

Working Paper Series
(ISSN 1211-3298)

589

**How Stress Affects Performance and
Competitiveness across Gender**

Jana Cahlíková
Lubomír Cingl
Ian Lively

CERGE-EI
Prague, May 2017

ISBN 978-80-7343-396-3 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-425-9 (Národohospodářský ústav AV ČR, v. v. i.)

How Stress Affects Performance and Competitiveness across Gender

Jana Cahlíková *

Lubomír Cingl

Ian Levely

May 2017

Abstract

Since many key career events, such as exams and interviews, involve competition and stress, gender differences in response to these factors could help to explain the labor-market gender gap. In a laboratory experiment, we manipulate psychosocial stress using the Trier Social Stress Test, and confirm that this is effective by measuring salivary cortisol. Subjects perform a real-effort task under both tournament and piece-rate incentives and we elicit willingness to compete. We find that women under heightened stress do worse than women in the control group when compensated with tournament incentives, while there is no treatment difference for performance under piece-rate incentives. For males, stress does not affect output under competition. We also find that stress decreases willingness to compete overall, and for women, this is related to performance. These results help to explain previous findings on gender differences in performance under competition both in and out of the lab.

JEL: C91, D03, J16, J33

Keywords: competitiveness, performance in tournaments, psychosocial stress, gender gap.

*Cahlíková: Max Planck Institute for Tax Law and Public Finance, Department of Public Economics, Marstallplatz 1, 80539 Munich, Germany and CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Prague, Czech Republic (email: jana.cahlíkova@tax.mpg.de); Cingl: Department of Institutional, Environmental and Experimental Economics, Faculty of Economics, University of Economics in Prague (email: lubomir.cingl@vse.cz); Levely: Development Economics Group, Wageningen University (email: Ian.Levely@wur.nl). We would like to thank Vojtěch Bartoš, Michal Bauer, Erwin Bulte, Alexander W. Cappelen, Belle Derks, Lorenze Goette, Michel Handgraaf, Johannes Haushofer, Štěpán Jurajda, Muriel Niederle, Haki Pamuk, Maroš Servátka, Simeon Schudy, Bertil Tungodden, and Bernadette von Dawans for helpful conversations at different stages of the project. This research was supported by the Grant Agency of Charles University (grant no. 1156213). All remaining errors are our own.

1. Introduction

Since students, job candidates and workers are often required to compete against peers in stressful settings, understanding differences in the way that men and women respond to competition under stress is vital in explaining the persistent gender gap in the labor market, and especially the under-representation of women in top positions in business, government, and academia. The events that are most influential on one's career—such as job interviews, university entrance exams and asking for promotion—involve competition, and in high-paying and influential positions both pecuniary compensation and prestige are often based heavily on one's performance relative to others. These career-determining interactions typically take place under heightened psychosocial stress, for example when job candidates or students are required to speak publicly and are judged in front of committees. We study how stress and competition affect men and women differently, by experimentally manipulating exposure to psychosocial stress and subsequently measuring performance in a real-effort task, under both competitive and non-competitive incentive schemes. We also test whether exposure to stress affects willingness to compete, and whether this differs by gender.

Previous research shows that women are less likely to enter competitive situations than men (Gneezy et al., 2003; Niederle and Vesterlund, 2007, 2011), and this has been highlighted as a potential explanation for the female wage gap. Lower willingness to compete could make women less likely to enter competitive fields than men with similar ability. Experimental measures of willingness to compete can, for example, partially explain female students' choices to enter less prestigious academic tracks (Buser et al., 2014). A substantial part of the gender wage gap is due not to choice of profession, but to quality of employers and advancement within fields (Card et al., 2016; Cardoso et al., 2016). This is plausibly related to lower willingness to compete, which could also make women less likely to ask for promotions or to apply for jobs that have a competitive application process or competitive compensation scheme (Flory et al., 2015).

In addition to preferences for competing, differences in performance under competition for men and women may also explain labor market outcomes. So far, the evidence is mixed. While Niederle and Vesterlund (2007) show that tournament incentives led to higher performance for both men and women, Gneezy et al. (2003) and Gneezy and Rustichini (2004) find that competitive incentives only increase performance among men. Using data from university exams, Ors et al. (2013) find that female students perform comparatively worse when competing against peers, and similarly Jurajda and Munich (2011) find that women do worse than their male counterparts on entrance exams only

when applying to more competitive programs.

Recent work shows that stress affects decision making and preferences (Starcke and Brand, 2012; von Dawans et al., 2012; Cahliková and Cingl, 2017) and given this, reaction to stressors might also shed light on gender differences in competitive behavior. The physiological and psychological aspects of the stress response can differ, depending on the type of stressor and the gender of the individual (Stroud et al., 2002). While “fight or flight” (Cannon, 1932) is the dominant model for understanding how humans (and other animals) respond to a perceived threat, females, who are less physically adapted to fight off foes and less mobile when caring for offspring, may have evolved a tendency to “tend-and-befriend,” by leveraging affiliation with social groups to avoid danger in some situations (Taylor et al., 2000; Taylor, 2006). We hypothesize that this difference in stress response between men and women could lead to gender differences in how stress affects performance under competition and willingness to compete; we designed the study to test this conjecture.

Specifically, we examine the effects of psychosocial stress on performance under tournament incentives and willingness to compete.¹ Psychosocial stress is related to psychological and social well-being, for example when an individual faces threats to status, self-image or social pressure, and is arguably more relevant for modern daily life than other forms of stress, such as physical stress, in which individuals face physical discomfort or threats to their survival (Dickerson and Kemeny, 2004).

Our design consists of an economic experiment with 95 male and 95 female university students, and uses a modified version of the Trier Social Stress Test for Groups (TSST-G) to manipulate stress (von Dawans et al., 2011). Subjects were assigned to either the *stress* or *control* treatment for the duration of the experiment. We measure salivary cortisol—a hormone related to stress²—and heart rate, in order to confirm that the stress manipulation was successful for both genders.

Using a laboratory experiment solves two problems that make causal inference of the effect of stress on behavior in naturally occurring competitions difficult. First, it avoids problems of self-selection into competitive and stressful situations. Second, since competitive situations can cause stress (Buckert et al., 2015; Fletcher et al., 2008), this makes it difficult to isolate the effects of stress and competition from one another in observational

¹In particular, we study acute (short-term) stress which has been shown to have psychological, neurological and behavioral effects distinct from those of long-term stress (McEwen, 2012). We concentrate on the former, which is more relevant for the labor market events that motivate this study.

²While the physiological effects of stress are complex, cortisol levels are released into the bloodstream at greater levels after exposure to stressors, and due to ease in measuring salivary cortisol levels, it is the most commonly used bio-marker for measuring physiological stress response (Hellhammer et al., 2009; Everly and Lating, 2013).

data.

The experiment measures the change in performance and willingness to compete under stress using a design based on [Niederle and Vesterlund \(2007\)](#). Subjects were compensated for adding up sets of four, two-digit numbers within a time limit. The payment scheme varied by condition: in the baseline condition, each correctly solved problem was rewarded with a fixed, piece-rate payment. This condition was repeated after the stress/control procedure to reveal the effect of stress on individual performance. Subjects also completed the task under a tournament incentive scheme, in which payoff depends on performance relative to another randomly selected participant. Next, subjects chose a linear combination of the piece-rate and tournament payment schemes, which is our measure of willingness to compete, after [Gneezy et al. \(2016\)](#). Subjects then performed the counting task again and were rewarded according to their choice. Afterwards, subjects made additional competition decisions that allow us to draw more precise conclusions regarding the underlying mechanism through which stress affects willingness to compete. We are able to rule out several channels, including risk aversion and confidence.

Our results show that stress indeed has a gender-specific effect on reaction to competition. Women in the stress treatment perform significantly worse in the tournament than do women in the control group. Interestingly, stress alone does not affect performance, and we find no treatment difference for female subjects' performance in the piece-rate rounds. Rather, the treatment difference in the tournament performance is due to an increase in performance for female subjects in the control group relative to that under piece-rate payment—women in the stress treatment actually perform slightly *worse* in the tournament. In contrast, men's performance under competition is not affected by stress: we find no statistically significant treatment difference in performance in the tournament. We also find that willingness to compete is lower in the stress treatment, overall. For women, this can be explained by differences in tournament performance across treatment.

This is consistent with a theory from the psychology literature that predicts lowered executive function with increased arousal, past a certain threshold ([Yerkes and Dodson, 1908](#)). Stress could interact with additional sources of strain on executive function that are present in the tournament and to which women might be more sensitive, such as higher stakes ([Ariely et al., 2009](#); [Azmat et al., 2016](#)), social comparison ([Schram et al., 2015](#)), or stereotype threat ([Schmader et al., 2008](#)).

