



NOTA DI LAVORO

95.2009

**Gender Differences in
Output Quality and Quantity
under Competition and Time
Constraints: Evidence from a
Pilot Study**

By **Olga Shurchkov**, Department of
Economics, Wellesley College

GLOBAL CHALLENGES Series

Editor: Gianmarco I.P. Ottaviano

Gender Differences in Output Quality and Quantity under Competition and Time Constraints: Evidence from a Pilot Study

By Olga Shurchkov, Department of Economics, Wellesley College

Summary

Gender gaps in income and level of position in the workplace are widespread. One explanation for this inequality is that the genders perform differently under competitive conditions, as previous experimental studies have found a significant gender gap in competitive tasks that are perceived to favor men. In this paper, we use a verbal task that is perceived to favor women and find no gender difference under competition per se. We also reject the hypothesis that a stereotype threat explains the inability of women to improve performance under competition. We propose an alternative explanation for gender inequality: namely, that women and men respond differently to time pressure. With reduced time pressure, competition in verbal tasks greatly increases the performance of women, such that women significantly outperform men. This effect appears largely due to the fact that extra time in a competition improves the quality of women's work, leading them to make fewer mistakes. On the other hand, men use this extra time to increase the quantity of work, which results in a greater number of mistakes.

Keywords: Gender Differences, Competition, Effects of Time Pressure

JEL Classification: C9, J16, J71

I am especially grateful to George-Marios Angeletos, for his guidance and support. I would also like to thank, Rachael Croson, Ernst Fehr, Robert Gazzale, Uri Gneezy, Lise Vesterlund and the participants of the 2008 International ESA Conference and the Wellesley College working seminar series for insightful comments. I acknowledge the generous financial support provided by the Shultz Fund (MIT Economics Department) and the Wellesley College Faculty Award. I would also like to thank the Harvard Business School Computer Lab for Experimental Research for providing me with the laboratory space and the access to the subject pool for this study. All remaining errors are my own.

This paper was one of the three winners of the first edition of the FEEM Award (2009), a prize organized jointly with the European Economic Association (EEA), aimed at rewarding the best research by young economists.

Address for correspondence:

Olga Shurchkov
Department of Economics
Wellesley College
106 Central St.
Wellesley, MA 02481
E- mail: oshurchk@wellesley.edu

Gender Differences in Output Quality and Quantity under Competition and Time Constraints: Evidence from a Pilot Study*

Olga Shurchkov[†]

Current Version: February 2009

First Version: February 2008

Abstract

Gender gaps in income and level of position in the workplace are widespread. One explanation for this inequality is that the genders perform differently under competitive conditions, as previous experimental studies have found a significant gender gap in competitive tasks that are perceived to favor men. In this paper, we use a verbal task that is perceived to favor women and find no gender difference under competition per se. We also reject the hypothesis that a “stereotype threat” explains the inability of women to improve performance under competition. We propose an alternative explanation for gender inequality: namely, that women and men respond differently to time pressure. With reduced time pressure, competition in verbal tasks greatly increases the performance of women, such that women significantly outperform men. This effect appears largely due to the fact that extra time in a competition improves the quality of women’s work, leading them to make fewer mistakes. On the other hand, men use this extra time to increase the quantity of work, which results in a greater number of mistakes.

Keywords: Gender Differences, Competition, Effects of Time Pressure

JEL Classification: C9, J16, J71

*I am especially grateful to George-Marios Angeletos, for his guidance and support. I would also like to thank, Rachael Croson, Ernst Fehr, Robert Gazzale, Uri Gneezy, Lise Vesterlund and the participants of the 2008 International ESA Conference and the Wellesley College working seminar series for insightful comments. I acknowledge the generous financial support provided by the Shultz Fund (MIT Economics Department) and the Wellesley College Faculty Award. I would also like to thank the Harvard Business School Computer Lab for Experimental Research for providing me with the laboratory space and the access to the subject pool for this study. All remaining errors are my own.

[†]Department of Economics, Wellesley College, 106 Central St., Wellesley, MA 02481.
e-mail: oshurchk@wellesley.edu.

1 Introduction

The study of gender differences in performance has a long history in the field of economics. Despite a recent policy push toward the equalization of men and women in the workplace and in society, considerable inequality persists, especially in high profile jobs. The possible explanations put forth in the literature can be sorted into three categories. The first explanation relies on gender differences in skills and preferences that lead to occupational self-selection (Polachek 1981; Macpherson and Hirsch 1995). The second explanation points to discrimination in the workplace which results in differential treatment of men and women of identical abilities and preferences (Black and Strahan 2001). The final and most recent class of explanations for the gender gap rests on the experimental evidence that women may be less effective than men in certain competitive environments (Gneezy, Niederle, and Rustichini 2003; Niederle and Vesterlund 2007). In this study, we would like to explore the latter explanation further and ask whether the nature of the competitive task at hand influences the outcome.

Society generally perceives men to be better than women at following directions and reading maps, while women supposedly tend to follow landmarks when driving (Bhattacharya 2005). When it comes to solving mazes, men are found to be overwhelmingly superior to women (Pease and Pease 2000, p. 107). Similarly, men are perceived to have higher math abilities relative to women, while women are perceived to have superior verbal skills. In particular, Pajares and Valiante (2001) note that differences in achievement of middle school students lie in the stereotyped beliefs about gender differences rather than gender itself. Girls report stronger motivation and confidence in writing and receive higher grades in language arts. Boys report stronger performance-approach goals (Pajares and Valiante 2001).

Previous experimental studies on gender differences under competition have focused primarily on tasks that are typically perceived to be better suited for men (Gneezy, Niederle, and Rustichini 2003; Niederle and Vesterlund 2007). In fact, these studies cite the so-called “stereotype threat” as a possible explanation for why women tend to “shy away from competition” in their experiments. The idea of a stereotype threat first appeared in the field of psychology. It describes the fear that certain behavior would confirm an existing stereotype of a group with which one identifies (Steele and Aronson 1995; Spencer, Steele, and Quinn 1999). Presumably, competition against men can bring out this stereotype threat in women, hindering their performance in historically male-dominated tasks.

Both Gneezy, Niederle, and Rustichini (2003) and Niederle and Vesterlund (2007) motivate their experimental research by the fact that gender gaps in income and social position are widespread. Bertrand and Hallock (2001) document this fact by gathering data on the five highest-paid executives of a large group of U.S. firms over the period of 1992–1997, where they find that only 2.5 percent of the executives in the sample are women.¹ Several authors have argued that this inequality is due to the innate inability of women to compete (see Baron-Cohen (2003), Lawrence (2006), and the citations in Barres (2006)). Gneezy, Niederle, and Rustichini (2003) and Niederle and Vesterlund (2007) argue that, in math-

¹A similar underrepresentation of women is found among CEOs at Fortune 500 companies (<http://money.cnn.com/magazines/fortune/fortune500/womenceos>), tenured faculty at leading research institutions (MIT 1999), and or top surgeons in New York City according to New York magazine (<http://nymag.com/bestdoctors/>).

emational tasks, women fail to compete against men, but not against other women. The authors claim that these results help explain gender inequality in the labor force. However, this conclusion deserves further thought, because the workplace is a complex environment. During the competition to reach the top in almost all spheres of business and politics, one's success is rarely measured solely based on performance in mathematical tasks that end in a matter of minutes. Women and men compete also on the basis of verbal and communication abilities, which are *not* associated with a stereotype threat against women. On the contrary, the aforementioned studies suggest that, in verbal tasks, the stereotype threat might actually negatively affect men, not women.

This study uses a verbal task in a competitive environment in order to shed light on the following two hypotheses. The first hypothesis is that competition hinders women's performance relative to the performance of men in *any* environment. In order to test this hypothesis, we run a controlled experiment where groups of men and women solve "Word-in-a-Word" puzzles. In the benchmark treatment, the subjects are paid according to their own performance in the task. Each group member is paid per valid word found within a larger word over a period of several minutes. We do not find statistically significant differences between men and women in this treatment.

To study the effects of competition, we use a tournament, where only the participant with the highest score is paid proportionally to his or her output. In contrast to previous studies which use tasks that stereotypically favor men, we find no significant gender differences in performance with the competitive payment scheme, which leads us to reject the first hypothesis.

The rejection of the first hypothesis leads us to consider the second hypothesis: gender differences stem not from competition *per se*, but rather from the effect of stereotype threat made salient in a competitive setting. In other words, stereotype threat has a symmetric effect on both genders. If the second hypothesis holds, just as women's performance suffers relative to men's when the competitive task is of a mathematical nature, symmetrically, men's performance should also suffer relative to women's when the competitive task is verbal. This can be explained by a rational inclination for the stereotyped-against group to compete less due to the presence of seemingly superior competitors (this stereotype-threat explanation is proposed, for example, by Gneezy, Niederle, and Rustichini 2003).

We observe that neither women nor men increase their performance in competition with verbal tasks. First, it seems that competition affects men in a manner that is consistent with the second hypothesis: they compete more in tasks that stereotypically favor men (math, mazes, etc.) and not in tasks that stereotypically favor women (verbal tasks). However, a true test of the hypothesis involves comparing the performance of men under piece-rate and under tournament in a single-sex environment which eliminates any potential effect of stereotype threat. We find that competition does not improve the performance of men in a single-sex environment, which implies that the second hypothesis can be rejected.

There is another potential reason for women not outperforming men in a tournament, even with tasks that favor them. Note that in all previous experiments, including those described in this paper so far, men and women have been solving various tasks under intense time pressure. Deadlines in the workplace, though typically strict, allow the workers an opportunity to think deeply about their decisions. Time constraints are rarely as pressing as they appear in previous laboratory experiments on the topic. At the same time, recent

evidence by Paserman (2008) leads us to believe that pressure has a larger detrimental effect on women’s performance under competition as compared to men’s performance.

In order to shed light on the impact of time pressure in competitive verbal tasks, we run the piece-rate and tournament treatments giving the subjects ample time to find as many words as they possibly can within a larger word. We also now provide them with an opportunity to give up before the time runs out, which allows us to test whether men and women have different attitudes regarding effort in these tasks.

Under this scenario, we find that extra time does not create a large difference between male and female performance in the piece rate condition. However, extended time in a competitive environment greatly increases the performance of women – so much so that they significantly outperform the men. A large portion of women’s improvement over the men comes from the quality of output they produce. Women seem to use the extra time to ensure that their words are correct, which results in a smaller relative number of mistakes. On the other hand, men tend to enter more invalid words, which results in a higher frequency of mistakes. In addition, men also give up slightly more than women in the competitive round. This evidence is consistent with the literature in evolutionary biology that suggests that men, as “hunters,” tend to have lower attention spans. On the other hand, women, as “gatherers,” tend to pay attention to detail and can stay focused on a singular task for a prolonged period of time. For example, men often flick through TV channels and do not have the patience to watch commercials, while women are not as averse to sitting through the boring breaks (Sullivan 2001, Pease and Pease 2000). Similarly, our findings are consistent with the notion that higher levels of the hormone testosterone in men are associated with a lower attention span (Sullivan 2001). This study therefore supports the research that “shows that we are more products of our biology than the victims of social stereotypes” (Pease and Pease 2000).

