on the exam are final, regardless of whether the student improves his/her score or not. As such, re-taking quizzes on the final exam implies a private risk of lowering initial quiz scores. Students who did not take the initial quizzes are required to answer the corresponding questions on the final exam or receive zero points for that question. Moreover, students with near-perfect seminar attendance during the course are automatically rewarded ten points on the final exam. Students who did not attend seminars have the option of answering an additional question on the exam with a maximum possible score of ten points.

The final exam therefore consists of two ten-point multiple choice questions akin to the two initial quizzes, one ten-point open-ended seminar credit question (not answered by students with near perfect seminar attendance) and seven ten point open-ended questions to be answered by all students. The maximum score possible on the final exam is, therefore, 100 points. Students have a maximum of five hours to complete the final exam. Grading is based on absolute performance only.

This examination procedure provides three ways of testing for gender differences in behavior. The first concerns gender differences in the probability of taking one or both of the voluntary quizzes offered during the course. Taking a voluntary quiz provides an opportunity to improve final exam scores as quiz takers are allowed to redo the corresponding quiz questions on the final exam. The second concerns gender differences in the probability of re-taking quiz questions on the final exam among those students that initially took one or both of the quizzes offered during the course. Students who took at least one quiz know their initial scores, can read and evaluate the corresponding questions on the final exam and can thereafter choose whether or not to attempt an improvement by re-doing the quiz questions. The third concerns gender differences in attending seminars as seminar attendance yields a bonus of ten points on the exam and more time to devote to other questions.

3 Data and empirical setup

3.1 Data

Data stems from individual information on quizzes and exams taken during first year microeconomics courses at the Department of Economics, Stockholm University, from the fall term 2006 to the spring term 2008. In total, 2,349 students were enrolled in this course during this time period and took the final exam. We restrict the analysis to the 2,021 students (47 percent of which are female) with background information on the high school math courses that are a pre-requisite for this course. Students at Stockholm University have the possibility of taking the first year microeconomics course via one of nine different academic programs or by registering independently for the course. In our sample, 58 percent of the female students and 56 percent of the male students are enrolled through programs, the remainder independently.[1]

In the empirical analysis two different samples are used. The first sample consists of all students who participated in the course. With this sample we can study gender differences in seminar attendance and the probability of taking one or both of the initial quizzes offered during the course. The second sample consists of student-quiz observations based on those students who took at least one quiz. With this sample, we can study gender differences in the probability of re-taking a quiz on the final exam.

Sample means on all students are presented in Table 1, by gender. Male and female students have on average the same scores on the final exam (63 points). Sample means also indicate that a significantly greater proportion of female students than male students take one or both of the initial quizzes. In addition, a greater proportion of female students than male students attend seminars. Female students are under-represented at the highest high school math levels and math scores also differ somewhat by gender.

[1] The academic programs are business and administration, accounting, retailing, mathematics and economics, economics and statistics, social planning, economics and political science and social science educators. The largest proportion of students enrolled via an academic program is in a business school program.

Table 1. Sample means individual observations (standard errors in parentheses)

	Female students	Male students	Female-male difference
Birth year	1983.9	1984.1	-0.23
	(0.13)	(0.11)	(0.17)
Total exam score (0-100)	62.9	62.6	0.25
	(0.59)	(0.57)	(0.82)
Wrote at least one quiz	0.944	0.899	0.045***
	(0.007)	(0.009)	(0.012)
Wrote two quizzes	0.835	0.787	0.048***
	(0.012)	(0.012)	(0.017)
Seminar attendance	0.876	0.812	0.063 ***
	(0.010)	(0.012)	(0.016)
Math level (1-4)			
Level 1 (lowest)	0.038	0.024	0.014
	(0.006)	(0.005)	(0.007)
Level 2	0.180	0.118	0.061***
	(0.012)	(0.010)	(0.015)
Level 3	0.526	0.428	0.098***
	(0.016)	(0.015)	(0.022)
Level 4 (highest)	0.256	0.430	-0.174***
	(0.014)	(0.015)	(0.020)
Math score (1-4)			
Fail	0.084	0.110	-0.026*
	(0.009)	(0.009)	(0.013)
Pass (lower level)	0.467	0.535	-0.068***
	(0.016)	(0.015)	(0.022)
Pass (higher level)	0.334	0.244	0.091***
	(0.015)	(0.013)	(0.020)
Pass with distinction	0.114	0.111	0.003
	(0.010)	(0.009)	(0.014)
No. of individuals	996	1,125	2,121

Note: Math scores are averaged for all students regardless of math level.

Sample means based on student-quiz observations are shown in Table 2. Among quiz takers, approximately 28 percent of both female and male quiz-takers re-take a quiz question on the exam.[2] It is difficult to interpret this as indication of no behavioral differences between males and females in re-take propensity. Rational behavior suggests that the probability of re-taking a quiz should be lower for those with higher initial quiz scores since attempting to improve one's score becomes more difficult with higher initial scores, all else equal. As shown in Table 2, female students have lower average initial quiz scores than male students. The distribution of initial quiz scores by

[2] Among students who took both initial quizzes, 25 percent of both male and female students retake one quiz question and 14 percent chose to retake both quiz questions.