While our paper is the first of which we are aware to study the effect of stress on performance under competition, it contributes to an emerging literature on the causal link between stress and willingness to compete. [Goette et al. \(2015\)](#) induce psychosocial stress, using the TSST-G procedure, and find no average effect on competitiveness. However,

they measure only decisions to compete based on the results of past performance—in this study, we replicate this (non)result. [Buser et al. \(2016\)](#) use an experimental game similar to that in our study to measure willingness to compete, but examine the effects of a physical stressor (putting a hand in ice-cold water), which likely has different physical and psychological effects than psychosocial stress ([Baum and Grunberg, 1997](#); [Haushofer and Jang, 2015](#)), and may be the reason that, in contrast to our study, they fail to find an impact on willingness to compete. We argue that our procedure produces a type of stress that is more relevant to labor market outcomes, as psychosocial stress is ubiquitous in professional settings, but physical stress is comparatively rare.³

Our results imply that women may be at a disadvantage when required to compete in stressful settings. This has broader implications for understanding how men and women approach competition, and adds to the discussion on the persistent under-representation of women in prestigious industries, high-paid jobs and leadership positions in politics and business. These careers involve both intense competition and stress. If women know that they do not perform well under these types of environments, they may decide to stay out. When hiring practices involve more competition or stress than the position itself, our results suggest that this could prevent employers from selecting women who are best suited for the job. If managers introduce competitive incentives as a means of boosting employee productivity, this may have the opposite effect for women in the presence of heightened psychosocial stress.

2. Design

All subjects completed several incentivized tasks which measure performance under piece-rate and tournament incentive schemes and willingness to compete. Our experimental manipulation consists of two treatments applied between-subjects: the stress treatment, in which subjects were exposed to a psychosocial stressor in the form of the TSST-G, and a control treatment. For a timeline of all tasks in the experiment, the stress/control procedures and cortisol measurements see Figure 1.

³In an unpublished paper, [Zhong et al. \(2016\)](#) also induce psycho-social stress and measure willingness to compete. While they do not find a significant effect of psychosocial stress on willingness to compete, they also find no gender differences in the control group. Since this is a result which has been replicated in numerous studies ([Gneezy et al., 2003](#); [Niederle and Vesterlund, 2007, 2011](#)), it is perhaps not surprising that they do not find the gender-specific differences in reaction to stress and competition that we do.

A. *Experimental tasks*

We measure competitiveness using a design based on [Niederle and Vesterlund \(2007\)](#) and [Gneezy et al. \(2016\)](#). Subjects completed a counting activity, twice under a non-competitive piece-rate scheme, then again under tournament incentives, after which they were asked which combination of these compensation schemes they preferred for the subsequent counting round.

The counting activity consisted of a series of addition problems, requiring subjects to add up four two-digit numbers in each. They had two minutes per task to solve as many problems as they were able to. Subjects familiarized themselves with the counting activity in an unpaid practice round; in following rounds correct results were incentivized according to two compensation schemes.

Under the *piece-rate* compensation scheme, participants earned CZK 25 (about EUR 1) per correct answer. Performance under the piece-rate scheme serves as a baseline measure of ability and effort. Subjects performed twice under the piece rate compensation scheme: once before the stress treatment/control procedure (Task 1, *Piece rate before treatment*) and once after (Task 2, *Piece rate under treatment*). Comparing the within-subject differences in performance in Task 1 and Task 2 across treatments allows us to measure the effect of the stress treatment on performance, controlling for baseline differences in ability.

In Task 3, *Tournament under treatment*, correct answers were rewarded according to the *tournament* compensation scheme: each participant was informed that he or she would be randomly matched with another participant in the room (there were always four males and four females present) and that whoever had more correct answers would receive CZK 50 per correct answer, while the subject with fewer correct answers would receive nothing in that task. In case of a tie, each participant received CZK 25 per correct answer, as in the piece-rate scheme. Comparing performance in Task 3 across treatments and with individual's performance in Tasks 1 and 2 allows us to assess how competitive incentives affect performance, and whether this changes by treatment and gender.

In Task 4, *Choice of compensation scheme for future performance*, subjects chose how they would be compensated before completing the counting portion of the task. They did so by splitting 100 points between the tournament and the piece-rate compensation schemes, as in [Gneezy et al. \(2016\)](#). For each point invested in the piece-rate scheme, they earned CZK 0.25 per correct answer. For each point invested into the tournament compensation scheme, they earned CZK 0.5 per correct answer, but only if they had more correct answers in Task 4 than another randomly selected participant, and received noth-

ing for each point invested in the tournament scheme if they answered fewer questions. In case of a tie, each point invested in the tournament account was rewarded according to the piece-rate scheme (CZK 0.25 per answer). Thus, if subjects invested all points into the piece-rate scheme, they were paid CZK 25 per correct answer, as in Task 1 and Task 2. If all points were invested in the tournament scheme, they received CZK 50 per question if they answered more questions than a randomly matched partner, 0 if they answered fewer problems and 25 in the event of a tie—as in Task 3. If they invested some points in the tournament scheme and some in the piece-rate scheme, they were paid according to a linear combination of the two compensation schemes. In order to make the decision easily understandable, before making their final choice subjects could experiment with different tournament investments and the resulting payoffs per correct question if they won and lost were displayed.

It is important to note that the choice of compensation scheme in Task 4 cannot be driven by pro-social concerns or beliefs about who self-selects into the tournament, as a subject's performance in Task 4 was compared to the Task 3 performance of another randomly selected subject. This information was highlighted in the instructions, and subjects knew that their decision to enter the tournament did not have payoff consequences for anyone else.

The choice of compensation scheme in Task 4 is our main measure of willingness to compete. To estimate the causal effect of stress, we compare the share of the 100 points invested in the tournament in Task 4 across the stress and control treatments. To determine the underlying mechanism, we implemented two additional tasks, in which subjects competed on past performance. This isolates preferences and beliefs related to performing in a competitive environment (which are relevant in Task 4 but not in Tasks 5-6) from willingness to compete and other beliefs and preferences that are present in all three (Niederle and Vesterlund, 2007).

In Task 5, *Choice of compensation scheme for past performance before treatment*, subjects again split 100 points between the tournament and piece-rate schemes, but were paid according to performance in Task 1. Subjects were reminded that Task 1 was incentivized with the piece-rate scheme and that it took place in the first room—indicating that it was completed before the stress/control procedure. Additionally, they were reminded of how many problems they correctly solved in Task 1. The decision in Task 5 captures willingness to compete, but, since the decision is made for past performance which occurred outside the stress treatment, preferences for engaging in a competitive activity or (beliefs about) the potential negative effect of stress on performance should not be relevant.

In Task 6, *Choice of compensation scheme for past performance under treatment*,

subjects also split 100 points between the tournament and piece-rate schemes, but were paid according to performance in Task 2. Instructions for Task 6 reminded subjects of their performance in Task 2, that this task took place after the stress/control procedure and that it was incentivized with the piece-rate scheme. Therefore, if stress negatively impacts performance, and thus possibly changes subjective beliefs about relative performance, this should influence the subjects' decisions in both Task 4 and Task 6. However, preferences for engaging in a competitive activity, and (beliefs about) performance in tournaments under stress are only relevant in Task 4.

In Task 7, we measure risk preferences using a design based on [Dohmen et al. \(2010\)](#). Subjects made a series of choices between a lottery, which paid CZK 240 or 0 with 50% probability each, and a safe payment. The safe payment varied across choices, gradually increasing from CZK 0 to CZK 240 in steps of CZK 20.

To estimate the role of confidence in competitiveness decisions, we asked non-incentivized questions regarding subjects' perceived rank among all eight participants in the given session for Tasks 1-3.

To limit possible hedging, subjects were informed that only two out of the seven tasks (Task 1-Task 7) would be randomly selected for payment at the end of the experiment.

Full experimental instructions for Tasks 1-7 are available in the online appendix.

B. Treatments

We experimentally induced stress in the laboratory, using a modified version of the Trier Social Stress Test for Groups (TSST-G) ([Kirschbaum et al., 1993](#); [von Dawans et al., 2011](#)). This procedure was intended to induce psychosocial stress in the stress treatment, with a control procedure designed to similarly prime subjects yet to keep stress levels constant. The TSST-G has been shown to be the most efficient experimental method of inducing stress, as measured by cortisol response ([Dickerson and Kemeny, 2004](#)).

The stress treatment protocol consisted of two parts: a public speaking task and a mental arithmetic task. The first part took place immediately before Task 2 and the second part immediately before Tasks 4. The reason for this is to ensure that subjects were under sufficient stress throughout the relevant time period (See Figure 1). In order to minimize the time between the TSST-G and decisions, subjects completed the tasks at computers in the same room, immediately after they finished the stress (or control) procedure. The full protocol for the TSST-G is included in the online appendix.

In both parts of the TSST-T, subjects spoke one-by-one in front of a committee of 2 experimenters, who sat at a table in front of the participants wearing white lab coats.

In order to increase subjects' level of psychosocial stress, the committee did not give any feedback and maintained a neutral facial expression throughout the procedure. The setting of the room is depicted in Appendix Figure A.1. Subjects were separated by dividers and wore headphones with ambient traffic noise during the entire TSST-G procedure, except when speaking to the committee, in order to prevent subjects from hearing others during the stress procedure and potentially developing subjective rankings in ability.