Related Literature. Gneezy, Niederle, and Rustichini (2003) test whether men and women differ in their propensity to perform in competitive environments using an experiment where groups of three men and three women perform the task of solving computerized mazes. In the piece-rate treatment, men and women do not differ significantly in terms of performance. Under competition, men significantly outperform the women, which leads to a large gender gap in performance. This gap disappears in a single-sex environment: women’s scores increase significantly in a tournament against other women. This finding is supported by a recent field study conducted by Gneezy, Leonard, and List (2008), who compare the effects of competition on gender performance in two distinct societies: a patriarchal society and a matrilineal society. Men compete at about twice the rate as the women in the former, while the opposite is true in the latter.² Gneezy and Rustichini (2004) obtain similar results as Gneezy, Niederle, and Rustichini (2003), when they analyze the performance of young boys and girls in a race over a short distance. Note that, just as with reading maps or following directions, men can be perceived to be more skilled at solving mazes or running. The benchmark treatment shows that this perception is most likely false. However, in a competitive setting, the social perceptions become more salient, which may inhibit female performance.

²Note that Gneezy, Leonard, and List (2008) use a task of tossing a ball into a bucket placed 3 meters away from the subject, which is a task that is likely to be perceived as better-suited for men.

Niederle and Vesterlund (2007) conduct a laboratory experiment where men and women add up five two-digit numbers with and without competition. Although there are no gender differences in performance under either of these compensation schemes, there is a substantial gender difference when participants subsequently choose the scheme they want to apply to their next performance. Twice as many men as women choose the tournament over the piece rate. This gender gap in tournament entry is not explained by performance either before or after the entry decision. However, as in the previous experiment, the task of adding up numbers, though very simple, is of mathematical nature and therefore perceived to be biased toward men. Again, the piece-rate treatment demonstrates the falsehood of this perception, since there is no difference in performance between men and women. On the other hand, the tournament treatment forces subjects to think more in terms of social norms and points out the idea that women are not supposed to be good at math.

Field data has also been used to provide evidence for a gender gap in performance in competitive settings. Paserman (2008) uses data from nine tennis Grand Slam tournaments played between 2005 and 2007 to assess whether men and women respond differently to competitive pressure in a setting with large monetary rewards. The author’s detailed point-by-point analysis reveals that, relative to men, women are substantially more likely to make unforced errors at crucial junctures of the match.

In all of the above studies, differences in performance between men and women may be explained by the salience of stereotype threat brought about by competition. The importance of non-gender stereotype threat for performance is documented in the broader literature. For example, Hoff and Pandey (2004) conduct a series of field experiments where low-caste male junior high school student volunteers in rural India performed the task of solving mazes under economic incentives. The authors find no caste differences in performance when caste is not publicly revealed, but making caste salient created a large and robust caste gap. Furthermore, when the link between performance and payoff was purely mechanical, the caste gap disappeared.

This paper is the first to document the effects of competition in a non-mathematical task on gender performance and to show that stereotype threat is not the predominant reason for gender differences. In addition, this paper demonstrates that competition does not hinder women’s performance per se. In fact, women outperform men in a tournament when time pressure is reduced. Another contribution of this study is the finding that, in a low-pressure tournament, women benefit from an increase in both the quality and the quantity of their work. On the other hand, men’s quality of work suffers due to an increase the quantity of output.

The rest of the paper is organized as follows. Section 2 describes the experimental design, procedures, and treatments. Section 3 presents the results of the data analysis. Section 4 concludes and discusses potential future research in this area.

2 Overview of the Experiment

In order to gauge the effects of different payment schemes on performance of men and women, we keep with the pervious literature and conduct a laboratory experiment in which the subjects solve a real verbal task.

2.1 The Task

The subjects were told in the instructions that their objective in this experiment was to work on a series of Word-in-a-Word puzzles.³ In particular, in each round, subjects would have three or fifteen minutes (depending on the treatment) to find as many sub-words that can be formed out of the letters of a big word as possible. In order to control for the level of difficulty across the various treatments within the same experiment, we picked puzzles that contain a similar number of sub-words and that have a similar maximum possible score. At the beginning of the experiment, each participant was given three minutes to solve a practice Word-in-a-Word puzzle in order to get familiar with the task.

The following are the general rules for all the puzzles.

- The sub-words must be 4-letters long or longer.
- Acceptable characters are the letters A-Z only. Any other symbol like a number or another symbol is automatically discarded.
- Proper nouns are not allowed.
- Plurals are allowed.
- Each letter in the puzzle word can only be used once within each new sub-word.

The following are the scoring rules for all the puzzles.

- Valid words add $(N - 3)$ points to the score, where N is the total number of letters in the word.
- Invalid words subtract $(N - 3)$ points from the score, where N is the total number of letters in the word.
- If the word is a duplicate, no points are subtracted and no points are added.
- If the word is too short, 1 point is subtracted from the total score (regardless of whether the word is 1, 2, or 3 letters long).

The time (three minutes in part 1 and fifteen minutes in part 2) ran out automatically. In the first part of the experiment, the subjects did not have the option of finishing the task earlier than three minutes. Once the time ran out, the subjects were presented with their score (in points) and the maximum possible score in a given puzzle. Since the program

³For a complete set of instructions, see Appendix A. The puzzles were computerized using the programming language Python and were based on the games provided by the website www.wordplays.com.

records the scores automatically, the subjects did not need to keep track of their winnings from round to round.⁴

Finally, changing the puzzle from round to round may prove problematic if the difficulty of the task varies dramatically. If the words were not similar, the results would not be directly comparable across treatments.⁵ Our main strategy is to carefully choose puzzles with the number of correct sub-words and the maximum number of points that is as similar as possible in all the treatments. In particular, the range for the number of sub-words is 77-85, and the range for the maximum possible number of points is 132-137. We also check how difficult it is to find the sub-words in any given puzzle. In particular, we count the number of permutations of letters needed to arrive at any one sub-word (the complexity factor). In particular, the average complexity factors for each of the big words (i.e., puzzles) are 2.51 for *carriageway*, 2.60 for *ordination*, 2.39 for *memorable*, 2.12 for *allopathy*, and 2.45 for *equitable*. The practice word is *infusate*.⁶

2.2 The Procedure

The experiment was conducted at the Computer Lab for Experimental Research (CLER) at Harvard Business School (HBS). All sessions were held in February-March of 2008. The total number of individual participants was 76 people (16 groups of two men and two women and three groups of all men).⁷ Most subjects were students at Harvard University (undergraduates and graduates), although students from other Boston-area universities, such as MIT and Boston University, also participated. Because all these institutions are highly competitive in terms of academics, we do not expect women or men to be particularly intimidated by competition against the opposite gender.

CLER recruits subjects via an online registration procedure. Subjects first register for the CLER subject pool. Then, they can sign up for certain studies of their choosing. At any point, a subject could remove him- or herself from the study if he or she is unable to attend. When the subjects arrived at the lab, we separated them into groups of four. Even though gender was not emphasized at any point during the study and explicit communication was

⁴Gneezy, Niederle, and Rustichini (2003) leave it to the subjects to record the number of correctly solved mazes, because they use the internet for the experiment. Although the authors claim to have been monitoring the subjects, in order to ensure that they did not lie about their performance, it is possible that some of them were left unwatched at some points during the task. If men tend to lie and overestimate their performance more in the tournament than in the piece-rate treatment, then this would potentially distort the results. That is the reason for why we did not use online word puzzles for this experiment, but rather programmed our own version of the game.

⁵This problem does not arise to the same extent in studies that use mathematical tasks. For example, Gneezy, Niederle, and Rustichini (2003) simply restrict their pool of mazes to a certain level of difficulty. Because every word is different, it is difficult to find truly identical tasks.

⁶Since the practice round is never used for any of the analysis, we choose an easier word (with 183 sub-words and 342 maximum number of points).

⁷We do not use the data generated by two subjects. One of the subjects (a woman) closed the puzzle program by accident and had to redo the task, which may have increased performance artificially. Another subject's data for the competition round are missing since they were mistakenly erased when the results were compiled for final payment.

not allowed, the subjects were encouraged to look around to see the composition of their group.

Upon entering the lab, the subjects were first asked to read through and sign informed consent forms for non-biomedical research.⁸ Paper copies of the instructions⁹ were distributed to the participants prior to the beginning of the experiment. In the instructions, the subjects could answer several control questions in order to familiarize themselves with the payment schemes and the experimental procedures. All questions were answered in private by the experimenter. At the end of the experiment, each participant filled out a brief questionnaire.¹⁰ The questionnaire asked the subjects some standard demographic questions and inquired about their strategies and beliefs throughout the experiment. At the very end, each subject was paid in cash a show-up fee equal to 10 US \$ and his or her earnings over the course of the session. Final income of each subject was first given in points and then converted to US \$ at the different rates according to the payment scheme detailed below. Average income (including the show-up fee) was \$17.84 across all groups and maximum income (including the show-up fee) was \$46.05. The approximate average duration of the sessions was 1 hour 10 minutes.

2.3 High Time Pressure Treatments

Our first goal is to establish whether men and women exhibit different levels of performance under competition in verbal tasks. As a benchmark measure of performance in these tasks, we use a non-competitive piece-rate payment scheme. To test whether there is a gender-specific effect of competition on performance, we introduce a competitive treatment in the following round. Note that the expected payoff in this competitive treatment is set to be identical to the expected payoff in the noncompetitive treatment.¹¹ However, payment is now uncertain. In order to disentangle the effects of competition and uncertainty on gender differences in performance, we also conduct a “random winner” treatment. In this treatment, payment is uncertain, yet independent of the performance of others.

2.3.1 Piece Rate Treatment

In this benchmark (non-competitive) treatment, the subjects have three minutes to solve one word puzzle (puzzle word: *carriageway*). Each subject receives 10 cents for every point earned in this round. No winner is announced, and everyone earns income according to one’s own performance.

2.3.2 Competitive Treatment

Once again, each subject is presented one word puzzle (puzzle word: *ordination*). However, the scoring is now different. The total score in this round is compared to the scores of the

⁸Copies of the informed consent forms are available upon request.

⁹Full copies of the instructions can be found in Appendix A. After each round, the subjects were reminded about the payoff scheme in the subsequent round.

¹⁰A copy of the questionnaire is available in Appendix A.

¹¹The expected payoff is equal to $\frac{1}{4}(Y_1+Y_2+Y_3+Y_4)$, where Y is the number of points of each group member i .

other three members in the group. The person with the highest score (“the winner”) in this round then receives 40 cents for every point earned. The other three members of the group receive 0 points. In case of a tie, the winner is determined randomly out of the top performers.

2.3.3 Random Winner Treatment

In this treatment, subjects also solve one word puzzle (puzzle word: *memorable*). As in the previous treatment, there is only one “winner.” However, “the winner” is chosen at random out of the 4 people in the group. This person receives 40 cents for every point earned. The other three members of the group receive 0 points.

2.4 Reduced Time Pressure Treatments

The second goal of this experiment is to understand the impact of time constraints on male and female performance with and without competition. For this purpose, we run a second experiment that closely resembles the above design, but differs from it along one dimension. In particular, the subjects now have ample amounts of time (five times the length of time in the previous experiment) to try to find as many sub-words within the larger word as possible. We now also give the subjects the opportunity to finish each treatment before the period of fifteen minutes allotted for each round elapses.