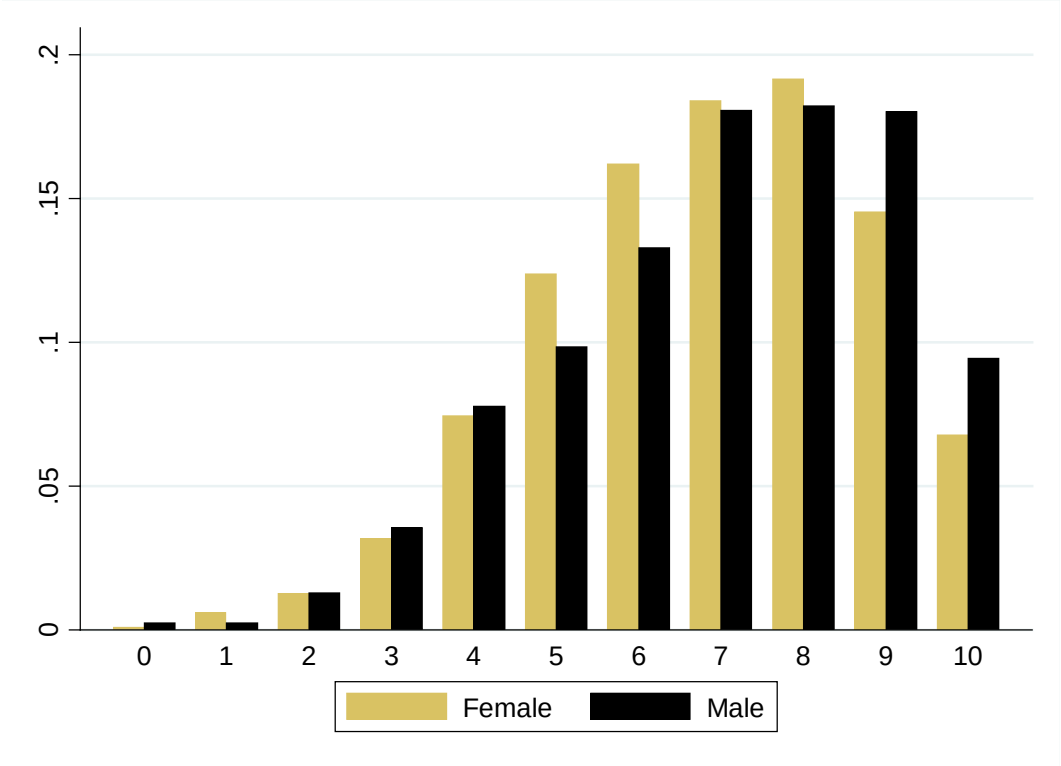
gender is shown in Figure 1. Female students have initial quiz scores of 5 and 6 to a relatively larger extent than male students while male students have initial quiz scores of 9 and 10 to a relatively larger extent than female students. Thus, in the empirical analysis, one should at the very least, compare gender differences in the probability of retaking the quiz conditioning on initial quiz scores.

Table 2. Sample means, quiz observations (standard errors in parentheses)

	Female students	Male students	Female-male difference
Birth year	1983.9	1984.2	0.23
	(0.10)	(0.09)	(0.13)
Total exam score (0-100)	64.5	65.1	0.61
	(0.42)	(0.40)	(0.29)
Initial quiz score	6.81	7.04	-0.23 ***
	(0.046)	(0.046)	(0.066)
Re-take quiz	0.288	0.278	0.009
	(0.011)	(0.010)	(0.015)
Math level (1-4)			
Level 1 (lowest)	0.035	0.025	0.010
	(0.004)	(0.004)	(0.006)
Level 2	0.179	0.123	0.056***
	(0.009)	(0.008)	(0.012)
Level 3	0.527	0.431	0.096***
	(0.012)	(0.011)	(0.016)
Level 4 (highest)	0.259	0.421	-0.162***
	(0.010)	(0.011)	(0.015)
Math score (1-4)			
Fail	0.082	0.104	-0.022**
	(0.007)	(0.007)	(0.010)
Pass (lower level)	0.464	0.541	-0.077***
	(0.012)	(0.011)	(0.016)
Pass (higher level)	0.334	0.245	0.089***
	(0.011)	(0.010)	(0.015)
Pass with distinction	0.120	0.110	0.009
	(0.008)	(0.007)	(0.011)
No. of Quiz Observations	1772	1896	3668
No. of Individuals	940	1,011	1,951

Note: Math scores are averaged for all students regardless of math level.