In the public speaking task, subjects were told to imagine a situation in which they had been caught cheating during an important academic examination and that they should defend themselves in front of a disciplinary committee. This scenario required participants to talk extensively about their personal qualities, and they were instructed to do their best. They were interrupted and asked additional questions if they spoke too fluently for too long.

In the second portion of our modified TSST-G procedure, subjects in the stress treatment were again called individually and asked to recite the alphabet backwards in steps of two, starting from a given letter. For example, if given the letter Z, they were required to recite Z, X, V,... Subjects engaged in this activity for a minute and were corrected if a mistake was made.

Our version of the TSST-G changes the standard protocol in several ways. We modified the speaking task to avoid possible priming effects: the original procedure is framed as a job-interview, which could have influenced competitiveness and performance in the experiment independently of the stress reaction. In the second portion of the task, subjects were instructed to recite the alphabet rather than counting in intervals. Likewise, this was done to avoid contaminating performance in the experiment, while still allowing us to use the addition of two-digit numbers as the real-effort task, consistent with previous work.⁴

The control procedure similarly primed subjects, and involved a similar degree of physical activity, but in a less stressful setting. Subjects were asked to read an article about academic dishonesty, silently for the first fourteen minutes and then aloud for two minutes. In the second part of the procedure, they collectively recited the alphabet out loud for a minute. Two experimenters were again present in the room during the control procedure, but wore casual clothes and behaved naturally. The subjects in the control group also wore headphones with ambient noise and were separated with dividers, to mimic conditions in the stress treatment group.

⁴Additionally, to avoid deception, participants in the stress treatment were not told that the panel members were trained in behavioral analysis, or that the video recordings would later be analyzed, as in [von Dawans et al. \(2011\)](#).

C. *Sample and procedures*

The experiment was carried out in 2014-2015, with 24 sessions in total. Subjects were recruited using a standard recruitment database, ORSEE (Greiner, 2004); no details about the nature of the experiment were mentioned in the invitation in order to avoid self-selection based on relevant personal characteristics, such as aversion to stressful or competitive situations. The stress and control treatments were randomized at the session level, for logistical reasons. Each session consisted of eight subjects, four males and four females, and though the gender composition was not directly mentioned (following Niederle and Vesterlund, 2007), it was easily observable—at the end of the experiment, 80% of subjects correctly reported the gender ratio of the session. The final sample is composed of 95 male and 95 female subjects, primarily undergraduate students (82%), majoring mostly in economics, business and related fields (61%).⁵

Decisions were made on computers, using the program z-Tree (Fischbacher, 2007). The experiment was conducted in the Czech language and sessions were administered by one experimenter (male), one assistant (female) and two separate committee members for the TSST-G procedure (a male and a female). The average length of the experiment was slightly less than 2 hours and the average payoff was CZK 516 (EUR 20).

The study was approved by the Internal Review Board of the Laboratory of Experimental Economics in Prague, where the experiment took place. We obtained informed consent from all participants, emphasizing that they were free to leave at any time. At the end of the session subjects in the stress treatment were debriefed on the true purpose of the stress procedure.

3. Results

A. *Physiological stress response*

To confirm that the TSST-G test was successful in producing a stress response, we analyze the salivary cortisol samples taken throughout the experiment. Results are presented in Figure 2. Baseline cortisol was measured in sample 1, before the stress procedure, and samples 2 and 3 were taken afterwards (see Figure 1). Cortisol sample 2 was collected only after the second portion of the TSST-G procedure. However, the cortisol response is typically delayed by 15-20 minutes after initial exposure to the stressor (Kemeny, 2003; Allen et al., 2014), and therefore sample 2 primarily captures the physiological response

⁵We dropped one female subject who left the experiment early and one male who did not meet the selection criteria.

to the first part of the TSST-G procedure, whereas sample 3 reflects responses to both parts 1 and 2.

While cortisol levels for subjects in the control group actually slightly decrease over the course of the experiment, levels for those in the stress treatment group more than doubled. For men in the stress treatment, cortisol levels in samples 2 and 3 increased by 130 and 113 percent of baseline, respectively (signed-rank test, $p = 0.000$ for both). For women, there was an increase of 109 percent on average, which remained constant in samples 2 and 3 (signed-rank tests, $p = 0.000$ for both). While it is difficult to directly compare cortisol levels between men and women due to biological differences, we find no evidence that the TSST-G was relatively more successful in either the males or females ($p = 0.200$ for the percentage increase between samples 1 and 2 and $p = 0.407$ for the percentage increase between samples 1 and 3).⁶

The timing of the cortisol response is difficult to measure precisely. Given this, we analyze heart-rate data to confirm that the elevated cortisol that we observe in the stress treatment is indeed a result of the TSST procedure, and that subjects remained under heightened stress during the tasks. During part one of the TSST protocol, the heart rate of the stress treatment group increases sharply, and is significantly higher than in the control group ($p = 0.000$). It stays significantly higher during Task 2 (piece-rate under treatment), Task 3 (tournament under treatment) and during the willingness to compete decision in Task 4 (see Appendix Table A.1).

B. Performance and competitive incentives

We next analyze the effect of stress on performance under competition, and whether this differs by gender. Recall that Tasks 1-4 included a counting activity. Since Task 1 took place before the treatment and was incentivized using the piece-rate scheme, this serves as the baseline for ability and motivation. Performance in Task 2 (piece-rate payment, after treatment) isolates the effect of stress on performance, and Task 3 measure how both stress and competition affect performance. If competitive incentives lead to increased performance—a common assumption—then subjects should be expected to complete more problems in Task 3 than in Tasks 1 and 2.

Results from performance in the counting portions of Tasks 1-4 are presented in the upper panel of Figure 3 and in Table 1. Under the piece-rate incentive scheme in Task 1, there is virtually no difference in the number of correctly solved problems between the treatment and control groups ($p = 0.931$). This demonstrates that randomization was

⁶All tests of equality of means are Wilcoxon rank-sum tests, unless otherwise noted.

successful. The same holds for both the male and female sub-samples, independently.

In Task 2 we do not find a statistically significant difference between treatments: subjects in the stress and control treatments correctly answered 6.37 and 6.56 problems, respectively ($p = 0.560$). As before, this result holds for both the male and female subsamples. This indicates that stress alone does not affect performance in the counting task.

In contrast, for performance in the tournament in Task 3 we see a significant treatment effect, with lower performance among the stress group, who solved only 6.24 problems correctly ($sd = 2.98$), compared to 7.14 ($sd = 2.74$) in the control group ($p = 0.018$). This difference is driven by female subjects: women in the stress treatment correctly solved 5.23 ($sd = 2.43$) problems on average, compared to 6.60 ($sd = 2.08$) in the control group ($p = 0.003$). The corresponding treatment difference for men is less than one third the size, 0.41, and is not statistically significant ($p = 0.562$).

In Table 2 we confirm this pattern by regressing performance under tournament incentives in Task 3 on a dummy that equals 1 if the subject was assigned to the stress treatment, gender, and baseline performance in Task 1, with standard errors clustered at the session level.⁷ We find that the stress treatment lowers performance by 0.84 correctly answered questions on average ($p < 0.001$).⁸ In column 2, we add an interaction term, *stress treatment*female*, and the results indicate that the effect of the stress treatment is specific to the female sub-sample ($p = 0.019$). In columns 3-4, we estimate the effects separately for males and females: the stress treatment lowers female subjects' performance by 1.45 questions ($p < 0.001$), while the coefficient for male subjects does not differ statistically from zero ($p = 0.513$).

Next, we consider the average difference in the number of problems each subject correctly solved in Tasks 3 and 2. The lower panel of Figure 3 demonstrates that tournament incentives influence performance within individuals, across treatment and gender. Overall, subjects in the control group correctly solved 0.57 more problems under the competitive compensation scheme in Task 3 than under the piece-rate scheme in Task 2 (signed-rank test $p = 0.034$). This holds independently for both men and women, who answered 0.42 (signed rank test, $p = 0.042$) and 0.73 (signed-rank test, $p = 0.007$) more questions correctly in Task 3 than in Task 2, respectively.

However, in the stress treatment, only 25.5 percent of female subjects did better in

⁷We cluster standard errors at the session level to account for shared experiences and because treatment was assigned at the session level. This results in 24 clusters. We confirm that the small number of clusters does not affect results by running a robustness check using the wild bootstrap method proposed by [Cameron et al. \(2008\)](#). Results are available upon request.

⁸We run similar regressions on performance in Tasks 1 and 2, which confirm the lack of treatment effect we observe through rank-sum tests. Results are presented in Appendix Table A.2.