We conduct two treatments in this reduced time-pressure environment: noncompetitive (puzzle word: *allopathy*) and competitive (puzzle word: *equitable*). Because the design of these treatments is otherwise identical to that of the high-time-pressure treatments above, we can conduct parallel comparisons within the two piece-rate and the two tournament scenarios, as well as compare between the piece-rate and the tournament schemes under the lowered time constraint.

Table I summarizes the various treatments for this experiment.

Table I.
Treatment Summary.

Payment Type	Time Pressure	
	High	Low
Practice	X	
Piece-Rate	X	X
Tournament	X	X
Random	X	

3 Data Analysis

3.1 Variables and Summary Statistics

In our analysis, the main dependent variable is *score* measured as the total number of points accumulated in any round.¹²

The main explanatory variables are dummy variables, *gender*, with 1 denoting a female subject, and *competition*, with 1 denoting a tournament payment scheme. We also include standard demographic control variables in our regressions. In particular, we control for *age*, whether the subject is a native English *language* speaker (with 1 denoting a non-native speaker), and whether the subject’s field of study is broadly categorized into a mathematical concentration or a humanities-related *concentration* (with 1 denoting humanities). Finally, during the experiment we attempt to gauge whether the subjects’ confidence changes across the different treatments. We measure *gender confidence* by asking each subject to report his or her belief about who would be better in these tasks on average, men or women (1 denotes the perception that women are better at the verbal task).

Table II provides descriptive statistics for all the experiments. First, note that women, achieve a slightly higher average score (16.9) than men (15.7) in the verbal task even in the piece-rate treatments (although the difference in those treatments is not statistically significant).

Second, competition seems to have very different effects in low-time-pressure and high-time-pressure environments. Average performance of both men and women is enhanced by competition in the former (for men, the score increases from 14.8 in the piece-rate to 20.7 in the tournament, and for women, the score increases from 16.2 in the piece-rate to 27.5 in the tournament, on average). In a high-time-pressure treatment, on the other hand, average performance of both genders drops when competition is introduced (men’s scores decline from 15.7 in the piece-rate to 13.2 in the tournament, and women’s scores decline from 16.9 to 13.8 in the tournament, on average).

Third, the gender confidence variable confirms the findings of previous studies that suggest that women are stereotyped to be better at verbal tasks than men: 81.9% of male subjects and 90% of female subjects report that they think that women would be, on average, better than men in these puzzles. Thus, it is fair to say, that in our experiment there is a stereotype threat against men, rather than against women. Note that the piece-rate treatments below suggest that the perception that women are better at verbal tasks than men is likely false.

¹²Note that we could also create an adjusted score variable by dividing the score by the maximum possible number of points or by the total number of sub-words in the puzzle. However, because we have already chosen the puzzle words to have the closest possible number of subwords, all the results using the score and the adjusted score variables would be qualitatively identical.

Table II.
Descriptive Statistics.

	Mixed Groups			Men
	Men	Women	Total	Only
Mean Score in Piece-Rate (High Time Pressure)	15.7	16.9	16.3	17.8
Mean Score in Competition (High Time Pressure)	13.2	13.8	13.5	14.2
Mean Score in Random Winner (High Time Pressure)	15.3	15.7	15.5	16.5
Mean Score in Piece-Rate (Low Time Pressure)	14.8	16.2	15.5	17.6
Mean Score in Competition (Low Time Pressure)	20.7	27.5	24.0	23.7
Max Score in Piece-Rate (High Time Pressure)	28	34	34	30
Max Score in Competition (High Time Pressure)	24	29	29	25
Max Score in Piece Rate (Low Time Pressure)	30	30	30	29
Max Score in Competition (Low Time Pressure)	36	42	42	32
Max Score in Random Winner (High Time Pressure)	31	35	35	32
Mean Time Spent in Piece-Rate (Low Time Pressure, Secs)	839	792	815	773
Mean Time Spent in Competition (Low Time Pressure, Secs)	741	862	801	860
Mean Age (Years)	22.3	21.0	21.7	21.5
Gender Confidence (% think own gender is better at task)	28.1	90.0	—	36.4
Percent Studying Humanities (%)	30	32	31	33
Percent Native English Speakers (%)	87.1	90.0	88.5	66.7
Number of Subjects	32	30	62	12

3.2 Performance in High Time Pressure Environment

Figure 1 and Figure 2 present the number of points (score) achieved by men and women in the piece rate and in the tournament treatments, respectively. The height of the bars in both figures corresponds to the share of male and female participants, respectively, who achieved the score in a given range.

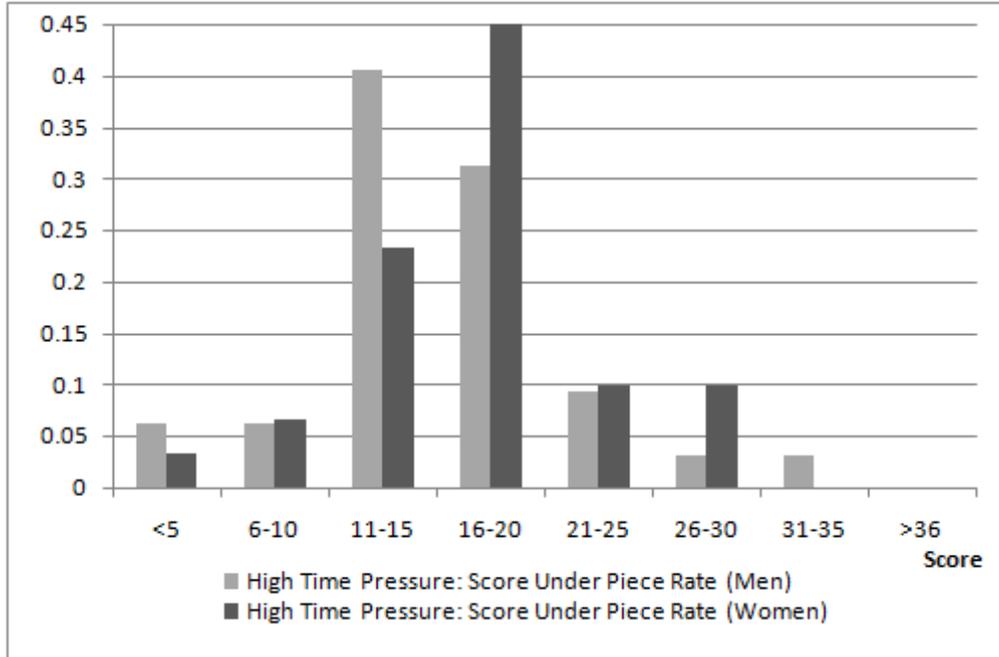


Figure 1: Score Under Piece Rate in High Time Pressure Environment

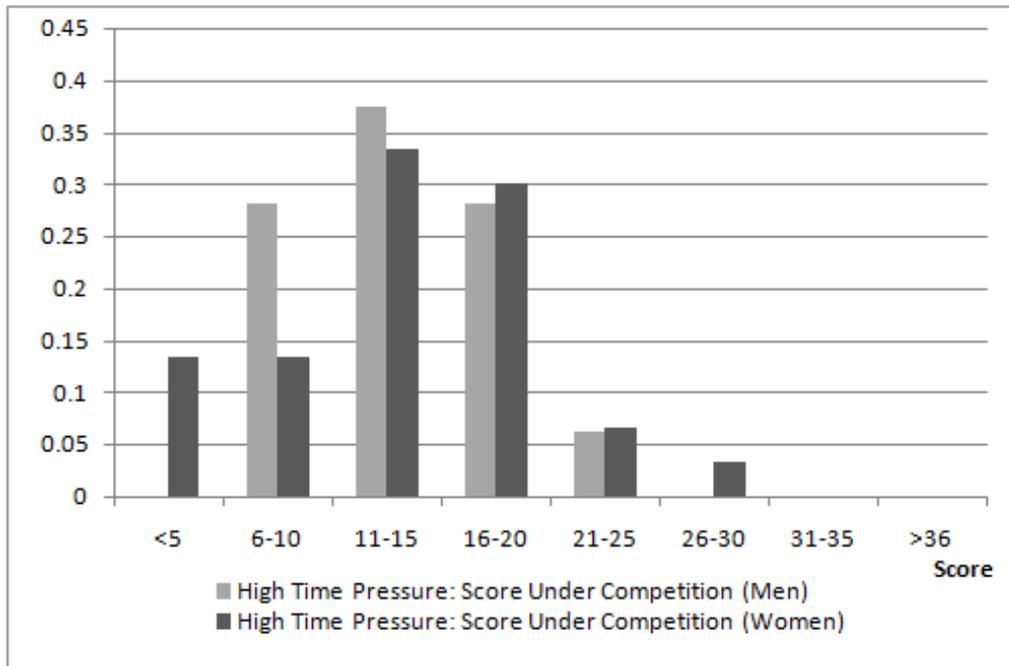


Figure 2: Score Under Competition in High Time Pressure Environment

The side-by-side comparison of the distributions of scores under piece rate and under mixed tournament (the competitive environment) leads to two observations. First, women and men do not appear to differ in performance within the two treatments. Second, both genders seem to do slightly worse in the competitive environment. In order to test the first observation, we follow the previous literature and perform the two-sided Mann-Whitney U test, which compares distributions. The p -value of the Mann-Whitney test is 0.3476 in the

piece-rate condition and 0.5723 in the tournament condition, which implies that there are no statistically significant differences in performance between men and women in either case.

Furthermore, in this high-pressure environment, competition seems to reduce the performance of men (with the p -value of the Mann-Whitney test of 0.073) as well as the performance of women (with the p -value of the Mann-Whitney test of 0.047).

Note that the tournament scheme differs from the piece rate condition in two ways: the payment depends on the performance of others and it is uncertain. For example, Dohmen and Falk (2006) attribute part of the gender difference in preferences for the competitive environment to differences in the degree of risk aversion. We therefore need to check that the results are truly driven by the effects of competition rather than risk aversion. In order to introduce uncertainty without competition, we run another treatment, where once again only one group member is paid, but that person is chosen at random. The results are similar in the random-winner treatment condition, where the p -value of the Mann-Whitney test is 0.7892.

We conclude that, unlike in mathematical tasks, in verbal tasks there is no significant difference between men and women when competition is involved. Competition does not boost performance of men. This allows us to reject the first hypothesis: competition alone does not seem to be the driving force behind gender differences in performance.

Next, we consider the second hypothesis proposed in the introduction: namely that the feelings of stereotype threat create gender differences under competition. Note that the above results for men could to be in line with the stereotype threat explanation: a man knows that in competition he is facing one person drawn from the same skill distribution as himself and two people who have higher verbal ability. Therefore, a man has a lower expectation of winning the tournament, which may hinder his performance. However, this reasoning would suggest that women should perform better under competition in verbal tasks, which we do not find. In order to resolve the question of the role of stereotype threat under competition, we compare the piece-rate and competition scores of men in a single-sex environment. We find that men do not perform significantly better relative to piece-rate when competing only against other men (the p -value of the Mann-Whitney test is 0.148).¹³ Therefore, we can reject the second hypothesis that stereotype threat is a major cause of gender differences in a competitive environment. Figure 3 provides a visual summary of the results. It shows that men and women do not perform differently in any of the treatments, and that men in a single-sex environment also do not perform significantly better than men (or women) in the mixed tournament.