Figure 1. Distribution of initial quiz scores, by gender



3.2 Empirical setup

To test for systematic gender differences in the propensity to take one or both of the initial quizzes offered during the course, we estimate the following linear probability models:

$$At\ least\ one\ quiz_i = \beta female_i + \mu_s + \gamma_c + \lambda_b + \kappa_g + \kappa_l + \delta_q + \varepsilon_i \qquad (1)$$

$$Two\ quizzes_i = \beta female_i + \mu_s + \gamma_c + \lambda_b + \kappa_g + \kappa_l + \delta_q + \varepsilon_i \qquad (2)$$

where *At least one quiz* is a dummy variable equal to one if student *i* took at least one of the initial voluntary quizzes offered during the course and zero otherwise. Correspondingly, *two quizzes* is a dummy variable equal to one if student *i* took both of the initial quizzes and zero otherwise. The variable of interest is the *female* dummy variable which takes the value one if student *i* is a female and zero otherwise and β measures the average differences in behavior between females and males. The remaining parameters represent a full set of controls for semester effects (μ_s, $s=1,...,5$), course code effects (γ_c, $c=1,...,10$), birth year effects (λ_b, $b=1952,...,1990$), quiz question

effects (δ_q, $q=1,2$), as well as high school math grades (κ_g, $g=1,.,.,4$) and math level (κ_l, $l=1,...,5$) effects.

An analogous regression for the probability of getting a seminar attendance bonus (10 points) on the exam due to near-perfect seminar attendance during the semester of study is also estimated. The equation is specified in the same way as equations (1) and (2) above except that no control for quiz question is included.

To test for gender differences in the propensity to re-take quiz questions on the exam, the following linear probability model, based on student-quiz level data, is estimated:

$$Retake\ quiz_{iq} = \beta female_i + q_{quiz\_score} + \mu_s + \gamma_c + \lambda_b + \kappa_g + \kappa_l + \delta_q + \alpha_a + \varepsilon_{iq} \qquad (3)$$

where *Retake quiz* is a dummy variable equal to one if student *i* re-takes the corresponding quiz question, *q*, on the exam and zero otherwise**.** Over and beyond the controls described above, estimation of re-take propensities includes a full set of controls for initial quiz scores q_{quiz_score} ($quiz\_score=0,\ldots,10$) and seminar attendance, α_a. The latter is important as those without the seminar bonus must answer an additional 10 point question on the exam implying differences in time constraints between these students and those with the bonus which may influence re-take propensities.

The ultimate goal with all models described above is to capture gender differences in behavior that do not stem from any other non-observable gender differences, such as underlying ability. The strategy used is inevitably based on selection on observables. We argue, however, that we control for the most crucial factors possible given the setting. By using information on math grades in high school we can be reasonably certain that our estimated gender differences in behavior do not reflect gender differences in the underlying skills relevant for the course in question. Moreover, entrance into a given program within this economics course is competitive and course code effects should capture any systematic differences in entrance requirements.

4 Results

4.1 Gender differences in the probability of taking initial quizzes

Results based on estimations of equations (1) and (2) are reported in Table 3. In Column 1 the dependent variable is whether the student took at least one quiz and in Column 2 the dependent variable is whether the student took two quizzes. Coefficient estimates show that female students are associated with a 3.8 percentage point higher probability of taking quizzes than male students. These effects correspond to a 4.1 (4.7) percent higher probability for female students to take at least one quiz (both quizzes). Gender differences in the propensity to attend seminars are also estimated and results reported in Column 3 (Table 3). Results indicate that female students are more likely to attend seminars and therefore receive the seminar bonus on the exam than likewise male students.[3]

Table 3. Gender differences in the probability of taking a quiz and attending seminars

	The probability of taking at least one quiz	The probability of taking two quizzes	The probability of getting a seminar attendance bonus
Female	0.038***	0.038**	0.055***
	(0.012)	(0.017)	(0.016)
Take quiz(zes) baseline	0.920	0.810	0.842
Percent effect	0.041	0.047	0.065
No. of observations	2,121	2,121	2,121

Note: Linear probability models on dummy variables indicating whether or not individuals took at least one quiz or both quizzes and if they attended seminars and received seminar bonus. Control for semester (1-5) , course code (1-10), birth year (1952-1990), math level (1-4) and math grade (1-4) including full interaction between math level and math grade.

4.2 Gender differences in the probability of re-taking quizzes on the exam

Results from estimations on the probability to re-take quizzes on the exam, i.e., Equation (3), are reported in Table 4. Results show that female students are significantly less likely than male students to re-take quiz questions on the exam. In percentage terms, female students are almost 10 percent less likely to re-take a quiz on the final

[3] As female and male students have significantly different levels of high school math prior to entering the undergraduate course in Microeconomics (see sample means in Table 1 and Table 2), we explore the sensitivity of reported results in Table 3 to gender differences in math skills. As reported results do not vary in models that control for math levels and math grades and models that also include an interaction term between math levels and grades, we depart from the simpler model with no interaction term. We re-estimate the simple model including instead an interaction between gender and math score/math level. Results remain positive and significant indicating that gender differences in math skills do not explain higher propensities among female students to take one or both quizzes and to attend seminars.

exam than male students, all else equal. Results are unaltered when we allow math levels and math grades to vary by gender (not shown).