Task 3 than Task 2, compared to 56.3 percent in the control, and 44.7 percent of women in the stress treatment did *worse* in Task 3 than in Task 2, compared to only 20.8 percent of those in the control. These proportions differ significantly across treatments (Chi-squared test, $p = 0.007$). On average, women in the stress treatment solved 0.49 fewer problems in Task 3 than in Task 1, which is marginally insignificant (signed-rank test, $p = 0.122$). For men, there is no treatment difference in the proportion of subjects who improved or did worse in Task 3 compared to Task 2 (Chi-squared test, $p = 0.453$), and men in the treatment group solved 0.23 more problems on average under the competitive incentive scheme than they did under piece-rate incentives in Task 2, though this is not statistically different from zero (signed-rank, $p = 0.290$).

The regression results also show a clear pattern: in column 5 of Table 2, we regress the difference between correctly answered problems in Tasks 3 and 2, which can be interpreted as the effect of the tournament incentive scheme on performance, on stress treatment and gender. The results indicate that the stress treatment diminishes the positive effect of the tournament incentive scheme by 0.70 questions on average ($p = 0.008$). As before, in columns 6-8 we see that this is driven by the female sub-sample, and that there is no statistically significant effect for men ($p = 0.586$).⁹

Overall, these results indicate that female subjects perform significantly worse in a competitive setting when under increased psychosocial stress.¹⁰ While competition increases performance in the control group, this is not the case for female subjects exposed to increased psychosocial stress. We do not find that stress, on its own, has a negative effect on performance for either gender: neither men nor women in the stress treatment perform significantly worse in Task 2, compared to either their own Task 1 performance or to the Task 2 performance of subjects in the control group. Moreover, we find that both men and women in the control group respond positively to competitive incentives, as they perform significantly better in Task 3 than in Task 2. However, the combination of stress and competition decreases performance for a large portion of female subjects. We do not find any such pattern for men, whose performance under tournaments is not significantly

⁹In Appendix Table A.3 we confirm that these results are stable with respect to additional controls: they do not depend on whether we control for baseline performance in Task 1 and they hold when we control for baseline cortisol levels, baseline heart-rate levels, baseline mood, personality traits (BFI inventory), for potential problems with understanding, for whether women take oral contraceptives and for the phase of their menstrual cycle (which can affect cortisol levels).

¹⁰Note that we do not arbitrarily divide the sample by gender; the study was designed with the principle aim of studying differences in reaction to stress by gender (project proposal available upon request). Moreover, the competition task includes only 6 experimental outcomes (Performance in Tasks 2-4 and Investment decisions in Tasks 4-6), which indicates that the risk of false positives due to multiple hypothesis testing is low.

affected by the stress treatment.¹¹

C. *Willingness to compete*

We now turn to investment in the tournament payment scheme in Task 4, which serves as our principal measure of willingness to compete. This decision captures both preferences for competitive outcomes, as well as those for engaging in a competitive activity as well as expectations of one's future performance under competition. Recall that in Task 4 subjects allocated 100 points between a tournament and piece-rate incentive scheme before completing the counting portion of the task. The results from Task 4 are presented in Figure 4 and panel A of Table 3. Overall, subjects allocated slightly less than half of the total amount, 46.68 points, into the tournament incentive scheme. We find that stress does indeed affect competitiveness: subjects in the stress treatment invested 7.72 fewer points in the tournament scheme than those in the control group ($p = 0.046$).

We confirm this result by regressing the points invested into the tournament scheme in Task 4 on the stress treatment dummy. We control for gender as well as baseline performance in Task 1 (i.e. before the treatment intervention) and cluster standard errors at the session level. As reported in column 1 of Table 4 we find that the stress treatment was associated with investing 7.59 fewer points in the tournament scheme ($p = 0.024$). Consistent with the literature, we find that gender has a strong influence on choices in Task 4, with women investing 25.27 fewer points in the tournament investment scheme than men ($p = 0.000$). This is confirmed by the regression results in column 1 of Table 4, in which we observe that women invested on average 22.06 fewer points, after controlling for treatment and baseline performance ($p = 0.000$).

The stress treatment has a similar effect on willingness to compete in both men and women. The negative effect of stress on investments in the tournament that we find on average in Task 4 holds separately for both the male and female sub-samples, though the treatment differences are not statistically significant, due to smaller sample sizes (Figure 4 and panel A of Table 3). In column 2 of Table 4, we add an interaction term between the female and stress treatment dummies and observe no statistically significant gender difference ($p = 0.926$). In columns 3-4, we run regressions separately on the male and female sub-samples and find that the coefficients for the stress treatment are virtually

¹¹We also consider performance in Task 4, though interpretation is less clear, since the incentive scheme is endogenous. On average, the stress treatment does not have a statistically significant effect on the number of correctly answered questions in Task 4 ($p = 0.650$). However, women in the stress treatment correctly completed 0.5 fewer correct problems ($p = 0.087$). Regression results confirm this; see Appendix Table A.4.

identical, though both coefficients are marginally insignificant: $p = 0.123$ and $p = 0.124$ for the male and female sub-samples, respectively.

Since most studies use a binary measure to measure willingness to compete, we perform a robustness test in which we classify subjects as competitive if they invest at least 50/100 points into the tournament incentive scheme in Task 4 and estimate the effects of the stress treatment and gender using a probit model. Results are similar to the linear measure (see Appendix Table A.5). We also confirm that results are robust to including additional controls (see Appendix Table A.6).

The decisions in Tasks 5 and 6 provide further insight into the mechanism behind the treatment effect we find for Task 4, which measures willingness to compete for future performance. In Task 5, subjects decided how much to invest in the tournament payment scheme based on the result of their performance in the counting portion of Task 1 (i.e. under the piece-rate payment scheme and before the stress/control treatment). In contrast to the competition decision in Task 4, we do not find a significant difference between the treatment groups for investment in the tournament in Task 5. On average, subjects in the control group invested 40.19 versus 41.20 in the stress treatment group ($p = 0.826$, see Panel B of Table 3). We do not find a statistically significant treatment difference for either men or women.

In Task 6, subjects also decided how much to invest in the tournament for past performance, this time based on the results from Task 2 (piece-rate, after the stress/control treatment). As in Task 5, we do not find a statistically significant difference in willingness to compete between treatments. Subjects in the control group invested 41.14 points into the tournament, while those in the stress treatment invested 39.64 points on average ($p = 0.702$). Results are presented in panel C of Table 3.¹²

Since Task 2 was completed after the stress treatment, changes in performance or perceived relative performance in response to the stressor should affect competitiveness in Tasks 4 and 6 similarly. The lack of treatment difference in Task 6 thus suggests that the difference in competitiveness we see in Task 4 is not caused by a difference in perceived ability as a result of the stress treatment alone. Together, the results from Tasks 5 and 6 indicate that the decrease in competitiveness that we see in the stress treatment in Task 4 is related to completing the task both under stress and in a competitive setting, rather than either element alone.

The gender difference in competitiveness that we observe in both treatments in Task 4 holds in Tasks 5 and 6 as well.

¹²We present regression results for investment in the tournament incentive scheme in Tasks 5 and 6 in Appendix Table A.7. We do not find any significant treatment effects for either men or women.

4. Discussion and Additional Results

In this section, we perform a series of robustness checks and additional analysis, with the goal of determining the mechanisms that drive our results.

A. *Physical stress response*

To begin, we relate the physiological stress response, measured as a relative increase in the salivary cortisol levels between the first and the second sample, to both willingness to compete in Task 4 and to tournament performance in Task 3. We examine the correlation between cortisol response and the outcome and estimate the average treatment effect on the treated, using the stress treatment as an instrument. Results are robust: the stronger the physiological stress response, the lower the willingness to compete and the worse the tournament performance for women (See Appendix Table A.8 and Table A.9). This is evidence that the stress treatment indeed affects behavior through stress, rather than some other channel.

B. *Stress and willingness to compete*

Next, we consider potential channels through which stress might lead to lower willingness to compete and, particularly for women, how this is related to performance under tournament incentives.

The first mechanism we consider is a change in preferences under heightened stress. Our design allows us to distinguish between willingness to compete for future performance and willingness to compete for past performance. Since we find a treatment difference only for the decision in Task 4 (future performance) but not in Tasks 5 or 6 (past performance), our results seem to rule out an effect of stress on preferences for competitive outcomes. This result is in line with [Goette et al. \(2015\)](#) who also find no effect of stress on competitiveness for past performance.¹³ This is in contrast to gender differences: consistent with findings in [Niederle and Vesterlund \(2007\)](#), we find that women are less competitive across all three investment decisions.

Second, our finding that women in the stress treatment perform worse under tournament incentives in Task 3 suggests one of two closely related underlying effects (or a combination thereof): stress may affect preferences for engaging in competition, which may in turn lower effort, or stress may lower the ability of women under tournament

¹³See Table 2 in [Goette et al. \(2015\)](#). Their principal finding is that stress affects low- and high-anxiety individuals differently. Task 5 of our experiment is similar to their design; we do not replicate this result and find no significant interaction effect between stress and trait anxiety. (Available upon request.)

incentives. Regardless, our results suggest that women in the stress treatment have a lower willingness to compete due to weaker performance in the tournament. Even though subjects were unaware of the number of questions that others correctly answered, they observed their own performance under both the piece-rate and tournament compensation schemes, and could have based their investment in the tournament on their relative performance in these rounds. To test this, we conduct a sensitivity analysis, which is reported in Appendix Table A.11. The results strongly suggest that for women stress affects willingness to compete in Task 4 principally by affecting performance, and that after controlling for Task 3 performance the stress treatment adds no explanatory power to the regression model.