¹³Note that so far we only have data for three single-sex groups. In the future, we plan to collect more data in this treatment in order to increase robustness of results.

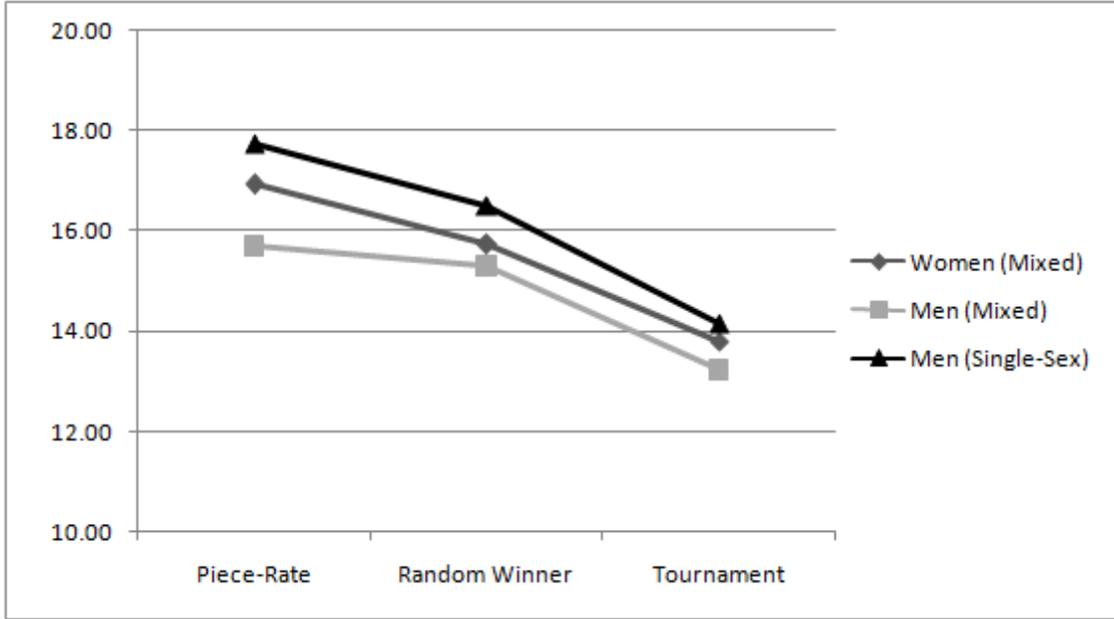


Figure 3: Average Score in the High-Time-Pressure Treatments

3.3 Performance in Low Time Pressure Environment

One potential reason behind women’s lack of improvement in performance under competition can be competition itself: women never compete.¹⁴ But another explanation is that time constraints and pressure affect men and women differently. In order to shed light on the latter question, we conduct an extended-time treatment, keeping the difficulty of the tasks and all other factors the same. In this low-time-pressure environment, we have two rounds: one with a piece-rate payment scheme and one with a tournament payment scheme. Figure 4 presents average scores of men and women for all the treatments in the experiment.

¹⁴We know from previous studies (Gneezy, Niederle, and Rustichini 2003) that women do not always shy away from competition. For example, women do perform better when facing only other women in a tournament.

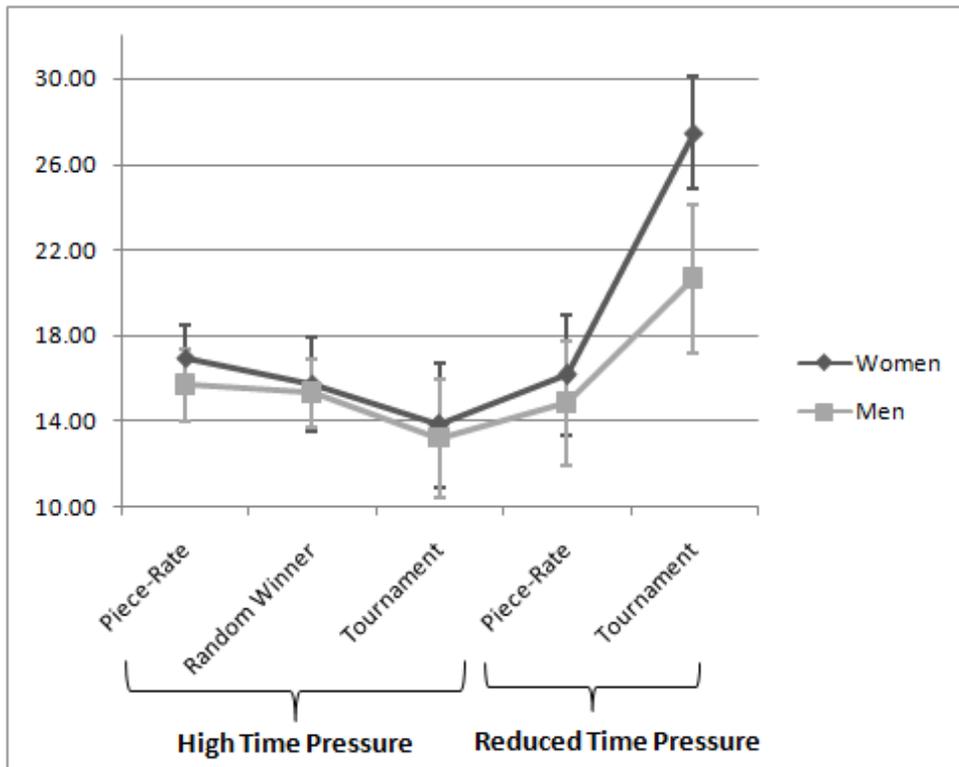


Figure 4: Average Score for All Treatments (Confidence Intervals at the 90% Confidence Level)

Figure 5 and Figure 6 present the number of points (score) achieved by men and women under low time pressure in the piece rate and in the tournament treatment, respectively.

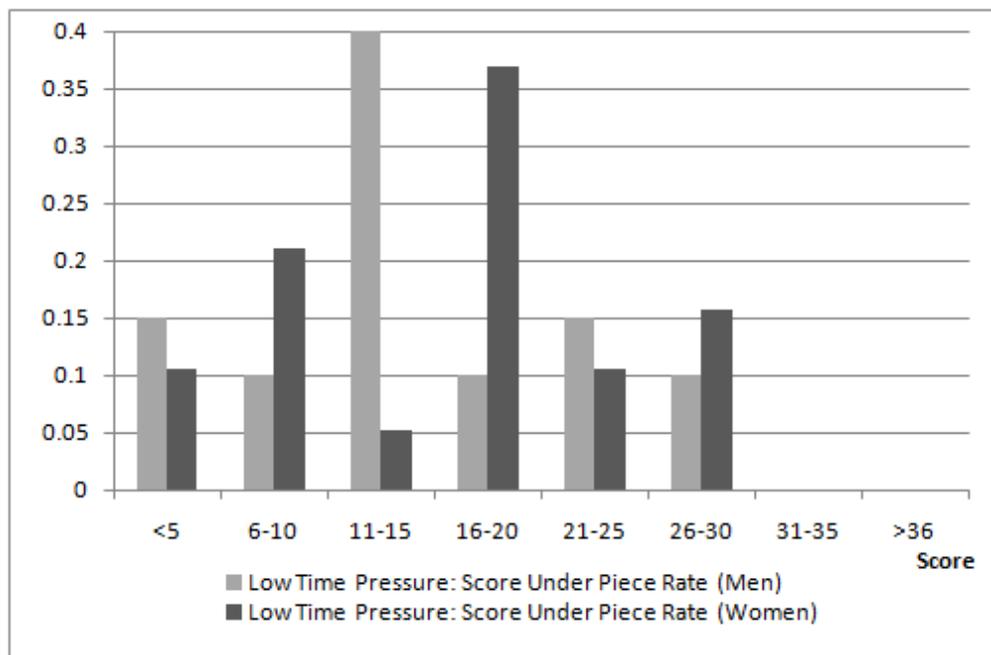


Figure 5: Score Under Piece Rate in Low Time Pressure Environment

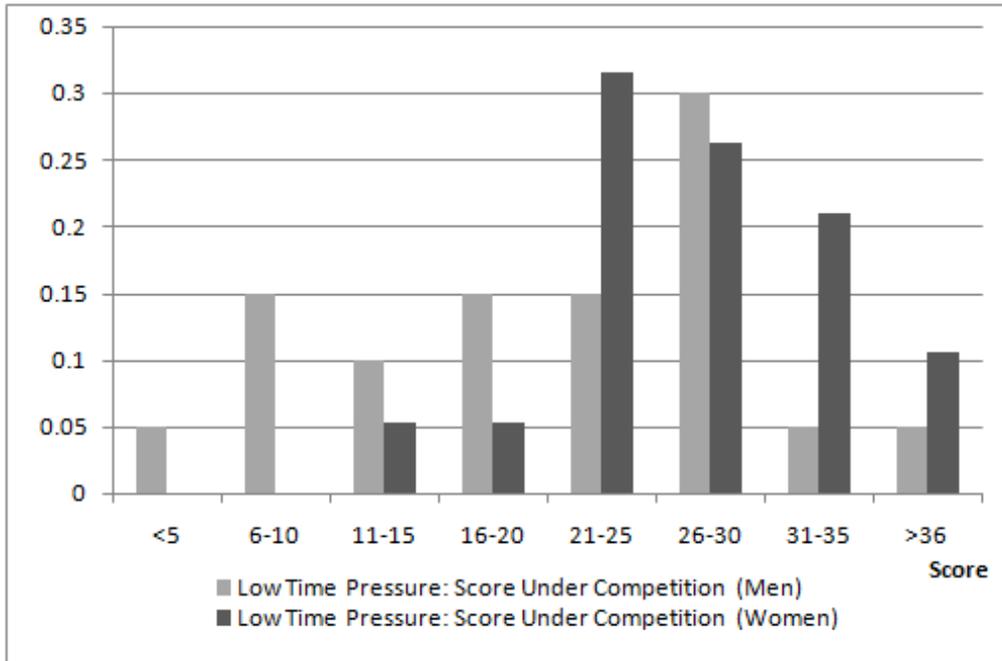


Figure 6: Score Under Competition in Low Time Pressure Environment

Comparing the distributions of scores for women and men in Figure 5 and Figure 6, we note that the availability of extra time significantly changes the outcomes. First, we observe that extra time increases the performance in the competitive round for both, men and women. Comparing the piece-rate treatment to the tournament, men increase performance with the p -value of the Mann-Whitney test of 0.0422 which is statistically significant at the 5 percent confidence level. The women increase performance with the p -value of the Mann-Whitney test of 0.0001 which is statistically significant at the 1 percent confidence level. Second, we find that competition with extended time boosts female performance beyond the performance of men. In order to test whether the apparent gender difference under competition is statistically significant, we again perform the Mann-Whitney U Test in both treatments. The p -value of the Mann-Whitney test is 0.673 in the piece-rate condition which implies that there is no statistically significant difference between the scores of men and women here. However, the p -value is 0.043 in the low-pressure tournament condition: women perform significantly better than men in this competitive environment.¹⁵

3.4 Women and Men: Quality vs. Quantity

In this section we look deeper into the mixed-group data in order to shed light on the origins of the differential performance of men and women in extended-time competitive environments.