Table 4. Gender differences in the probability of re-taking quizzes on the final exam (quiz observations)

	Re-take Propensity
Female	-0.028**
	(0.013)
Re-take (quiz) baseline	0.283
Percent effect	-0.097
No. of observations	3,668

Note: Linear probability models on dummy variables indicating whether or not individuals re-took at least one quiz. Control for quiz question (1-2), seminar attendance, semester (1-5) , course code (1-10), birth year (1952-1990), math level (1-4) and math grade (1-4) including full interaction between math level and math grade and quiz question and initial quiz score including full interaction between quiz question and initial quiz score. Robust standard errors in parentheses.

Average difference can mask differences across the skills distribution. There are a number of studies that suggest that women who choose to compete, especially in male dominated environments, are as competitive as the men in these environments.[4] Estimation of gender differences in re-take propensities across the distribution of initial quiz scores is explored by re-estimating model 3 including an interaction between the female dummy variable and each initial quiz score.[5] Results, depicted in Figure A 1 in the Appendix, suggest that gender differences in re-take propensities are driven by those with an initial quiz score of five. In other words, female students who have an initial quiz score of five are significantly less likely to re-take the quiz on the exam than male students with an initial quiz score of five. Re-take propensities are similar for female and male students with higher initial quiz scores confirming results in previous studies that suggest smaller gender differences in behavior at the high end of the skills distribution.

4.3 Potential costs of gender differences in exam behavior

Before discussing potential explanations behind observed gender differences in examination behavior, let us first look at how these behavioral differences (quiz-taking, seminar attendance and quiz re-taking) correlate with other examination outcomes.

[4] See for example: Master and Meier (1988), Birley (1989), Johnson and Powell (1994), Dwyer *et al*. (2002), Atkinson *et al*. (2003), Datta Gupta *et al*. (2005), Nekby *et al*. (2008) and Garratt *et al*. (2011).

[5] Initial quiz scores of zero to three are grouped together due to a low number of observations in these cells.

The difference between quiz-takers and non-quiz-takers in final exam scores is shown in Table 5. Quiz-taking is associated with higher final exam scores. Estimation controlling for all relevant covariates yields results showing that taking at least one quiz (or two quizzes) is associated with approximately 12 higher points on the final exam in comparison to not taking any quizzes (which corresponds to 17-20 percent higher final exam scores for quiz-takers).

Table 5. The correlation between quiz-taking and final exam scores

	(1)	(2)	(3)	(4)
At least one quiz	12.1***	11.8***		
	(1.61)	(1.87)	--	--
Two quizzes			10.9***	12.0***
	--	--	(1.00)	(1.28)
Female	-1.61	-2.40	-1.60**	0.48
	(0.74)	(3.26)	(0.73)	(1.83
Female* at least one quiz		0.85		
	--	(3.26)	--	--
Female* two quizzes				-2.54
	--	--	--	(1.96)
Final exam baseline		62.7		
Percent effect	19.5	19.0	17.4	19.1
No. of observations		2,121		

Note: OLS estimation of exam scores (0-100) on quiz taking, defined as dummy variables equal to one if an individual has taken at least one quiz, alternatively both quizzes, and zero otherwise. Control for quiz question (1-2), seminar attendance, semester (1-5) , course code (1-10), birth year (1952-1990), math level (1-4) and math grade (1-4) including full interaction between math level and math grade. Robust standard errors in parentheses.

Thus, quiz-takers do better on final exams, either due to a systematically different type of study behavior which voluntary quizzes give rise to (including the possibility of re-taking quiz questions on the final exam) or because quiz-takers are selected among the more ambitious and scholarly students enrolled in the course. As we control for differences in relevant math skills, more weight should perhaps be given to the first explanation. Female quiz-takers who take at least one initial quiz do not appear to differ from likewise male quiz-takers in terms of the quiz premium on final exams as indicated by the insignificant coefficient on the gender-quiz interaction terms (Columns 2). The quiz premium for those who take two initial quizzes appears to differ somewhat between female and male students (Columns 4) but this difference is not significant.

Turning instead to an examination of the correlation between re-taking quizzes on the exam and final exam scores (based on quiz observations), we see from results reported in Table 6 that quiz re-takers are associated with a significantly higher final

quiz scores in comparison with those that abstain from re-taking the quiz. Quiz re-takers are associated with, on average, 1.3 higher points than those that do not re-take the quiz on the final exam. No gender differences in quiz improvement are noted as shown by the insignificant estimated coefficient on the interaction between the female dummy variable and the re-take dummy variable.