Alternatively, stress could affect subjective beliefs about performance. To this end, we measured subjective confidence for each round—after the experiment—by eliciting beliefs about rank among the 8 subjects in the session, for Tasks 1-3. Appendix Figure A.2 confirms that beliefs are highly correlated with the performance results: women in the stress treatment have lower confidence in their performance in tournaments compared to women in the control group, but there is no significant difference in confidence in performance under piece-rate incentives, either before or after the stress/control manipulation. Stress has no effect on men’s confidence. In Appendix Table A.12, we conduct a similar sensitivity analysis and find that the stress treatment does not additionally explain (subjective) confidence in tournament performance after controlling for actual performance.

Another possibility is that stress influences competitiveness through risk preferences. [Cahlíková and Cingl \(2017\)](#) find that a similar version of the TSST-G leads to higher levels of risk aversion, especially for men. Since the tournament incentive scheme increases subjects’ exposure to risk, greater risk aversion might lead to lower willingness to compete. However, a change in risk-preferences would also affect willingness to compete for past performance, and we do not observe any effect in Tasks 5-6. Moreover, in our sample, we fail to find any significant relationship between the stress treatment and risk preferences elicited in Task 7. In fact, those in the stress treatment actually had slightly higher certainty equivalents than those in the control group on average, though the difference is not statistically significant ($p = 0.334$). This is consistent within both the male and female sub-samples, ($p = 0.501$ and $p = 0.698$, respectively).¹⁴ These results suggest that risk preferences are not a mechanism by which stress affects willingness to compete

¹⁴Since we randomly chose two of the seven tasks (1-7) for payment, it is possible that decisions in the risk task were affected by decisions in previous rounds. As stress lowered willingness to compete, leading to lower risk exposure in Task 3, this may have caused subjects in the stress treatment to make riskier decisions in Task 7, independent of risk preferences. Therefore, our measure of risk preferences should be interpreted with caution.

in our sample.

To summarize, for men, we do not find evidence that lower ex-ante willingness to compete is related to a change in performance, confidence, risk-preferences, or understanding, and by process of elimination we conclude that lower willingness to compete among men is driven by preferences for engaging in competition under stress. For female subjects, we conclude that psychosocial stress lowers willingness to compete because women under stress perform worse under tournament incentives. Based on this, women react by investing less in the tournament incentive scheme when given a choice.

C. Stress and tournament performance

Why do tournament incentives lead to poorer performance for women in the stress treatment? Holding preferences and beliefs stable, under tournament incentives individuals should exert more effort than in the piece-rate scheme, as long as their marginal cost of effort is sufficiently low and they believe that their relative ability is sufficiently high. This seems to be the case for men and women in the control group, who perform better in the tournament. However, there are also several reasons why performance might decrease under competition. We now consider how these factors might be related to stress and gender, and whether they are plausible explanations for our results.

First, tournament incentives might induce additional stress and there could be a threshold level of stress beyond which performance suffers. However, this seems unlikely to explain our findings. While cortisol levels increase sharply after the TSST-G manipulation for subjects in the stress treatment, the cortisol levels of subjects in the control group continue to decrease over the course of the experiment, despite the fact that control subjects also take part in the tournament. It is possible that cortisol would have decreased by even more in the absence of a tournament, but in any case, this suggests that the effect of the tournament on cortisol—and thus stress—is small. [Buser et al. \(2016\)](#), using a similar design as ours, find that cortisol levels increase on average by 3-5% when subjects perform a task compensated with a piece-rate scheme and by 12-15% when they perform under tournament incentives. In comparison, we find that the TSST-G protocol increases cortisol levels by 109% in women and 130% in men. Additionally, the heart-rate data from our study, which provides a more time-specific measure of stress levels, shows that the response of the stress group to the treatment is much stronger than the control group's response to the tournament in Task 3 (Appendix Table A.13). Any such stress threshold would therefore have to be extremely sensitive in order to explain the effect we observe. Moreover, it would have to be specific to women, as men's performance in the tournament

is not affected by the stress treatment.

Second, potentially subjects in the stress treatment were under increased cognitive load as a result of the TSST-G procedure, and this affected either their ability to concentrate on the task or their motivation to perform. If this were the case though, we would also expect to see a difference in Task 2 performance, since this also occurred after the stress treatment. To further rule out this possibility, we run two robustness checks. We had subjects complete a d2 attention test ([Brickenkamp and Zillmer, 1998](#)) at the end of the experiment and find no treatment differences.¹⁵ Subjects were also asked to rate their understanding of the experimental instructions in the questionnaire at the end of the session and we find no significant treatment difference ($p = 0.291$). The treatment difference in Task 3 performance remains significant after dropping all women who reported less than perfect understanding ($n=53$, $p=0.016$).

Third, the results might be explained by the higher stakes inherent to competition. For example, if a subject was matched with a partner in Task 3 who correctly solved the median number of 6 problems, the difference in payoffs between solving 5 and 7 problems would be CZK 350. In the piece-rate tasks, the same 2-problem difference in performance would only change payoff by CZK 50. This is relevant because large stakes have been shown to have a detrimental effect on performance, causing individuals to “choke” under increased pressure. This is demonstrated by a series of experiments in [Ariely et al. \(2009\)](#), who link results to the [Yerkes and Dodson \(1908\)](#) law, a long-standing principle from the psychology literature according to which arousal increases executive function up to a point, but the relationship is defined by an inverted U-shaped function, and ability declines with increased stimulation after passing a threshold. In other words, performance is worse when stakes are either too high or too low. If psychosocial stress also stimulates subjects, the Yerkes-Dodson law may explain why stress and competition lower performance in combination, but not individually. There is emerging evidence that women might be more sensitive to this effect: [Azmat et al. \(2016\)](#), using a data from non-competitive exams, show that increasing the stakes hurts women’s performance, but not men’s.

A fourth plausible explanation is that tournament incentives accentuate social comparison, which can be detrimental to performance, as shown by [Ashraf et al. \(2014\)](#) and [Schram et al. \(2015\)](#), who find that the anticipation of being ranked by status leads to

¹⁵In the d2 attention test, subjects’ selective attention and concentration are tested by presenting them with a series of letters and instructing them to identify any letter “d” with two adjacent marks. We find no treatment difference in the number of correctly marked characters for either men ($p = 0.710$) or women ($p = 0.835$) We also find no treatment difference for men or women in the number of attempted problems or error rate.

worse performance among women, while men actually perform better. Since the counting task in our study might be perceived as a male-dominated activity, it plausibly produces a “stereotype threat” among female subjects, in which they are confronted with perceived negative stereotypes about women and math (Spencer et al., 1999). Even though we do not observe negative effects of tournament incentives in the control group, stereotype threat might further lower executive function already challenged by stress—including working memory, which is crucial for solving the addition problems—resulting in worse performance according to the Yerkes-Dodson law. Schmader et al. (2008) provide a theoretical framework for understanding how stereotype threat (social identity threat) affects performance, and argue that this could be exacerbated by increased cortisol, and “physiological stress response could play a direct role in impairing task performance under stereotype threat” (p.343). Our results are in line with this hypothesis, which offers an explanation as to how the combination of higher stakes in the tournament and psychosocial stress could lower performance for women, who plausibly face a stereotype threat in the counting task, but not for men.

5. Conclusion

This article presents new evidence on the effects of stress on performance under competition and individual willingness to compete. We experimentally induce psychosocial stress in the laboratory using a modified TSST-G protocol and find that subjects in the stress treatment group are less competitive, investing less in the tournament compensation scheme than those in the control group. However, this is only true when the willingness to compete decision is made before the competitive task. By examining salivary cortisol, which increases after exposure to the stress treatment, we confirm that the treatment difference in the willingness to compete is driven by stress; the cortisol response is negatively correlated with willingness to compete. In the tasks for which subjects made willingness-to-compete decisions for past performance, we find no treatment effect.

For women, we find that performance under competition is worse in the stress treatment than in the control group. While female subjects in the control group perform significantly better under tournament incentives than under the piece-rate scheme, female subjects in the stress treatment actually do slightly worse. It is the combination of stress and tournament incentives which is detrimental to performance, and for women, this explains the lower willingness to compete we observe in the stress treatment group. We do not find such a link among men, for whom there is no treatment difference in tournament performance. The lower willingness to compete among men in the stress treatment seems

to be driven by a link between stress and preferences for engaging in competition.

We propose that the most plausible explanation for decreased performance under stress and competition among women is that the counting task was perceived as a male-dominant activity, and women therefore faced a stereotype threat. Stress, higher stakes in the tournament and stereotype threat are all factors linked to executive function in the Yerkes-Dodson framework, and the combination of all three could decrease certain abilities, including impairment of working memory, thus affecting performance in the counting task (Schmader et al., 2008).