¹⁵Note that it is possible that some of the positive effect of competition on both genders stems from the usage of a different puzzle word in the two treatments. However, first of all, according to the complexity factor, the tournament puzzle word is actually slightly more difficult than the piece-rate word. Secondly, the effect of competition and reduced time pressure is strong enough in magnitude to dominate any of these word effects. In future work, we plan to test this hypothesis by switching the order of the words in piece-rate and competition rounds.

Most importantly, the significant gender gap that we find here arises independently of word choice and order, since all subjects face the same puzzle in a particular round.

In particular, we focus on a measure of mistakes in various treatment conditions. This measure is simply the number of points lost due to entering invalid words.

First, we compare mistakes made under piece-rate and tournament in the high pressure environment. The Mann-Whitney U Test p -values of 0.5474 and 0.4292 in piece-rate and competition, respectively, suggest that the numbers of mistakes made by women and men do not differ significantly in either treatment when time pressure is relatively high. With reduced time pressure, the p -value in the piece rate condition is 0.9653, which once again implies that men and women make the same number of mistakes. However, in the low-time-pressure tournament, the p -value is 0.0275, which means that, at the 5 percent confidence level, we can reject the null hypothesis that men and women make the same number of mistakes.

Figure 7 documents the average number of mistakes made by men and women in all treatments.

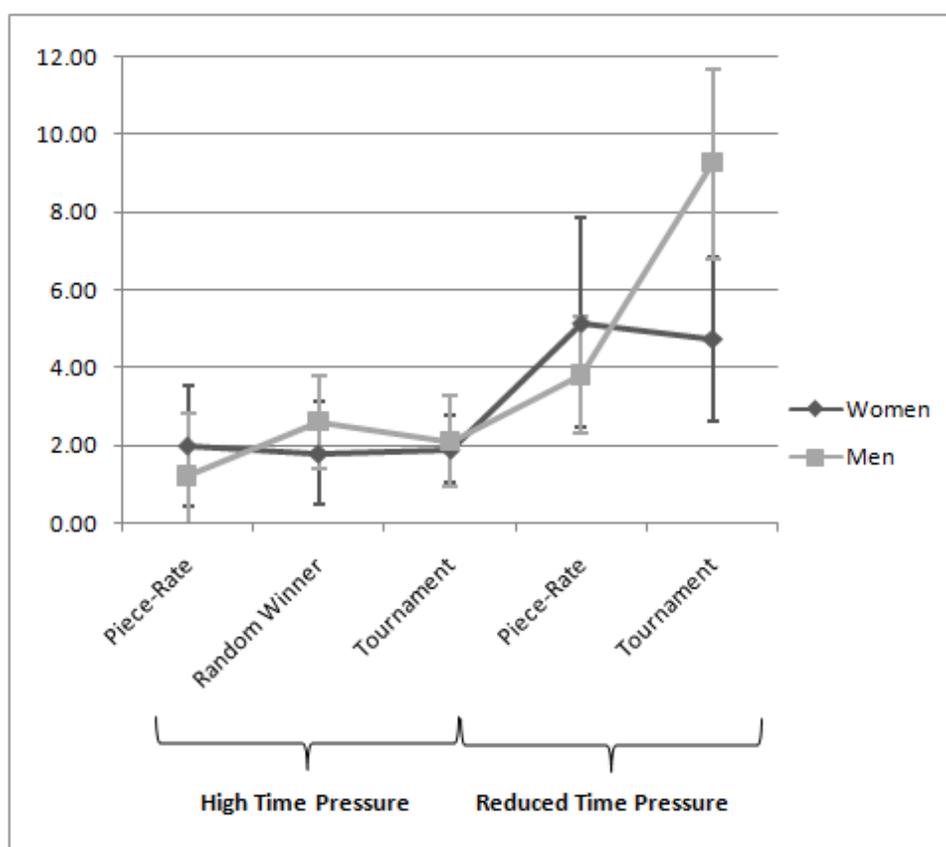


Figure 7: Average Mistakes (in terms of Points) for All Treatments (Confidence Intervals at the 90% Confidence Level)

In the low pressure environment, competition slightly reduces the number of mistakes made by women from the average of 5.2 points to 4.7 points. However, the important factor contributing to the men falling behind the women in this treatment condition is that reduced time pressure increases the number of mistakes made by men significantly from the average of 3.8 points in the piece-rate condition to the average of 9.3 points in the tournament. This is an average increase of 5.5 points, which is a large effect, given that the mean number

of points earned by men in the entire experiment is approximately 16 points. (The Mann-Whitney U Test rejects the null hypothesis that men do not make more mistakes significantly at the 1 percent confidence level with a p -value of 0.0084.)

The results suggest that while women might see the increase in time they can spend on the task as a way to increase the quality of their words, men seem to view the extended time only as a way to come up with more words. Table III documents the ratio of points lost due to invalid words (the quality measure) to the total number of points that the subject could have achieved if all the words were valid (the quantity measure). In the reduced-time-pressure tournament, men’s percentage of mistake points out of all possible points is almost double that of women (30.6 percent vs. 16.2 percent).

Table III.
The Percentage of “Mistake” Points Relative to the Total Possible
Points Across Treatments (Quality-to-Quantity Ratio, %).

	Men	Women
High Time Pressure Piece-Rate	6.1	10.1
High Time Pressure Tournament	12.5	12.1
Low Time Pressure Piece-Rate	19.5	28.0
Low Time Pressure Piece-Rate	30.6	16.2

On average, in the reduced time-pressure environment, men can potentially achieve a slightly higher score than women (30.2 vs. 29.2) because they enter more words. Unfortunately for men, this increase in quantity seems to greatly reduce quality which results in an overall reduced performance.

Recall also that in the part of the experiment with reduced time pressure, the subjects are able to withdraw from the game at any point during the round by clicking a “finish” button. The instructions clearly explained to the participants that they may click this button after they had come up with all the sub-words they possibly can. Looking back at the descriptive statistics (Table II), we note that women tend to spend more time (give up less) in the competitive round than in the piece-rate. The opposite is true for men: they seem to give up more in the competitive round. The effect is statistically significant at the 13% confidence level for women, according to the Mann-Whitney U Test and may also add to the explanation of the gender difference between men and women.

4 Conclusion

This paper uses controlled experiments to address two major issues in the literature regarding gender differences. First, we explore further a hypothesis put forth in previous studies that men outperform women in *all* competitive environments. Previous studies used tasks of mathematical nature, which are stereotyped to be better suited for men. In this paper, we use verbal tasks, which are typically thought to favor women. We find no gender differences in performance in either non-competitive or competitive payment conditions. Neither gender shows improvement in performance when the payment scheme changes from piece-rate to tournament. These findings lead to a rejection of the first hypothesis.

Second, we test the hypothesis that stereotype threat affects both men and women in a symmetric fashion, favoring the gender for which the task is perceived to be most well-suited. Our study finds that with verbal tasks, men do not improve performance in a single-sex tournament relative to men in a mixed tournament. This evidence implies a rejection of the second hypothesis.

Next, we propose an alternative explanation for the lack of a positive effect of competition on women. This explanation relies on the differential effect of time pressure on the performance of men and women. Our main finding is that women perform significantly better than men in competitive verbal tasks once the time constraint is relaxed. An important factor contributing to this gender difference is that men and women respond differently to reduced time pressure. While women seem to use the extra time to increase the quality of their work (effectively reducing the number of mistakes), men use the time to increase the quantity, producing a higher volume of work, but also increasing the share of mistakes.

We conclude that competition is not the cause of the gender gap in performance per se. Its effect depends greatly on the type of the task at hand. In the workplace, women and men face competition not only in terms of their ability to perform jobs of mathematical nature, but also in terms of their verbal abilities, such as writing reports, creating presentations, and talking to clients. According to previous studies, competition favors men in the former. We find that, in the latter, competition favors women, at least when they are given ample time to complete the task. Using a more creative task in this study, such as word puzzles, also allows us to show that women tend to do a more thorough job than men when given the opportunity. For example, with a task like solving mazes, quantity is all that matters, since an incorrectly solved maze is not penalized. With a verbal task, on the other hand, quality matters as well, which is what eventually separates women from men in competition.

The evidence documented in this paper suggests that gender inequality may be explained by the inherent differences in the responses of men and women to time pressure, rather than by societal conditioning, such as stereotype threat. This result implies that policies seeking to alleviate gender gaps in the workforce may want to subsidize a work culture that promotes a more relaxed atmosphere and lowered time pressure.

The results so far suggest several directions for future research. In addition to the two hypotheses proposed and tested in this paper, we consider a third hypothesis that stems from the rejection of the first two. In particular, the hypothesis claims that stereotype threat only affects women and not men in a competitive setting. Note that we observe that competition in verbal tasks does not seem to help women. Previous studies (for example, see Gneezy, Niederle, and Rustichini (2003)) suggest that women may do worse relative to men in competition precisely because they expect men to have superior mathematical abilities. If men were affected by stereotype threat in our study, then the finding that women do not improve performance under competition would lead to a rejection of the third hypothesis. However, in our study, stereotype threat does not have an effect on men, and therefore, we cannot test the third hypothesis with available data. In other words, we cannot distinguish that women do not do better under competition because they are affected by male presence even in verbal tasks or because they do not change their behavior for strategic reasons, anticipating that stereotype threat would not have an effect on men. However, the extended-time experimental framework provides an opportunity for testing this third hypothesis in the future. The new treatment that would provide such a test is one

where men and women perform tasks of a mathematical nature in a reduced-time-pressure environment. Comparing performance of men and women in the piece-rate and tournament conditions in this treatment would allow for a direct comparison of the responses of men and women to stereotype threat.

Another important robustness check that has yet to be conducted involves running more sessions of the experiment and changing the order of the competitive and the non-competitive treatments. This is necessary because we are interested in within-subject comparison and therefore keep the same subjects from piece-rate to competition. Reversing the order of the treatments will eliminate potential order effects.

So far, we have documented that women's *performance* is higher than men's in the extended-time competitive environment with a verbal task, which allows us a direct comparison to some of the previous literature's results. However, previous literature has established another important result, namely, that women choose to enter into competitions with male-oriented tasks much less than men (Niederle and Vesterlund (2007)). In order to provide a parallel using a verbal task, we propose to run sessions where men and women are asked to *self-select* into one of the two types of environments: non-competitive and competitive. We hypothesize that women will self-select more than men into the competitive verbal task only when there is reduced time pressure.

Finally, extra single-sex sessions are also required in order to add data to our currently small sample.