Table 6. The correlation between quiz re-taking and final exam scores (quiz observations)

	Quiz score	
Quiz re-take	1.266***	1.242***
	(0.063)	(0.083)
Female	-0.015	-0.028**
	(0.033)	(0.013)
Quiz re-take* female		0.046
	--	(0.112)
	Other exam question scores	
Quiz re-take	6.18***	7.00***
	(0.659)	(0.81)
Female	0.44	0.91
	(0.536)	(0.62)
Quiz re-take* female		-1.63
	--	(1.01)
No. of observations	3,668	3,668

Note: OLS estimation of quiz scores (0-10) and exam scores (0-100) on quiz re-taking, defined as dummy variable equal to one if an individual that has taken at least one of the initial voluntary quizzes and re-taken it on the final exam. All estimation controls for quiz question (1-2), seminar attendance, semester of study (1-5), course code (1-10), birth year (1952-1990), math level (1-4) and math grade (1-4) as well as a full interaction between math level and math grade. Robust standard errors reported in parenthesis.

Re-taking quizzes on the exam implies a time investment. Students must weigh in to what degree they will benefit from re-taking quiz questions given that less time can be invested in the remaining mandatory questions on the exam. Students are allotted a maximum five hours to complete final exams. Abstaining from re-taking quiz questions on the exam may therefore be a rational strategy for students who feel that their time is better invested in answering other final exam questions. If such is the case, we would expect that non re-takers have higher scores on the other exam questions than re-takers, especially for female students. Results in the second panel of Table 6, however, show that re-takers are associated with higher points (a 6-7 point advantage) on the remaining exam questions and that there is no gender difference in this association. These correlations suggest that it is not time constraints driving female quiz takers to abstain from re-taking quizzes on the exam, rather there appears to be spill-over effects on other exam questions implying that the selection of students who prepare for the possibility of

re-taking quizzes on the exam improve their scores over and beyond the improvement on the quiz questions alone. Again no gender differences in this correlation are found.

Taken together, results show that quiz re-takers are associated with higher quiz scores on the exam as well as higher marks on other exam questions. As such, quiz re-taking is also associated with an approximate eight-point advantage in terms of higher total final exam scores.

One way of measuring the implications of observed gender differences in examination behavior is to estimate how final exam scores would be affected if female students behaved like male students. We therefore calculate a Blinder-Oaxaca decomposition based on the following equation using the full sample of students:[6]

$$Final\ Exam\ Score_i = \beta_1 Seminar + \beta_2 Quiz + \beta_3 Re\text{-}take + q_{quiz\_score} + \mu_s + \gamma_c + \lambda_b + \kappa_g + \kappa_l + \delta_q + \varepsilon_i \qquad (4)$$

Table 7. Predicted final exam scores if female students behaved like male students

	Explained and unexplained final exam scores	Predicted female final exam scores based on male characteristics
Female-Male difference in final exam scores	62.9 - 62.6 = 0.24	
Unexplained	0.66	
Explained:	-0.42	63.2
By seminar attendance	0.79	62.2
By taking at least one quiz	0.48	62.5
By re-taking quiz	-0.81	63.1

Note: Re-taking quiz is evaluated at an initial quiz score of five.

In the decomposition, we depart from the female coefficients, i.e. female remuneration in terms of final exam scores for the given productivity characteristics. Of the total gender difference in final exam scores (0.24 points), -0.42 points is explained by differences in characteristics or differences in behavior. In other words, female students have lower attributes or behave in a manner that yields lower scores than male students given similar remuneration for these traits and behaviors. Broken down by different examination behaviors, we see, as expected, that female students gain final exam points by higher seminar attendance (0.79 points) and by a higher propensity to take at least

[6] The dummy for re-take is re-coded to be equal to zero for all students that (1) did not re-take a quiz question on the exam (among quiz-takers) and (2) for those students that did not take either of the original quizzes.

one quiz (0.48 points) but lose by their lower propensity to re-take quizzes on the final exam (-0.81 points). Predicted final exam scores if female students behaved like male students but were remunerated according to a female payment structure are shown in the second column. Clearly, female students would win overall by behaving more like male students (63.2 points instead of 62.9 points), a result that is driven by a higher propensity among male students to re-take quizzes on the exam.[7]

5 Potential explanations for gender differences in examination behavior

There are a number of potential mechanisms behind our observed behavioral differences in examination strategies having to do with, for example, gender differences in confidence, risk-aversion, competitiveness and/or time usage (such as procrastination), which may be difficult to tease out in the type of real world setting examined here. Below we discuss the credibility of some of these mechanisms in explaining observed results.

Seminar attendance and quiz taking are two ways of gaining more time for other questions during the final exam. The seminar bonus automatically yielded the highest possible scores on one of the final exam questions and students that were satisfied with initial quiz scores could refrain from re-taking corresponding quiz questions on the final exam and devote this time to other questions. Taking one or both of the initial quizzes also provided students with a second chance to improve their scores on the final exam. As such, seminar attendance and quiz taking must be considered risk-averse strategies in examination behavior. Our results showing higher propensities among female students to take initial quizzes and attain a seminar bonus are therefore consistent with women being, on average, more risk avert than men.