Our findings help to explain past results regarding the effect of tournament incentives on performance. While some studies have found a positive effect for both genders, others show a positive effect only for men (Niederle and Vesterlund, 2007; Gneezy et al., 2003; Gneezy and Rustichini, 2004). Potentially, the environments differed in the degree of stress involved. Moreover, our results also support the claim made in Niederle and Vesterlund (2010) that gender gaps in math test scores may not necessarily reflect differences in math ability. Especially when test results come from highly-competitive and stressful settings, such as university entrance exams, women's performance may fall below their ability.

While a gender-neutral or female-dominated task might produce different results, the labor market settings that motivate this research, such as business or academia, are similarly perceived as stereotypically male. Many competitive situations that affect one's career trajectory—exams, job interviews and asking for a promotion—are also stressful. If women perform worse under competition and stress, this will directly affect labor market outcomes, and could dissuade women from entering competitive environments in the first place.

If employers make hiring decisions in stressful, competitive settings, our results suggest that this will lead to inefficient outcomes as women may under-perform. This is especially relevant if the position itself is not particularly stressful or competitive, compared to the hiring process. Moreover, a combination of tournament incentives and increased pressure to perform are often used to boost output in firms, and while this incentive structure may be effective in motivating male workers, our results show that such policies could have unintended consequences when applied to female workers.

This phenomenon could produce path-dependence across sectors: if the initial composition of a particular field is dominated by males or females, the optimal management and hiring strategies might differ with respect to stress and competition. Both hiring and self-selection of workers into the field would then exacerbate gender disparities in response to the dominant practices and norms.

Furthermore, in organizations where competition is inherent, as in many firms or universities, identifying and taking measures to reduce stress could be a viable policy for improving performance, as (especially female) workers might be performing below their potential.

References

- Allen, A. P., P. J. Kennedy, J. F. Cryan, T. G. Dinan, and G. Clarke (2014). Biological and psychological markers of stress in humans: Focus on the Trier Social Stress Test. *Neuroscience and Biobehavioral Reviews* 38, 94–124.
- Ariely, D., U. Gneezy, G. Loewenstein, and N. Mazar (2009). Large Stakes and Big Mistakes. *Review of Economic Studies* 76(2), 451–469.
- Ashraf, N., O. Bandiera, and S. Lee (2014). Awards unbundled: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization* 100.
- Azmat, G., C. Calsamiglia, and N. Iriberry (2016). Gender Differences in Response to Big Stakes. *Journal of the European Economic Association* 14(6), 1372–1400.
- Baum, A. and N. E. Grunberg (1997). Measuring Stress Hormones. In S. Cohen, R. C. Kessler, and L. U. Gordon (Eds.), *Measuring Stress: A Guide for Health and Social Scientists*, pp. 175–192. Oxford: Oxford University Press.
- Brickenkamp, R. and E. Zillmer (1998). *The d2 test of attention*. Hogrefe & Huber Pub.
- Buckert, M., C. Schwieren, B. M. Kudielka, and C. J. Fiebach (2015). How stressful are economic competitions in the lab? An investigation with physiological measures. *Discussion Paper Series, University of Heidelberg, Department of Economics* (No. 0593).
- Buser, T., A. Dreber, and J. Mollerstrom (2016). The Impact of Stress on Tournament Entry. *Experimental Economics* forthcoming.
- Buser, T., M. Niederle, and H. Oosterbeek (2014). Gender, Competitiveness and Career Choices. *Quarterly Journal of Economics* 129(3), 1409–1447.
- Cahlíková, J. and L. Cingl (2017). Risk preferences under acute stress. *Experimental Economics* 20(1), 209–236.
- Cameron, A., J. Gelbach, and D. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics* 90(3), 414–427.
- Cannon, W. B. (1932). The wisdom of the body. *American Journal of the Medical Sciences* 184(6), 864.

- Card, D., A. R. Cardoso, and P. Kline (2016). Bargaining, sorting and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *Quarterly Journal of Economics* (2), 633–686.
- Cardoso, A. R., P. Guimarães, and P. Portugal (2016). What drives the gender wage gap? a look at the role of firm and job-title heterogeneity. *Oxford Economic Papers* 68(2), 506–524.
- Dickerson, S. S. and M. E. Kemeny (2004). Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin* 130(3), 355–391.
- Dohmen, B. T., A. Falk, D. Huffman, and U. Sunde (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review* 100(3), 1238–1260.
- Everly, G. S. and J. M. Lating (2013). The Anatomy and Physiology of the Human Stress Response. In *A Clinical Guide to the Treatment of the Human Stress Response*, pp. 17–51. Springer New York.
- Fischbacher, U. (2007). Z-TREE: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2), 171–178.
- Fletcher, T. D., D. Major, and D. D. Davis (2008). The interactive relationship of competitive climate and trait competitiveness with workplace attitudes, stress, and performance. *Journal of Organizational Behavior* 29(7), 899–922.
- Flory, J. A., A. Leibbrandt, and J. A. List (2015). Do Competitive Work Places Deter Female Workers? A Large-Scale Natural Field Experiment on Gender Differences in Job-Entry Decisions. *Review of Economic Studies* 82(1), 122–155.
- Gneezy, U., M. Niederle, and A. Rustichini (2003). Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics* 118(3), 1049–1074.
- Gneezy, U., A. Pietrasz, and S. Saccardo (2016). On the size of the Gender Difference in Competitiveness. *Management Science* forthcoming.
- Gneezy, U. and A. Rustichini (2004). Gender and Competition at a Young Age. *American Economic Review* 94(2), 377–381.

- Goette, L., S. Bendahan, J. Thoresen, F. Hollis, and C. Sandi (2015). Stress pulls us apart : Trait anxiety modulates self-confidence under stress. *Psychoneuroendocrinology* 54, 115–123.
- Greiner, B. (2004). An online recruitment system for economic experiments. In K. Kremer and V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen GWDG Bericht* 63, pp. 79–93. Gottingen: Ges. für Wiss. Datenverarbeitung.
- Haushofer, J. and C. Jang (2015). Stress and Temporal Discounting: Do Domains Matter? *Unpublished*.
- Hellhammer, D. H., S. Wüst, and B. M. Kudielka (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology* 34, 163–171.
- Jurajda, S. and D. Munich (2011). Gender Gap in Performance under Competitive Pressure: Admissions to Czech Universities. *American Economic Review: Papers & Proceedings* 101(3), 514–518.
- Kemeny, M. E. (2003). The Psychobiology of Stress. *Current Directions in Psychological Science* 12(4), 124–129.
- Kirschbaum, C., K. M. Pirke, and D. H. Hellhammer (1993). The Trier Social Stress Test - A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology* 28(1-2), 76–81.
- McEwen, B. S. (2012). Brain on stress: how the social environment gets under the skin. *Proceedings of the National Academy of Sciences of the United States of America* 109(Supplement 2), 17180–5.
- Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122(3), 1067–1101.
- Niederle, M. and L. Vesterlund (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of Economic Perspectives* 24(2), 129–144.
- Niederle, M. and L. Vesterlund (2011). Gender and Competition. *Annual Review of Economics* 3(1), 601–630.
- Ors, E., F. Palomino, and E. Peyrache (2013). Performance gender gap: Does competition matter? *Journal of Labor Economics* 31(3), 443–499.

- Schmader, T., M. Johns, and C. Forbes (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review* 115(2), 336–356.
- Schram, A., J. Brandts, and K. Gerxhani (2015). Status Ranking: A Hidden Channel to Gender Inequality under Competition. *Barcelona GSE Working Paper No. 817*.
- Spencer, S. J., C. M. Steele, and D. M. Quinn (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology* 35(1), 4–28.
- Starcke, K. and M. Brand (2012). Decision making under stress: a selective review. *Neuroscience and Biobehavioral Reviews* 36(4), 1228–48.
- Stroud, L. R., P. Salovey, and E. S. Epel (2002). Sex differences in stress responses: social rejection versus achievement stress. *Biological Psychiatry* 52(4), 318–327.
- Taylor, S. E. (2006). Tend and Befriend: Biobehavioral Bases of Affiliation Under Stress. *Current Directions in Psychological Science* 15(6), 273–277.
- Taylor, S. E., L. C. Klein, B. P. Lewis, T. L. Gruenewald, R. A. Gurung, and J. A. Updegraff (2000). Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychological Review* 107(3), 411.
- von Dawans, B., U. Fischbacher, C. Kirschbaum, E. Fehr, and M. Heinrichs (2012, June). The social dimension of stress reactivity: acute stress increases prosocial behavior in humans. *Psychological Science* 23(6), 651–60.
- von Dawans, B., C. Kirschbaum, and M. Heinrichs (2011). The Trier Social Stress Test for Groups (TSST-G): A new research tool for controlled simultaneous social stress exposure in a group format. *Psychoneuroendocrinology* 36(4), 514–522.
- Yerkes, R. M. and J. D. Dodson (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology* 18(5), 459–482.
- Zhong, S., I. Shalev, D. Koh, R. P. Ebstein, and S. H. Chew (2016). Competitiveness and Stress. *Unpublished*.