References

- [1] Andreoni, James, and Lise Vesterlund, “Which Is the Fair Sex? Gender Differences in Altruism,” *Quarterly Journal of Economics*, 116(1), February 2001, pp. 293–312.
- [2] Barres, Ben A. “Does Gender Matter?” *Nature*, 442(13), July 2006, pp. 133-136.
- [3] Baron-Cohen, Simon. “The Essential Difference. Men, Women, and the Extreme Male Brain.” 2003, London: Allan Lane, p. 256
- [4] Bertrand, Marianne and Hallock, Kevin F. “The Gender Gap in Top Corporate Jobs.” *Industrial and Labor Relations Review*, 55(1), October 2001, pp. 3-21.
- [5] Bhattacharya, Shaoni. “Gay Men Read Maps Like Women.” *Behavioral Neuroscience*, 119, February 2005, p. 311, Downloadable from <http://www.newscientist.com/article.ns?id=dn7069>
- [6] Black, Sandra E. and Strahan, Philip E. “The Division of Spoils: Rent-Sharing and Discrimination in a Regulated Industry.” *American Economic Review*, 91(4), September 2001, pp. 814-831.
- [7] Gneezy, Uri, Niederle, Muriel, and Rustichini, Aldo. “Performance in Competitive Environments: Gender Differences.” *Quarterly Journal of Economics*, 118(3), August 2003, pp. 1049-74.
- [8] Gneezy, Uri and Rustichini, Aldo. “Gender and Competition at a Young Age.” *American Economic Review*, 94(2), May 2004, pp. 377-81.
- [9] Gneezy, Uri, Leonard, Kenneth L., and John A. List. “Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society.” NBER Working Paper 13727, January 2008.
- [10] Lawrence Peter A. “Men, Women, and the Ghosts in Science.” *PLOS Biology*, 4(1), January 2006, pp. 13-15.
- [11] Macpherson, David A. and Hirsch, Barry T. “Wages and Gender Composition: Why Do Women’s Job Pay Less?” *Journal of Labor Economics*, 13(3), July 1995, pp. 426-471.
- [12] Niederle, Muriel, and Vesterlund, Lise. “Do Women Shy Away from Competition? Do Men Compete Too Much?” *Quarterly Journal of Economics*, August 2007, forthcoming.
- [13] Massachusetts Institute of Technology. A Study on the Status of Women Faculty in Science at MIT, 1999. Downloadable from <http://web.mit.edu/fnl/women/women.pdf>.
- [14] Pajares, Frank and Valiante, Giovanni. “Gender Differences in Writing Motivation and Achievement of Middle School Students: A Function of Gender Orientation?” *Contemporary Educational Psychology*, 26(3), July 2001, pp. 366-381.

- [15] Paserman, M. Daniele. "Gender Differences in Performance in Competitive Environments: Evidence from Professional Tennis Players." CEPR Discussion Paper 6335. February 2008.
- [16] Pease, Allan and Barbara Pease. "Why Men Don't Listen and Women Can't Read Maps: How We're Different and What to Do About It." New York, NY: Welcome Rain, 2000.
- [17] Polachek, Solomon W. "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure." *Review of Economics and Statistics*, 63(1), February 1981, pp. 60-69.
- [18] Spencer, Steven J., Steele, Claude M., and Quinn, Diane M. "Stereotype Threat and Women's Math Performance." *Journal of Experimental Social Psychology*, 35(1), January 1999, pp. 4-28.
- [19] Steele, Claude M. and Aronson, Joshua. "Stereotype Threat and the Intellectual Test Performance of African-Americans." *Journal of Personality and Social Psychology*, 69, 1995, pp. 797-811.

5 Appendix A: Experimental Documents

Instructions for the experiment (Part 1)

Thank you very much for participating in this experiment which involves solving word puzzles. Please read the following instructions very carefully. You will receive all the information you require for participation in the experiment. If you follow these instructions, you will have an opportunity to earn real money that will be paid to you, privately and in cash, at the end. If you do not understand something, please raise your hand and wait for the experimenter to come to your place and answer your question privately.

Communication between participants is absolutely forbidden during the experiment! Not obeying this rule will lead to immediate exclusion from the experiment and all payments.

Groups:

Throughout the experiment, you will be a member of a **group of 4 people** (two men and two women). The other three members of your group are sitting in the same row of desks as you.

Puzzles:

You will be solving **Word-in-a-Word** puzzles. You will have 3 minutes to find as many smaller sub-words that can be formed out of the letters of the big word as you possibly can. Everyone will be working on the same puzzle at the same time.

General rules:

- The words must be 4-letters long or longer
- Acceptable characters are letters A-Z only. Any other symbol like a number or another symbol will be automatically discarded
- Proper nouns are not allowed
- Plurals are allowed
- Each letter in the word can only be used once

Scoring rules:

- Valid words add $(N - 3)$ points to your score, where N is the total number of letters in the word
- Invalid words subtract $(N - 3)$ points from your score, where N is the total number of letters in the word

- If the word is a duplicate, no points are subtracted, no points are added
- If the word is too short, 1 point is subtracted from the total score (regardless of whether the word is 1, 2, or 3 letters long)

Example: Big word PERSUASIVELY. You enter: persuasive (+7), live (+1), live (0), lyve (-1), lyver (-2), sap (-1), as (-1). The total number of points is 3.

Rounds and Payment:

The first part of the experiment consists of **4 rounds**. In each round, you will have a chance to work on a different word puzzle. You will have 3 minutes to work on each puzzle (**3 minutes per puzzle**).

Round 1: In this round, you will be given 3 minutes to solve one word puzzle for **practice**. This puzzle will not add or subtract from your total score.

Round 2: In this round, you will be asked to solve one word puzzle. Each of you will receive **10** cents for every point you earn in this round.

Round 3: In this round, you will again solve one word puzzle. However, the scoring will be different. Your total score in this round will be compared to the scores of the other three members in your group. The person with the highest score (“the winner”) in this round will receive **40** cents for every point earned. The other three members of the group will receive 0 points. Note that, in case of a tie, the winner will be determined randomly out of the two top performers.

Round 4: In this round, you will also solve one word puzzle. As in Round 3, there will only be one “winner” in this round. However, “the winner” will be chosen at random out of the 4 people in the group and will receive **40** cents for every point earned. The other three members of the group will receive 0 points.

Exact Procedure

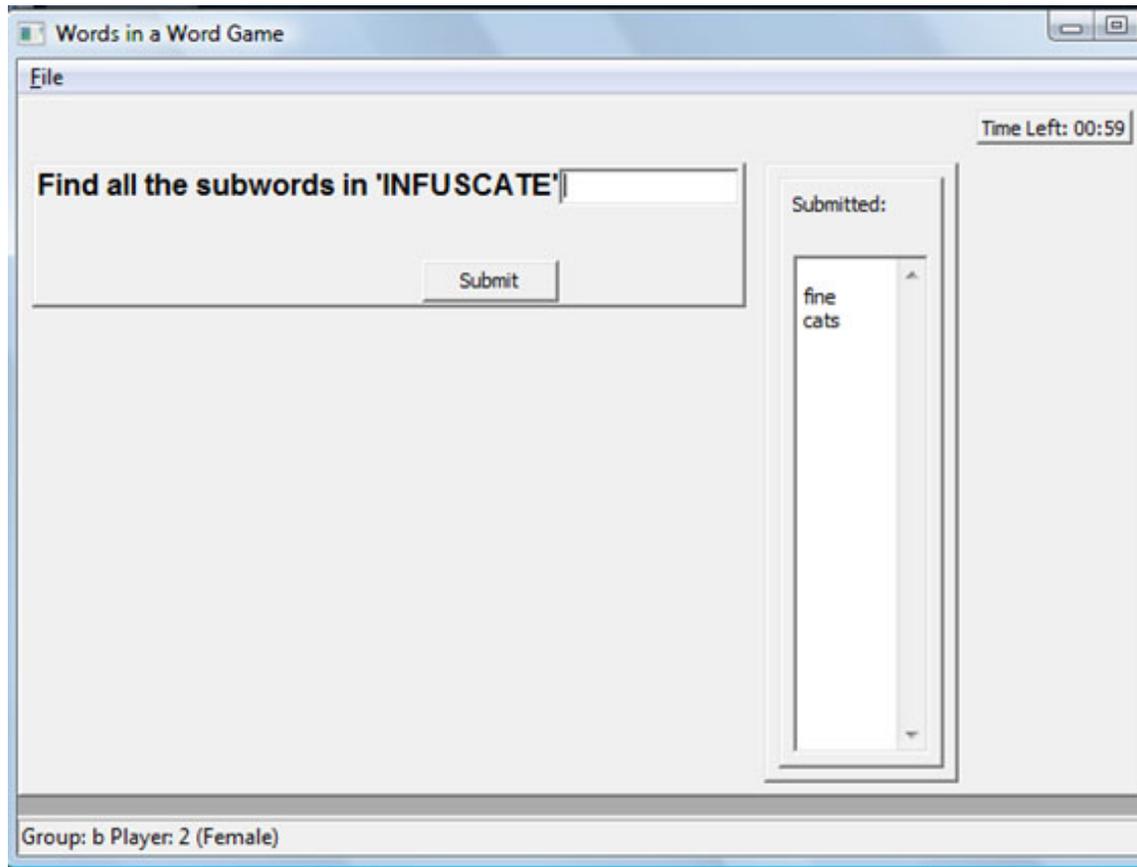
At the beginning of the experiment, you will receive an ID number. For example, members of the same group might receive ID numbers A1, A2, A3, and A4. Members of another group might receive ID numbers B1, B2, B3, and B4.

On the first screen you see, you will enter your ID number. Next, you will click “Submit.”

