Available online at www.sciencedirect.com
SCIENCE DIRECT®
Economics Letters 86 (2005) 199–203
economics letters
www.elsevier.com/locate/econbase

# Gender and overconfidence

Claes Bengtsson[a], Mats Persson[b,*], Peter Willenhag[a]

[a]Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden
[b]Institute for International Economic Studies, Stockholm University, SE-106 91 Stockholm, Sweden

Received 7 February 2004; received in revised form 29 June 2004; accepted 20 July 2004
Available online 10 November 2004

**Abstract**

Do males differ from females in terms of self-confidence? Using a large set of exam data from Stockholm University, we find that exam behavior is gender-specific: male students are more inclined than female students to aim for a higher grade.


*Keywords:* Overconfidence; Gender differences; Exam behavior; Exam results

*JEL classification:* A2; J2; J16

## 1. Introduction

There is a small but growing literature indicating that men are more overconfident than women.[1] While highly relevant to our understanding of social phenomena, studies of the link between gender and overconfidence are hampered by a lack of good data. The purpose of the present paper is to exploit a new data source to shed light on the issue.

Stockholm University has used a particular design of the written exam for the first-year courses in economics. There are three grades: Very Good (VG), Pass (P), and Fail (F). The exam consists of four questions, and in order to get the grade P on the entire exam, you need a P on each one of these four

* Corresponding author. Tel.: +46 8 162224; fax: +46 8 161443.
*E-mail address:* mp@iies.su.se (M. Persson).
[1] Cf. Barber and Odean (2001) and Correll (2001).

0165-1765/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2004.07.012

questions. For the student aiming for a VG on the exam, however, there is a fifth question. In order to get a VG, you need firstly a VG on each of the first four questions and secondly a satisfactory answer to Question 5. If one has a mere P on one or more of the first four questions, one can never get a VG on the exam, regardless of whether or not one gives a good answer to the fifth question. At the time when one decides whether or not to answer Question 5, one does not yet know how good the answers to Questions 1–4 are. Thus self-assessment enters the picture. Furthermore, if one is satisfied with a mere P on the exam, one has no incentive to answer the fifth question—that would only be a waste of time.

This exam structure provides an interesting opportunity for studying gender differences in self-assessment among the students. The issue we analyze is whether female students are less prone to answer the fifth question than are male students. For this purpose, we use historical exam data consisting of enough observations so that reliable significance tests can be obtained.

We find, in fact, that there is such a difference in exam behavior between male and female students; men tend to be more inclined to answer Question 5 than women. But this is not the entire story. There are interesting patterns within each gender group. Dividing the students into "good" and "mediocre" ones (depending on their results on Questions 1–4), we find that the gender difference in self-assessment was significant in both groups, but more so in the mediocre one. Furthermore, we find an age effect: the gender difference in self-assessment is limited to the younger students.

## 2. The empirical picture

### *2.1. Exam results*

We have used data from five Microeconomics I exams, from the Fall Term 2001 through the Spring Term 2004. In total, there were 2217 students who took those exams, of which 1102 (49.7%) were female and 1115 (50.3%) were male.

Among the female students, 78.8% (869 out of 1102) passed the exam, i.e. received the grade P or VG. Among the male students, 76.5% (853 out of 1115) passed the exam. Is 78.8% significantly different from 76.5%?

To answer this question, we use the test statistic

$$t = \frac{\hat{\mu}_{\mathrm{F}} - \hat{\mu}_{\mathrm{M}}}{\sqrt{\frac{\hat{\sigma}_{\mathrm{F}}^2}{n_{\mathrm{F}}-1} + \frac{\hat{\sigma}_{\mathrm{M}}^2}{n_{\mathrm{M}}-1}}},$$

where $\hat{\mu}_j \equiv n_j^i/n_j$, $j$=F, M, are the sample fractions among female and male students, respectively, who passed the exam, and where $\hat{\sigma}_j^2 \equiv \hat{\mu}_j(1-\hat{\mu}_j)$, $j$=F, M, are the sample variances. This statistic is $t$ distributed with $n_{\mathrm{F}}+n_{\mathrm{M}}-2df$.[2]

The null hypothesis is now that the percentage of female students who pass the exam is the same as the percentage of male students who pass the exam. Substituting $n_{\mathrm{F}}$=1102, $n_{\mathrm{F}}^i$=869, $n_{\mathrm{M}}$=1115 and $n_{\mathrm{M}}^i$=853, where $i$ indicates "passed the exam", into the equation above, we obtain $t$=1.3312, which is significant at the 90% confidence level. Thus the observed frequencies 78.8% and 76.5% are sufficiently

[2] See Hogg and Tanis (1997, pp. 305–309).

different for us to be able to reject the null hypothesis that male students pass the exam as often as female students do. This result is in line with earlier evidence that females seem to perform slightly better than males at school.[3]

What about the VG grade? There were 130 (11.8%) female students with a VG, and 180 (16.1%) male students; the difference in proportions achieving a VG are statistically significant at the 99% level ($t=-2.9574$).

There are thus striking differences between male and female students in terms of outcome: female students are slightly better at passing the exam, but male students are much better at getting the highest grade. Whether this pattern is due to differences in innate intellectual capacity, study habits, or some other factor, is an open question that deserves further investigation. In this paper, we confine the analysis to investigating whether there are any gender differences in exam behavior.