Re-taking a quiz on the final exam is associated with a private risk of lowering initial quiz scores. Although observed gender differences in re-taking propensities are driven by students with medium to low initial quiz scores, five or below on a ten point scale, there is still a risk involved in attempting to improve initial scores. Our result that female students are less inclined to re-take the quizzes, all else equal, is therefore also consistent with an interpretation that female students are more risk averse than male

[7] Estimation of mean differences in re-take behavior by gender is evaluated at an initial quiz score equal to five.

students. Note that when analyzing gender differences in re-take propensities, we use a selection of students who took the initial quizzes. Since female students took quizzes to a larger extent, and if taking a quiz is associated with a higher degree of risk aversion, it must be the case that we have stronger selection of risk-averse males students compared to female students among initial quiz-takers. As such, it is likely that we underestimate gender differences in risk aversion as measured by re-take propensities. A lower propensity to re-take quiz questions on the exam may also be explained by a female preference for open-ended questions. Female students may prefer to invest more time in the other exam questions than the multiple choice quiz questions.[8]

An alternative interpretation of results is that male students are more confident than female students. Overconfidence is a relative concept. Men could be confident (and correct) in their interpretation of their abilities whereas women are under-confident, or men could be overconfident (i.e., have an incorrect estimation of their ability) whereas women have a correct estimation of their ability. Irrespective of whether male students are overconfident or female under-confident, given that male students have the same ambition as female students to maximize final exam scores, our results may reflect male perceptions of a higher capability to receive high final exam scores without the bonus, in terms of time or scores, of taking an initial quiz and/or attending seminars. Higher confidence, on average, can also explain why male students re-take quizzes to a large extent, all else equal.

One way to explore the degree to which gender differences in examination behavior are due to male overconfidence rather than female risk aversion is to compare initial quiz scores with the quiz scores on the final exam for those that re-take quizzes on the exam. If men are more overconfident there should be a greater spread of male differences in final and initial quiz scores due to re-taking as overconfident male students are more likely to win and lose than female students. If not, then female risk aversion is more likely to be behind this particular difference in behavior. The distribution of the difference in final and initial quiz scores is shown in Figure A 2 in the Appendix. In this figure, we depart from an initial quiz score of five as results showing gender differences in re-take probabilities are driven by students with this

[8] Given a higher preference for open-ended questions among female students, taking one or both of the initial quizzes during the term is still a risk-averse strategy as it allows female students a larger chance to act on this preference during the final exam.

initial quiz score. The variance of the difference between initial and final quiz scores is indeed larger for male students but this is due to the fact that male students are more likely to win compared to female students.[9] This suggests that male overconfidence is not part of the story behind gender differences in re-take behavior.

The exam procedure under study here did not involve any direct rewards that depend on students' relative performance implying that gender differences in competitive behavior are unlikely to be a relevant explanation for presented results. There is one other potential explanation which concerns gender differences in time-use. Male students may be more prone to procrastinate implying that they study less at the beginning of the course and more intensely close to the date of the final exam. Gender differences in procrastination may then explain why male students take initial quizzes to a lesser extent and have lower seminar attendance. According to course evaluations of students who attended the microeconomics course in 2011, a larger fraction of female than male students spend more than 30 hours a week studying for the course. This does not imply that male students procrastinate to a larger extent than female students, but may be consistent with a story where male students spend less time on their studies and, perhaps, distribute this time differently than female students. Higher male procrastination is also consistent with higher male confidence about doing well on final exams. However, procrastination cannot explain why male students, given that they took an initial quiz, re-take these on the final exam to a larger extent than female students.

Taken together, the observed gender differences in examination behavior found in this study seem consistent with conclusions from the experimental literature that women are more risk averse, on average, than men. Studying gender differences in a natural setting allows us to provide tentative conclusions on how gender differences in preferences affect real outcomes. Attending seminars, taking initial quizzes and re-taking quizzes on the final exam are all correlated with higher final exam scores. As such, the risk-averse behavior of female students in terms of higher seminar attendance and taking initial quizzes, seems to be a better strategy than that followed by male

[9] This also means that the average difference in initial and final quiz scores is higher for male students given that we conditioned on an initial quiz score of five. Note that there are no gender differences in the correlation between re-taking a quiz and final quiz score if we do not condition on a given initial quiz score (see Table 6).

students. On the other hand, female students could improve their results on the final exam by behaving more like men when it comes to re-taking quizzes on the final exam.

6 Conclusion

This study has used information on examination behavior among first year micro-economics students at Stockholm University to analyze gender differences in preferences in terms of examination behavior. Students in this course have the possibility of attaining a seminar bonus on the final exam for near-perfect seminar attendance and are given two voluntary initial quizzes during the semester. At the final exam, the scores received on initial quizzes can either be accepted as is, or students can attempt to improve their marks by answering similar quiz questions on the exam. Results suggest that female students are more likely to take initial quizzes and receive a seminar bonus but are less likely to re-take quiz-questions on the final exam. Attending seminars, taking initial quizzes and re-taking quizzes on the final exam are all correlated with higher final exam scores. As such, the risk-averse behavior of female students in terms of higher seminar attendance and taking initial quizzes, seems to be a better strategy than that followed by male students. On the other hand, female students could improve their results on the final exam by behaving more like men when it comes to re-taking quizzes on the final exam.