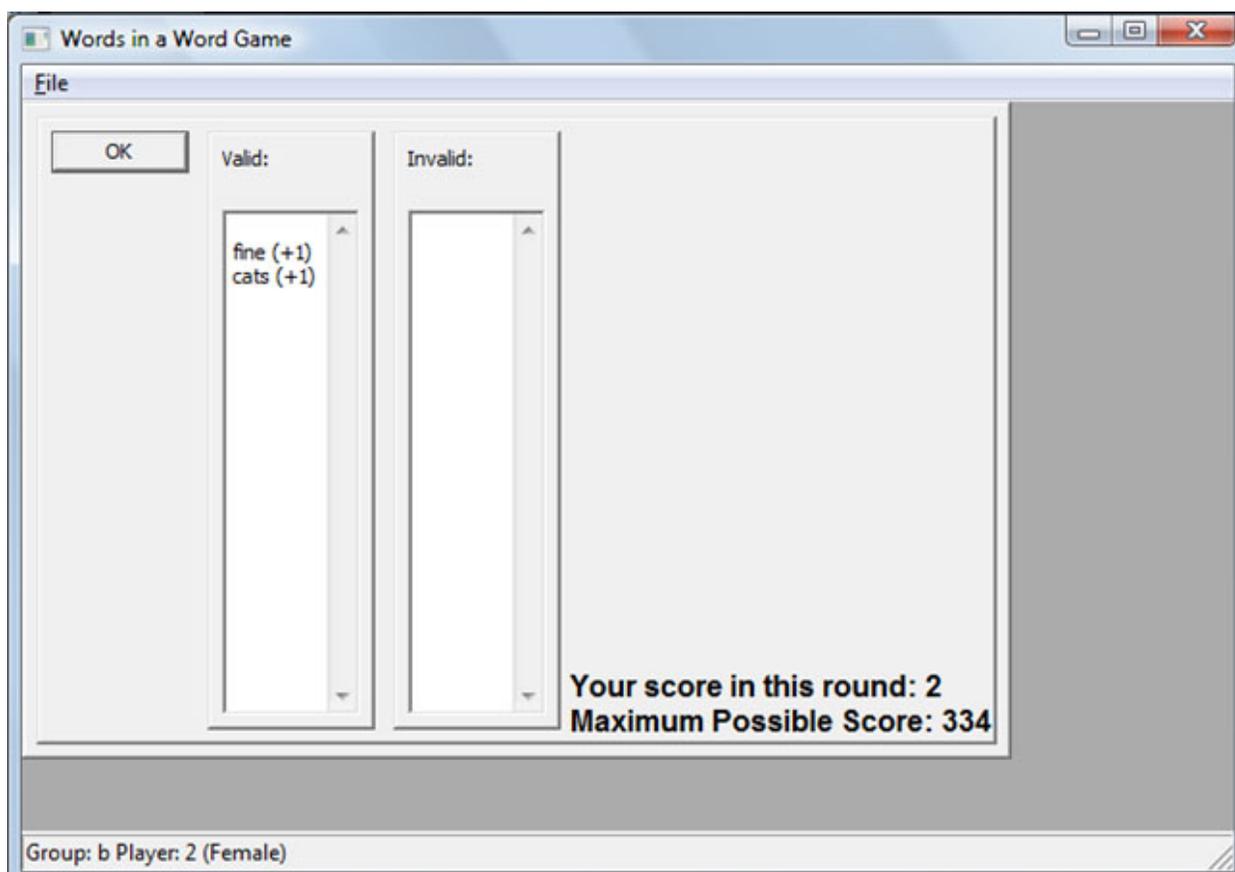
To start the first round, press “Start” on the screen that follows.

On the next screen, you will be presented with the practice puzzle (Round 1). Remember that you will have 3 minutes to come up with as many valid words as you can. After typing in your sub-words, you can either press the “Enter” key on the keyboard or click “Submit.”

Make sure you check the spelling of your sub-words carefully, since you will NOT be able to go back and edit them once they have been submitted.



The time will run out automatically. Once it does, you will see your score (the number of points you earned in this practice round) and the total possible number of points you could have earned in this puzzle. Please, click "OK" to proceed.



Please, be patient. You might need to wait 30 seconds to 2 minutes in between rounds!

Clicking “OK” will bring you to an information screen. Please, click the button labeled with “Proceed to next Round,” at which point the second round will start. The procedures for Rounds 2, 3, and 4 are identical to the procedure in Round 1.

You will receive your payment at the end of the experiment. Payment calculation for Part 1:

	Your income from Round 1 [= \$0]
+	Your income from Round 2
+	Your income from Round 3
+	Your income from Round 4
<hr/>	
=	Total income for Part 1

Control Questions

1. You are in Round 2. Your big word is CREATIVITY. You find words Reactivity, Race, Rat, Creative, and Crate.
 - (a) Your score in points is _____ points. (Answer: 4 Points)
 - (b) Your income made on this word puzzle is \$ _____ (or _____ cents)

2. You are in Round 3. Your score is 30 points. The other members of your group scored 14 points, 25 points, and 29 points.
 - (a) Your income made on this word puzzle is \$ _____ (or _____ cents)
 - (b) The income of the person with the score of 25 points is \$ _____ (or _____ cents)

3. You are in Round 3. Your score is 30 points. The other members of your group scored 14 points, 25 points, and 30 points.
 - (a) Your income is \$ _____ (or _____ cents) with probability of $\frac{1}{2}$ and \$ _____ (or _____ cents) with probability of $\frac{1}{2}$.
 - (b) The income of the person with the score of 25 points is \$ _____ (or _____ cents)

4. You are in Round 4. Your score is 30 points. The other members of your group scored 14 points, 25 points, and 39 points.
 - (a) Your income is \$ _____ (or _____ cents) with probability of $\frac{1}{4}$ and \$ _____ (or _____ cents) with probability of $\frac{3}{4}$.

Instructions for the experiment (Part 2)

The second part of this experiment is identical to the first part with the exception that you will now have 15 minutes to solve each word puzzle. At any point during the 15 minutes, you will be able to finish the round by clicking the button labeled “Finish.” **Please, DO NOT CLICK “Finish”** until you are absolutely sure that you cannot come up with any more words.

Rounds and Payment:

Round 1: In this round, you will be asked to solve a word puzzle. You should have as much time as you need to come up with as many words as you can (up to 15 minutes). Each of you will receive 5 cents for every point you earn in this round.

Round 2: In this round, you will again get a new word puzzle, and you will again have as much time as you need to come up with as many words as you can (up to 15 minutes). However, the scoring will be different. Your total score in this round will be compared to the scores of the other three members in your group. The person with the highest score (“the winner”) in this round will receive 20 cents for every point earned. The other three members of the group will receive 0 points. Note that, in case of a tie, the winner will be determined randomly out of the two top performers.

Exact Procedure

The exact procedure is identical to the procedure described in the instructions for Part 1. The payment calculation for Part 2:

$$\begin{array}{r} + \quad \text{Your income from Round 1} \\ + \quad \text{Your income from Round 2} \\ \hline = \quad \text{Total income for Part 2} \end{array}$$

Total Income in the Experiment

$$\begin{array}{r} \text{Show-up fee [= \$10]} \\ + \quad \text{Total income for Part 1} \\ + \quad \text{Total income for 2} \\ \hline = \quad \text{Total income in the Experiment} \end{array}$$

Questionnaire

We ask you now to fill out a brief questionnaire. Please, answer as honestly and completely, as you possibly can, since your answers will help us tremendously!

1. What is your ID number (i.e. A1, B3, etc.)? _____
2. What is your gender? _____
3. What is your age? _____
4. Are you a native English speaker (i.e., is English your first language)? _____
5. Are you currently a student? _____
And if yes, what is your field of study? _____

6. In your opinion, who would be better in these tasks on average, men or women (circle one)?

Men

Women

7. How did changing the rules from non-competitive to competition change your effort and performance in the puzzles? Why?

8. Circle one answer:

Competition:	a. helped me.	b. hurt me
I tried more in:	a. non-competitive rounds	b. competitive rounds
I gave up:	a. never	b. more in non-competitive rounds
	c. more in competitive rounds	d. in both type of rounds

THANK YOU VERY MUCH FOR YOUR PARTICIPATION IN THIS STUDY!!!

NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI

Fondazione Eni Enrico Mattei Working Paper Series

Our Note di Lavoro are available on the Internet at the following addresses:

<http://www.feem.it/Feem/Pub/Publications/WPapers/default.htm>

<http://www.ssrn.com/link/feem.html>

<http://www.repec.org>

<http://agecon.lib.umn.edu>

<http://www.bepress.com/feem/>

NOTE DI LAVORO PUBLISHED IN 2009

- SD 1.2009 Michael Hoel: [Bush Meets Hotelling: Effects of Improved Renewable Energy Technology on Greenhouse Gas Emissions](#)
- SD 2.2009 Abay Mulatu, Reyer Gerlagh, Dan Rigby and Ada Wossink: [Environmental Regulation and Industry Location](#)
- SD 3.2009 Anna Alberini, Stefania Tonin and Margherita Turvani: [Rates of Time Preferences for Saving Lives in the Hazardous Waste Site Context](#)
- SD 4.2009 Elena Ojea, Paulo A.L.D. Nunes and Maria Loureiro: [Mapping of Forest Biodiversity Values: A Plural Perspective](#)
- SD 5.2009 Xavier Pautrel : [Macroeconomic Implications of Demography for the Environment: A Life-Cycle Perspective](#)
- IM 6.2009 Andrew Ellul, Marco Pagano and Fausto Panunzi: [Inheritance Law and Investment in Family Firms](#)
- IM 7.2009 Luigi Zingales: [The Future of Securities Regulation](#)
- SD 8.2009 Carlo Carraro, Emanuele Massetti and Lea Nicita: [How Does Climate Policy Affect Technical Change? An Analysis of the Direction and Pace of Technical Progress in a Climate-Economy Model](#)
- SD 9.2009 William K. Jaeger: [The Welfare Effects of Environmental Taxation](#)
- SD 10.2009 Aude Pommeret and Fabien Prieur: [Double Irreversibility and Environmental Policy Design](#)
- SD 11.2009 Massimiliano Mazzanti and Anna Montini: [Regional and Sector Environmental Efficiency Empirical Evidence from Structural Shift-share Analysis of NAMEA data](#)
- SD 12.2009 A. Chiabai, C. M. Travisi, H. Ding, A. Markandya and P.A.L.D Nunes: [Economic Valuation of Forest Ecosystem Services: Methodology and Monetary Estimates](#)
- SD 13.2009 Andrea Bigano, Mariaester Cassinelli, Fabio Sfera, Lisa Guarrera, Sohbet Karbuz, Manfred Hafner, Anil Markandya and Ståle Navrud: [The External Cost of European Crude Oil Imports](#)
- SD 14.2009 Valentina Bosetti, Carlo Carraro, Romain Duval, Alessandra Sgobbi and Massimo Tavoni: [The Role of R&D and Technology Diffusion in Climate Change Mitigation: New Perspectives Using the Witch Model](#)
- IM 15.2009 Andrea Beltratti, Marianna Caccavaio and Bernardo Bortolotti: [Stock Prices in a Speculative Market: The Chinese Split-Share Reform](#)
- GC 16.2009 Angelo Antoci, Fabio Sabatini and Mauro Sodini: [The Fragility of Social Capital](#)
- SD 17.2009 Alexander Golub, Sabine Fuss, Jana Szolgayova and Michael Obersteiner: [Effects of Low-cost Offsets on Energy Investment – New Perspectives on REDD –](#)
- SD 18.2009 Enrica De Cian: [Factor-Augmenting Technical Change: An Empirical Assessment](#)
- SD 19.2009 Irene Valsecchi: [Non-Uniqueness of Equilibria in One-Shot Games of Strategic Communication](#)
- SD 20.2009 Dimitra Vouvaki and Anastasios Xeapapadeas: [Total Factor Productivity Growth when Factors of Production Generate Environmental Externalities](#)
- SD 21.2009 Giulia Macagno, Maria Loureiro, Paulo A.L.D. Nunes and Richard Tol: [Assessing the Impact of Biodiversity on Tourism Flows: A model for Tourist Behaviour and its Policy Implications](#)
- IM 22.2009 Bernardo Bortolotti, Veljko Fotak, William Megginson and William Miracky: [Sovereign Wealth Fund Investment Patterns and Performance](#)
- IM 23.2009 Cesare Dosi and Michele Moretto: [Auctioning Monopoly Franchises: Award Criteria and Service Launch Requirements](#)
- SD 24.2009 Andrea Bastianin: [Modelling Asymmetric Dependence Using Copula Functions: An application to Value-at-Risk in the Energy Sector](#)
- IM 25.2009 Shai Bernstein, Josh Lerner and Antoinette Schoar: [The Investment Strategies of Sovereign Wealth Funds](#)
- SD 26.2009 Marc Germain, Henry Tulkens and Alphonse Magnus: [Dynamic Core-Theoretic Cooperation in a Two-Dimensional International Environmental Model](#)
- IM 27.2009 Frank Partnoy: [Overdependence on Credit Ratings Was a Primary Cause of the Crisis](#)
- SD 28.2009 Frank H. Page Jr and Myrna H. Wooders (lxxxv): [Endogenous Network Dynamics](#)
- SD 29.2009 Caterina Calsamiglia, Guillaume Haeringer and Flip Klijn (lxxxv): [Constrained School Choice: An Experimental Study](#)
- SD 30.2009 Gilles Grandjean, Ana Mauleon and Vincent Vannetelbosch (lxxxv): [Connections Among Farsighted Agents](#)
- SD 31.2009 Antonio Nicoló and Carmelo Rodríguez Álvarez (lxxxv): [Feasibility Constraints and Protective Behavior in Efficient Kidney Exchange](#)
- SD 32.2009 Rahmi İlkiliç (lxxxv): [Cournot Competition on a Network of Markets and Firms](#)
- SD 33.2009 Luca Dall'Asta, Paolo Pin and Abolfazl Ramezanzpour (lxxxv): [Optimal Equilibria of the Best Shot Game](#)
- SD 34.2009 Edoardo Gallo (lxxxv): [Small World Networks with Segregation Patterns and Brokers](#)
- SD 35.2009 Benjamin Golub and Matthew O. Jackson (lxxxv): [How Homophily Affects Learning and Diffusion in Networks](#)