### 2.2. Exam behavior

Among the female students, there were 506 who were qualified to answer Question 5, i.e. who scored a VG on each of the Questions 1–4. Of these, 424 (83.8%) answered Question 5. Among the male students, 480 were qualified, of whom 418 (87.1%) answered Question 5. Applying the above test statistic, we find that qualified females are less prone to answer Question 5 than qualified males. The difference is significant at the 90% level ($t=-1.4653$).

There were 596 unqualified females (i.e. those who had not received a VG on each one of the Questions 1–4). Of these, 248 (41.6%) still tried to answer Question 5. Among the 635 unqualified males, 309 (48.6%) were similarly optimistic. The difference is statistically significant at the 99% level ($t=-2.4891$). We can thus reject the null hypothesis that females are as prone as males to go for Question 5. This pattern holds for both qualified and unqualified students, but it is more pronounced among the unqualified ones.

There are other patterns in the data. For instance, we divided the unqualified students into two groups: those who are "mediocre" (having at least the grade P, but not the grade VG, on the Questions 1–4) and those who are "bad" (having an F on at least one of the Questions 1–4, i.e. failing the exam). In both groups, males were significantly more prone to answer Question 5.

To see whether there was an age effect involved, we divided the students into two age groups of roughly equal size: those between the ages 18 and 22 (inclusive), and those between 26 and 62 (inclusive). We then tested the following null hypothesis: Within each age group, the propensity to answer Question 5 is the same for men and women.

Running the same tests as above for the younger group of qualified men and women, we obtained $t=-1.0280$, i.e. there was no significant difference in behavior between younger qualified men and women. Running the test for the younger group of unqualified men and women, we obtained $t=-1.5317$, i.e. there was a significant difference (at the 90% level) between younger unqualified men and women. In the older group, however, we did not observe any significant difference between men and women. Breaking the data down into smaller and smaller subgroups, however, may make it difficult to discern significant differences due to the correspondingly fewer degrees of freedom.

[3] Cf. McNabb et al. (2002).

Summing up, there is a clear difference in self-assessment between men and women. With the qualification for the problem of degrees of freedom, this difference seems to be more pronounced in the younger age group.[4]

## 3. Some explanations

There are clear gender differences both in exam results and exam behavior. An economist would attribute this difference in behavior to constraints, preferences, and/or technology.

The most immediate constraint is the budget constraint, which in this case is a time constraint: answering the optional fifth question takes time. If women's time were more valuable than men's time, it would be rational for the former not to waste time on Question 5. Collecting data on how much time each student spent answering the questions, we tested whether females seemed to be more in a hurry than males. This turned out not to be the case: there were no significant gender differences in this respect.[5]

Another aspect of the time budget constraint could be that a student might be inclined to skip Question 5 in order to have more time for Questions 1–4. To test this, we divide all students into two groups: those who answered all five questions, and those who answered only four of them. The result was clear: while the students who answered Question 5 spent on average 30 min more on the exam than those who did not answer all five questions, there was no significant gender difference in this respect.

The distinction between preferences and informal institutional constraints is of course ambiguous. A difference in preferences that may be the result of social norms could be, for instance, optimising versus satisficing behavior. If men were optimizing, while women were satisficing, this would be consistent with the observed differences in exam behavior. Our data does not, however, allow for a sharper test of this hypothesis.

Another preference trait could be that of risk aversion. This does not, however, provide an explanation to the observed gender differences. Question 5 could be regarded as a lottery ticket. If such a lottery ticket were costly, more risk-averse persons would be less inclined to buy it. But if the ticket were offered for free, everybody would take it, regardless of his or her degree of risk aversion. Now, if the time cost of answering Question 5 is negligible, failing to answer the question would be like not picking up a free lottery ticket. Differences in the propensity to pick up such a ticket cannot be explained by differences in risk aversion.

Differences in technology would here mean differences in ability. One possibility is that unqualified men who choose to answer Question 5 really are marginally better than the females—but this difference is not evident in our coarse grading system. We tested this hypothesis, without finding any support for such a difference.[6]

Another possibility that has to do with technology is that men may be more uncertain about the outcome of the exam. It would then be a rational strategy for them to take a chance with Question 5, provided that their time is not too valuable.

---

[4] Another interesting question is whether there is an age effect within each gender group. It turned out that this is the case at the 99% level. In our sample, younger women were more prone to answer Question 5 than older women ($t$=2.52), and younger men were much more prone to answer Question 5 than older men ($t$=3.04).

[5] See Bengtsson et al. (2004) for details.

[6] See Bengtsson et al. (2004) for details.

It is hard to find out the variances of people's subjective probability distributions, but we can at least test whether the actual exam scores of men have a variance different from that of women. Using different ways of assigning numerical values to the different questions, we found a robust result: The variance over the results on Questions 1–4 among the males was significantly higher than among the females. There are thus a number of potential explanations, in terms of constraints, preferences and technology that are consistent with the data. To discriminate between these explanations would require quite another data set, a data set that is presently not available.

## References

Barber, B.M., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Quarterly Journal of Economics 116, 261–292.

Bengtsson, C., Willenhag, P., Persson, M., 2004. Gender and overconfidence. Stockholm University, Institute for International Economic Studies, Working Paper No. 730.

Correll, S.J., 2001. Gender and the career choice process: the role of biased self-assessments. American Journal of Sociology 106, 1691–1730.

Hogg, R.V., Tanis, E.A., 1997. Probability and Statistical Inference. Fifth ed. Prentice-Hall, Upper Saddle River, NJ.

McNabb, R., Pal, S., Sloane, P., 2002. Gender differences in educational attainment: the case of university students in England and Wales. Economica 69, 481–503.