References

Atkinson, S., Boyce, B., and Melissa, F. (2003), "Do Female Fund Managers Manage Differently?," *Journal of Financial Research*, 26, 1-18.

Bertrand, M. and Hallock, K. (2001), "The Gender Gap in Top Corporate Jobs," *Industrial and Labor Relations review*, 55, 3-21.

Bertrand, M., (2010), "New Perspectives on Gender", *Handbook of Labor Economics*, Volume 4b, Elsevier Ltd., pp.1545-1592.

Birley, S. (1989), "Female Entrepreneurs: Are They Really Any Different?," *Journal of Small Business Management*, 27: 32–37.

Croson, R. and Gneezy, U. (2009), "Gender Differences in Preferences," *Journal of Economic Literature*, 47(2), 448–74.

Datta Gupta, N., Poulsen, A., and Villeval, M-C. (2005), "Male and Female Competitive Behavior: Experimental Evidence," IZA Working Paper no. 1833, Institute for the Study of Labor.

Dwyer, P. D., Gilkeson, J. H. and List, J. A., (2002), "Gender Differences in Revealed Risk Taking: Evidence from Mutual Fund Investors, " *Economics Letters*, 79:151-158.

Garratt, R. J., Weinberger, C. and Johnson, N., (2011), "The State Street Mile: Age and Gender Differnces in Competition Aversion in the Field", *Economic Inquiry*, forthcoming.

Johnson, J. E. V., and Powell, P. L. (1994), "Decision Making, Risk and Gender: Are Managers Diffferent?," *British Journal of Management*, 5(2): 123–38.

Master, R. and Meier, R. (1988), "Sex Differences and Risk Taking Propensity of Entrepreneurs," *Journal of Small Business Management*, 26: 31–35.

Nekby, L., Vahtrik, L., and Skogman Thoursie, P. (2008) "Gender and Self-Selection Into a Competitive Environment: Are Women More Overconfident Than Men?," *Economics Letters*, 100, 405-407.

Appendix

Figure A 1. Female/initial quiz score interaction in the propensity to re-take quizzes on the final exam (quiz observations)

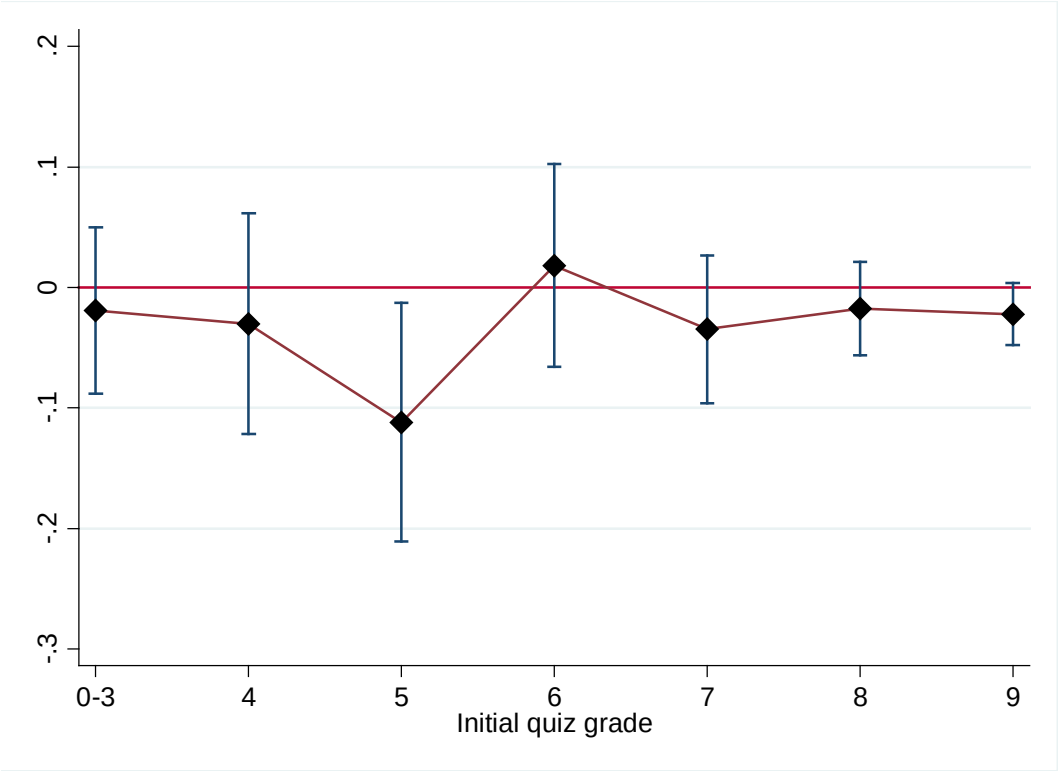
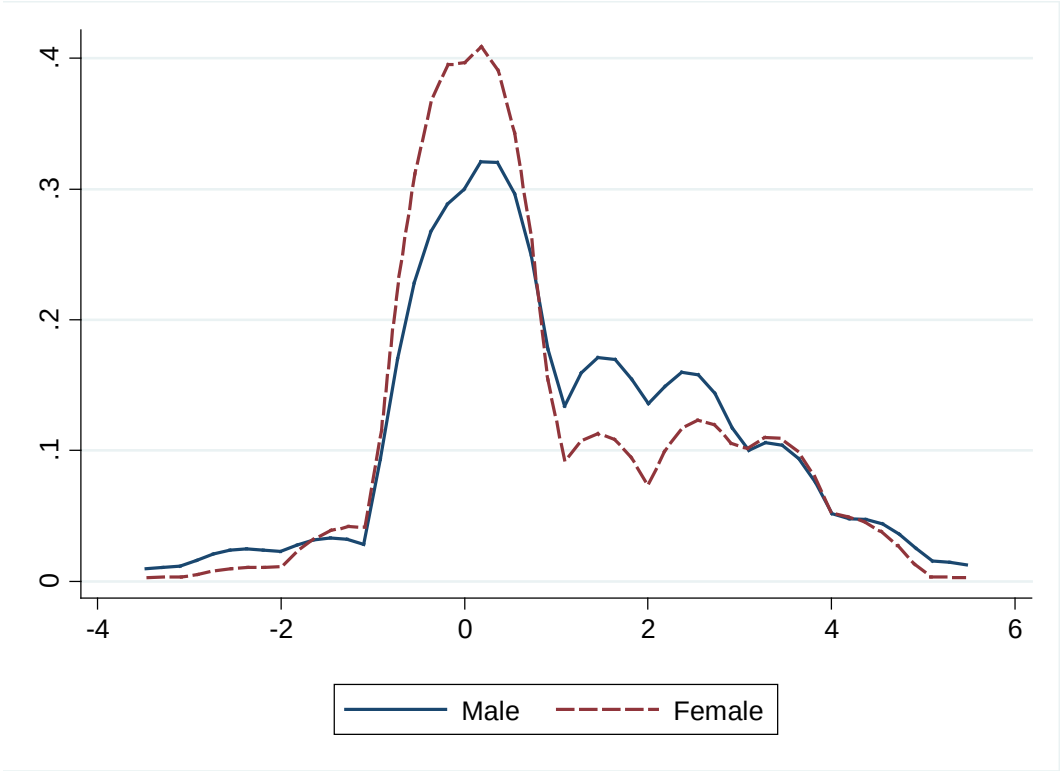