FIGURE 1: TIMELINE OF THE EXPERIMENT

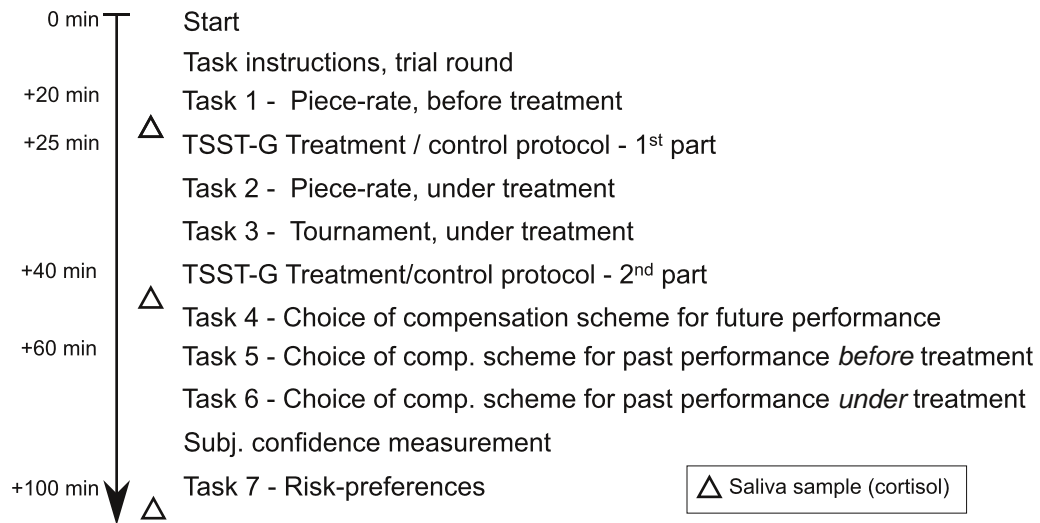
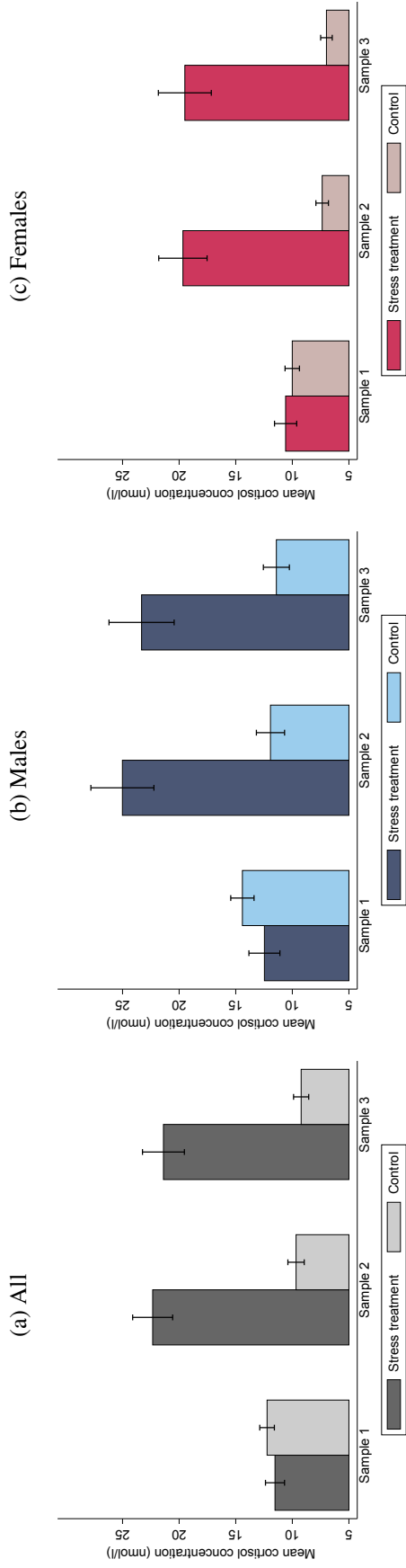
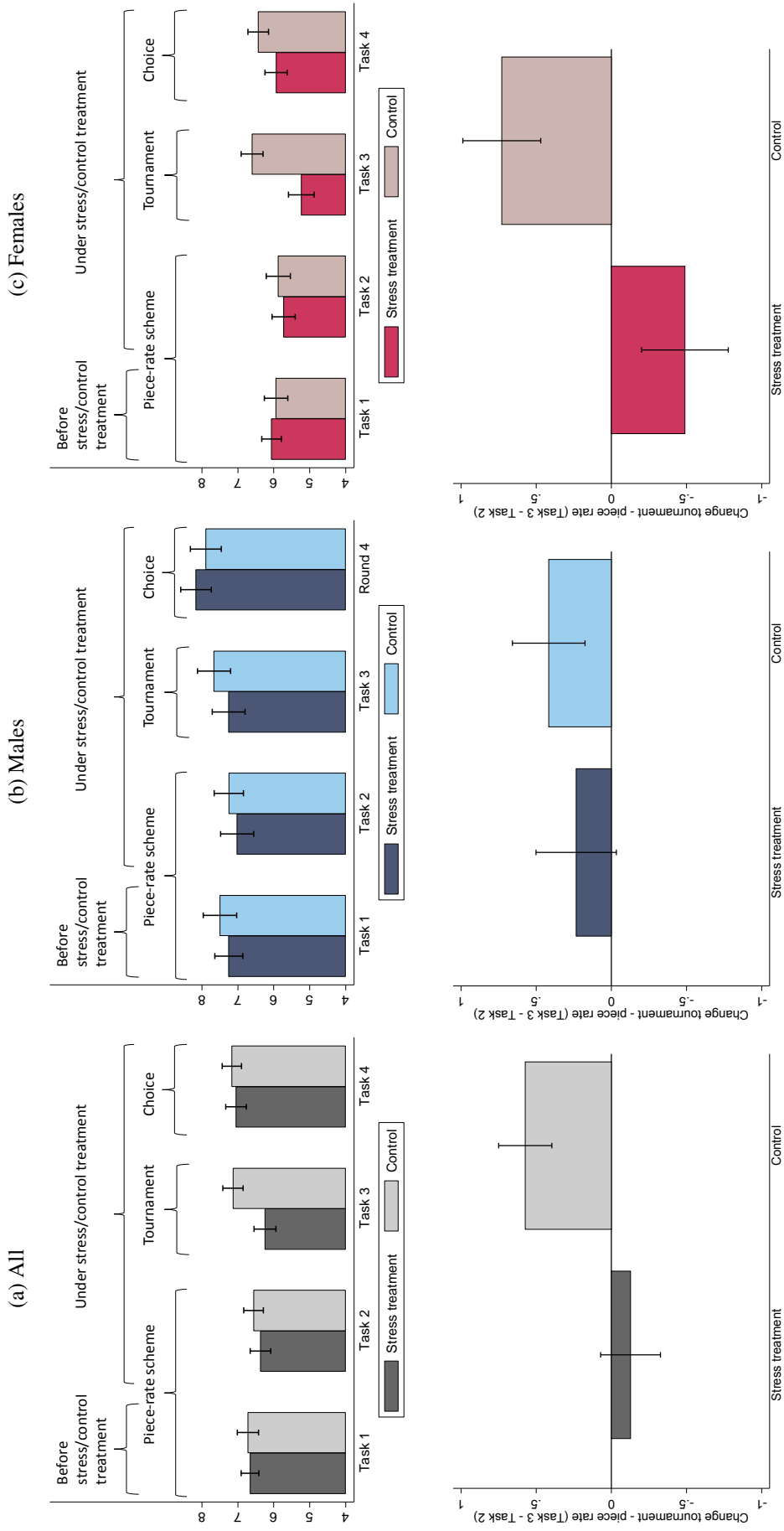


FIGURE 2: CORTISOL RESPONSE, BY TREATMENT



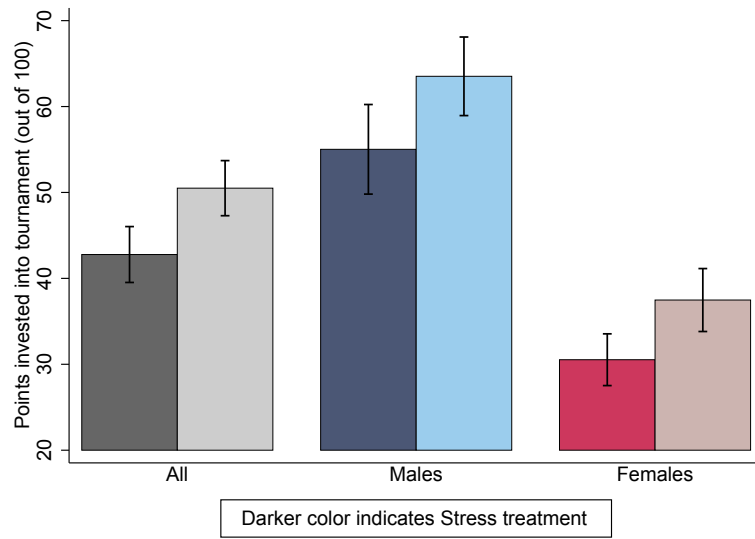
Notes: Mean salivary cortisol concentration (in nmol/l) over the course of experiment, by treatment. “Stress treatment” indicates that the TSSST-G stress procedure. Sample 1 was collected prior to the TSSST-G stress/control procedure. Sample 2 was collected after the second part of the TSSST-G protocol (the counting task) and Sample 3 was collected at the end of the experiment. For details see Figure 1. Bars indicate mean \pm standard error.