SD	36.2009	Markus Kinateder (lxxxv): Team Formation in a Network
SD	37.2009	Constanza Fosco and Friederike Mengel (lxxxv): Cooperation through Imitation and Exclusion in Networks
SD	38.2009	Berno Buechel and Tim Hellmann (lxxxv): Under-connected and Over-connected Networks
SD	39.2009	Alexey Kushnir (lxxxv): Matching Markets with Signals
SD	40.2009	Alessandro Tavoni (lxxxv): Incorporating Fairness Motives into the Impulse Balance Equilibrium and Quantal Response Equilibrium Concepts: An Application to 2x2 Games
SD	41.2009	Steven J. Brams and D. Marc Kilgour (lxxxv): Kingmakers and Leaders in Coalition Formation
SD	42.2009	Dotan Persitz (lxxxv): Power in the Heterogeneous Connections Model: The Emergence of Core-Periphery Networks
SD	43.2009	Fabio Eboli, Ramiro Parrado, Roberto Roson: Climate Change Feedback on Economic Growth: Explorations with a Dynamic General Equilibrium Mode
GC	44.2009	Fabio Sabatini: Does Social Capital Create Trust? Evidence from a Community of Entrepreneurs
SD	45.2009	ZhongXiang Zhang: Is it Fair to Treat China as a Christmas Tree to Hang Everybody's Complaints? Putting its Own Energy Saving into Perspective
SD	46.2009	Eftichios S. Sartzetakis, Anastasios Xepapadeas and Emmanuel Petrakis: The Role of Information Provision as a Policy Instrument to Supplement Environmental Taxes: Empowering Consumers to Choose Optimally
SD	47.2009	Jean-François Caulier, Ana Mauleon and Vincent Vannetelbosch: Contractually Stable Networks
GC	48.2009	Massimiliano Mazzanti, Susanna Mancinelli, Giovanni Ponti and Nora Piva: Education, Reputation or Network? Evidence from Italy on Migrant Workers Employability
SD	49.2009	William Brock and Anastasios Xepapadeas: General Pattern Formation in Recursive Dynamical Systems Models in Economics
SD	50.2009	Giovanni Marin and Massimiliano Mazzanti: Emissions Trends and Labour Productivity Dynamics Sector Analyses of De-coupling/Recoupling on a 1990-2005 Namea
SD	51.2009	Yoshio Kamijo and Ryo Kawasaki (lxxxv): Dynamics, Stability, and Foresight in the Shapley-Scarf Housing Market
IM	52.2009	Laura Poddi and Sergio Vergalli: Does Corporate Social Responsibility Affect the Performance of Firms?
SD	53.2009	Valentina Bosetti, Carlo Carraro and Massimo Tavoni: Climate Change Mitigation Strategies in Fast-Growing Countries: The Benefits of Early Action
GC	54.2009	Alireza Naghavi and Gianmarco I.P. Ottaviano: Firm Heterogeneity, Contract Enforcement, and the Industry Dynamics of Offshoring
IM	55.2009	Giacomo Calzolari and Carlo Scarpa: On Regulation and Competition: Pros and Cons of a Diversified Monopolist
SD	56.2009	Valentina Bosetti, Ruben Lubowski and Alexander Golub and Anil Markandya: Linking Reduced Deforestation and a Global Carbon Market: Impacts on Costs, Financial Flows, and Technological Innovation
IM	57.2009	Emmanuel Farhi and Jean Tirole: Collective Moral Hazard, Maturity Mismatch and Systemic Bailouts
SD	58.2009	Kelly C. de Bruin and Rob B. Dellink: How Harmful are Adaptation Restrictions
SD	59.2009	Rob Dellink, Michel den Elzen, Harry Aiking, Emmy Bergsma, Frans Berkhout, Thijs Dekker, Joyeeta Gupta: Sharing the Burden of Adaptation Financing: An Assessment of the Contributions of Countries
SD	60.2009	Stefania Tonin, Anna Alberini and Margherita Turvani: The Value of Reducing Cancer Risks at Contaminated Sites: Are More Heavily Exposed People Willing to Pay More?
SD	61.2009	Clara Costa Duarte, Maria A. Cunha-e-Sá and Renato Rosa: The Role of Forests as Carbon Sinks: Land-Use and Carbon Accounting
GC	62.2009	Carlo Altomonte and Gabor Békés: Trade Complexity and Productivity
GC	63.2009	Elena Bellini, Gianmarco I.P. Ottaviano, Dino Pinelli and Giovanni Prarolo: Cultural Diversity and Economic Performance: Evidence from European Regions
SD	64.2009	Valentina Bosetti, Carlo Carraro, Enrica De Cian, Romain Duval, Emanuele Massetti and Massimo Tavoni: The Incentives to Participate in, and the Stability of, International Climate Coalitions: A Game-theoretic Analysis Using the Witch Model
IM	65.2009	John Temple Lang: Article 82 EC – The Problems and The Solution
SD	66.2009	P. Dumas and S. Hallegatte: Think Again: Higher Elasticity of Substitution Increases Economic Resilience
SD	67.2009	Ruslana Rachel Palatnik and Roberto Roson: Climate Change Assessment and Agriculture in General Equilibrium Models: Alternative Modeling Strategies
SD	68.2009	Paulo A.L.D. Nunes, Helen Ding and Anil Markandya: The Economic Valuation of Marine Ecosystems
IM	69.2009	Andreas Madestam: Informal Finance: A Theory of Moneylenders
SD	70.2009	Efthymia Kyriakopoulou and Anastasios Xepapadeas: Environmental Policy, Spatial Spillovers and the Emergence of Economic Agglomerations
SD	71.2009	A. Markandya, S. Arnold, M. Cassinelli and T. Taylor: Coastal Zone Management in the Mediterranean: Legal and Economic Perspectives
GC	72.2009	Gianmarco I.P. Ottaviano and Giovanni Prarolo: Cultural Identity and Knowledge Creation in Cosmopolitan Cities
SD	73.2009	Erik Ansink: Self-enforcing Agreements on Water allocation
GC	74.2009	Mario A. Maggioni, Francesca Gambarotto and T. Erika Uberti: Mapping the Evolution of "Clusters": A Meta-analysis
SD	75.2009	Nektarios Aslanidis: Environmental Kuznets Curves for Carbon Emissions: A Critical Survey
SD	76.2009	Joan Canton: Environmentalists' Behaviour and Environmental Policies
SD	77.2009	Christoph M. Rheinberger: Paying for Safety: Preferences for Mortality Risk Reductions on Alpine Roads

IM	78.2009	Chiara D'Alpaos, Michele Moretto, Paola Valbonesi and Sergio Vergalli: <u>"It Is Never too late": Optimal Penalty for Investment Delay in Public Procurement Contracts</u>
SD	79.2009	Henry Tulkens and Vincent van Steenberghe: <u>"Mitigation, Adaptation, Suffering": In Search of the Right Mix in the Face of Climate Change</u>
SD	80.2009	Giovanni Bella: <u>A Search Model for Joint Implementation</u>
SD	81.2009	ZhongXiang Zhang: <u>Multilateral Trade Measures in a Post-2012 Climate Change Regime?: What Can Be Taken from the Montreal Protocol and the WTO?</u>
SD	82.2009	Antoine Dechezleprêtre, Matthieu Glachant, Ivan Hascic, Nick Johnstone and Yann Ménière: <u>Invention and Transfer of Climate Change Mitigation Technologies on a Global Scale: A Study Drawing on Patent Data</u>
SD	83.2009	László Á. Kóczy: <u>Stationary Consistent Equilibrium Coalition Structures Constitute the Recursive Core</u>
SD	84.2009	Luca Di Corato and Michele Moretto: <u>Investing in Biogas: Timing, Technological Choice and the Value of Flexibility from Inputs Mix</u>
SD	85.2009	Valentina Bosetti, Enrica De Cian, Alessandra Sgobbi, and Massimo Tavoni: <u>The 2008 WITCH Model: New Model Features and Baseline</u>
IM	86.2009	Rocco Macchiavello: <u>Vertical Integration and Investor Protection in Developing Countries</u>
SD	87.2009	Massimiliano Mazzanti and Antonio Musolesi: <u>Carbon Kuznets Curves: Long-run Structural Dynamics and Policy Events</u>
GC	88.2009	Gianmarco I.P. Ottaviano and Christian Volpe Martincus: <u>SMEs in Argentina: Who are the Exporters</u>
GC	89.2009	Gianpaolo Rossini and Cecilia Vergari: <u>Input Production Joint Venture</u>
SD	90.2009	Angelo Antoci, Simone Borghesi and Marcello Galeotti: <u>Environmental Options and Technological Innovation: An Evolutionary Game Model</u>
GC	91.2009	Cristina Cattaneo: <u>The Decision to Migrate and Social Capital: Evidence from Albania</u>
SD	92.2009	Valentina Bosetti and Jeffrey Frankel: <u>Global Climate Policy Architecture and Political Feasibility: Specific Formulas and Emission Targets to Attain 460 ppm CO2 Concentrations</u>
IM	93.2009	Benno Bühler: <u>Do International Roaming Alliances Harm Consumers?</u>
GC	94.2009	Elena Fumagalli and Laura Fumagalli: <u>Like Oil and Water or Chocolate and Peanut Butter? Ethnic Diversity and Social Participation of Young People in England</u>
GC	95.2009	Olga Shurchkov: <u>Gender Differences in Output Quality and Quantity under Competition and Time Constraints: Evidence from a Pilot Study</u>

(lxxxv) This paper has been presented at the 14th Coalition Theory Network Workshop held in Maastricht, The Netherlands, on 23-24 January 2009 and organised by the Maastricht University CTN group (Department of Economics, http://www.feem-web.it/ctn/12d_maa.php).