Figure A 2. Distribution of difference in initial and final quiz score (quiz re-takers)



Note: The above distribution is calculated for an initial quiz score of five.

Publication series published by IFAU – latest issues

Rapporter/Reports

2012:20 Ulander-Wänman Carin ”Flexibilitet – en dominerande diskurs i de anställningsvillkor som rör sysselsättningstrygghet för arbetstagare i kommun och landsting”

2012:21 Wikström Christina and Magnus Wikström ”Urval till högre utbildning – Påverkas betygens prediktionsvärde av ålder?

2012:22 Karimi Arizo, Erica Lindahl and Peter Skogman Thoursie ”Effekter av föräldrapenning på arbetsutbud”

2012:23 Sibbmark Kristina ”Arbetsmarknadspolitisk översikt 2011”

2012:24 Liljeberg Linus, Sara Martinson and Jonas Thelander ”Vad innebär det att bli coachad? En utvärdering av jobbcoachningen vid Arbetsförmedlingen”

2012:25 van den Berg Gerard J., Petter Lundborg and Johan Vikström ”De långsiktiga ekonomiska och sociala konsekvenserna av att förlora ett barn”

2012:26 Hägglund Pathric, Per Johansson and Lisa Laun ”Rehabiliteringsgarantin”

2012:27 Engdahl Mattias and Olof Åslund ”Ekonomiska drivkrafter och studieresultat – effekter av Sfi-bonus”

2012:28 Lindgren Karl-Oskar ”Arbetsplatsstorlek och sjukfrånvaro”

Working papers

2012:20 Angelov Nikolay and Arizo Karimi ”Mothers’ income recovery after childbearing”

2012:21 Wikström Christina and Magnus Wikström ”University entrance selection and age at admission”

2012:22 Karimi Arizo, Erica Lindahl and Peter Skogman Thoursie ”Labour supply responses to paid parental leave”

2012:23 van den Berg Gerard J., Petter Lundborg and Johan Vikström ”The economics of grief”

2012:24 Åslund Olof and Mattias Engdahl ”The value of earning for learning: performance bonuses in immigrant language training”

2012:25 Collet Francois and Peter Hedström ”Endogenous tie formation mechanisms in a directed network generated by employee mobility”

2012:26 Lindgren Karl-Oskar ”Workplace size and sickness absence transitions”

2013:1 Nekby Lena, Peter Skogman Thoursie and Lars Vahtrik ”Examination behavior – Gender differences in preferences?”

Dissertation series

2011:1 Hensvik Lena “The effects of markets, managers and peers on worker outcomes”