FIGURE 3: PERFORMANCE UNDER TOURNAMENT AND PIECE-RATE INCENTIVE SCHEMES, BY TREATMENT AND GENDER



Notes: Performance in the experiment, by gender. “Stress treatment” indicates that the subject was exposed the TSST-G stress procedure. Task 1 is the baseline conducted before the treatment, with piece-rate incentives. Task 2 and Task 3 took place after the first part of the TSST-G, under piece-rate and tournament incentives, respectively. Task 4 took place after the second part of the TSST-G, after the subjects had chosen their preferred incentives scheme for that round. Graphs in the bottom panel depict the change between Task 3 and Task 2, which captures the effect of tournament incentives by treatment. Bars indicate mean \pm standard error.

FIGURE 4: WILLINGNESS TO COMPETE, BY TREATMENT



Notes: Mean willingness to compete, by treatment. Willingness to compete is the investment in the tournament in Task 4; 0 indicates all points invested in the piece-rate scheme, 100 indicates all points invested in the tournament. Bars indicate mean \pm standard error.

TABLE 1: SUMMARY STATISTICS, PERFORMANCE IN THE COUNTING TASK

<i>Sample</i>	<i>Number of problems solved correctly</i>				<i>Ranksum (p-value)</i>	<i>N</i>
	<i>Treatment group</i>			<i>Diff.</i>		
	<i>All</i>	<i>Stress</i>	<i>Control</i>			
(1)	(2)	(3)	(4)	(5)		
<i>Task 1 - Piece rate before treatment</i>						
All	6.69	6.66	6.72	-0.06	0.931	190
Male	7.38	7.26	7.50	-0.24	0.943	95
Female	6.00	6.06	5.94	0.13	0.880	95
<i>Task 2 - Piece rate under treatment</i>						
All	6.47	6.37	6.56	-0.19	0.560	190
Male	7.14	7.02	7.25	-0.23	0.761	95
Female	5.80	5.72	5.88	-0.15	0.632	95
<i>Task 3 - Tournament under treatment</i>						
All	6.69	6.24	7.14	-0.89	0.018	190
Male	7.46	7.26	7.67	-0.41	0.562	95
Female	5.93	5.23	6.60	-1.37	0.003	95
<i>Task 4 - Chosen scheme under treatment</i>						
All	7.11	7.05	7.17	-0.11	0.650	190
Male	8.03	8.17	7.90	0.27	0.538	95
Female	6.19	5.93	6.44	-0.50	0.087	95

Notes: Mean performance in the tasks under different compensation schemes, by treatment and gender. “Stress treatment” indicates that the subject was exposed to the TSST-G stress procedure. All differences are tested using a Wilcoxon rank-sum test.

TABLE 2: THE EFFECT OF STRESS, COMPETITION AND GENDER ON PERFORMANCE

Dep. Variable	Problems Solved Correctly								
	Task 3				Task 3 - Task 2				
	Tournament Incentives		Effect of Tournament Incentives		All		Males		Females
Incentive scheme	All	Males	Females	All	Males	Females	All	Males	Females
Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(8)
Stress treatment	-0.84*** (0.21)	-0.21 (0.29)	-0.19 (0.28)	-1.45*** (0.34)	-0.70*** (0.24)	-0.17 (0.30)	-0.15 (0.28)	-1.21*** (0.33)	
Female	-0.41 (0.27)	0.22 (0.34)			-0.14 (0.28)	0.39 (0.36)			
Stress treatment*Female		-1.26** (0.50)				-1.05** (0.43)			
Solved Task 1 (baseline)	0.81*** (0.04)	0.82*** (0.04)	0.91*** (0.05)	0.63*** (0.12)	0.04 (0.06)	0.05 (0.06)	0.11** (0.05)	-0.09 (0.10)	
Constant	1.88*** (0.36)	1.53*** (0.38)	0.85 (0.50)	2.85*** (0.75)	0.34 (0.48)	0.05 (0.51)	-0.44 (0.52)	1.24** (0.53)	
Observations	190	190	95	95	190	190	95	95	95
R-squared	0.61	0.62	0.72	0.39	0.04	0.06	0.04	0.10	0.10

Notes: OLS. Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the number of addition problems that were correctly completed within the time limit in the specified task. Both Task 2 and Task 3 were completed under treatment, under piece rate and tournament incentives, respectively. "Stress treatment" is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE 3: SUMMARY STATISTICS, WILLINGNESS TO COMPETE BY TREATMENT AND TASK

<i>Sample</i>	<i>Number of points invested into competition (out of 100)</i>					<i>N</i>
	<i>All</i>	<i>Treatment group</i>		<i>Diff.</i>	<i>Ranksum</i> (p-value)	
		<i>Stress</i>	<i>Control</i>			
(1)	(2)	(3)	(4)	(5)	(6)	
Panel A: Task 4 Choice - Future performance under treatment						
All	46.68	42.78	50.50	-7.72	0.046	190
Male	59.32	55.02	63.52	-8.50	0.190	95
Female	34.04	30.53	37.48	-6.95	0.159	95
Panel B: Task 5 Choice - Past performance before treatment						
All	40.69	41.20	40.19	1.01	0.826	190
Male	47.85	48.21	47.50	0.71	0.988	95
Female	33.53	34.19	32.88	1.32	0.688	95
Panel C: Task 6 Choice - Past performance under treatment						
All	40.39	39.64	41.14	-1.50	0.702	190
Male	51.09	51.60	50.60	0.99	0.967	95
Female	29.69	27.68	31.67	-3.99	0.710	95

Notes: Mean decisions regarding willingness to compete, across tasks, treatments, and gender. Panel A presents the competitiveness decision for future performance, while Panels B and C present the two competitiveness for past performance decisions. “Stress treatment” indicates that the subject was exposed to the TSST-G stress procedure. All differences are tested using a Wilcoxon rank-sum test.

TABLE 4: THE EFFECT OF STRESS AND GENDER ON WILLINGNESS TO COMPETE

Dep. Variable <i>Sample</i>	<i>Willingness to Compete</i> <i>points invested into tournament ex ante (Task 4)</i>			
	<i>All</i>		<i>Males</i>	<i>Females</i>
	(1)	(2)	(3)	(4)
Stress treatment	-7.59** (3.14)	-7.93 (5.11)	-7.99 (5.00)	-7.31 (4.58)
Female	-22.06*** (4.14)	-22.41*** (6.23)		
Stress treatment*Female		0.69 (7.38)		
Solved Task 1 (baseline)	2.33** (0.95)	2.32** (0.95)	2.07* (1.15)	2.85** (1.34)
Constant	45.89*** (8.05)	46.08*** (8.33)	48.02*** (9.53)	20.55** (8.06)
Observations	190	190	95	95
R-squared	0.21	0.21	0.05	0.09

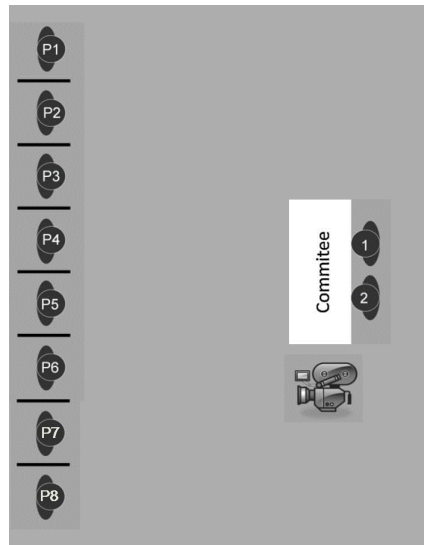
Notes: OLS. Standard errors are clustered at a session level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is investment in the tournament compensation scheme in Task 4, where the choice was made before the counting portion of the task. 0 indicates all points invested in the piece-rate scheme; 100 indicates all points invested in the tournament compensation scheme. “Stress treatment” is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

Appendix

Appendix Figures A.1-A.2 and Appendix Tables A.1-A.13 are below. Experimental instructions for Tasks 1-7 and the protocol for the TSST-G stress/control procedure are included in the Online Appendix.

FIGURE A.1: TSST-G STRESS PROCEDURE: SETTING OF THE ROOM

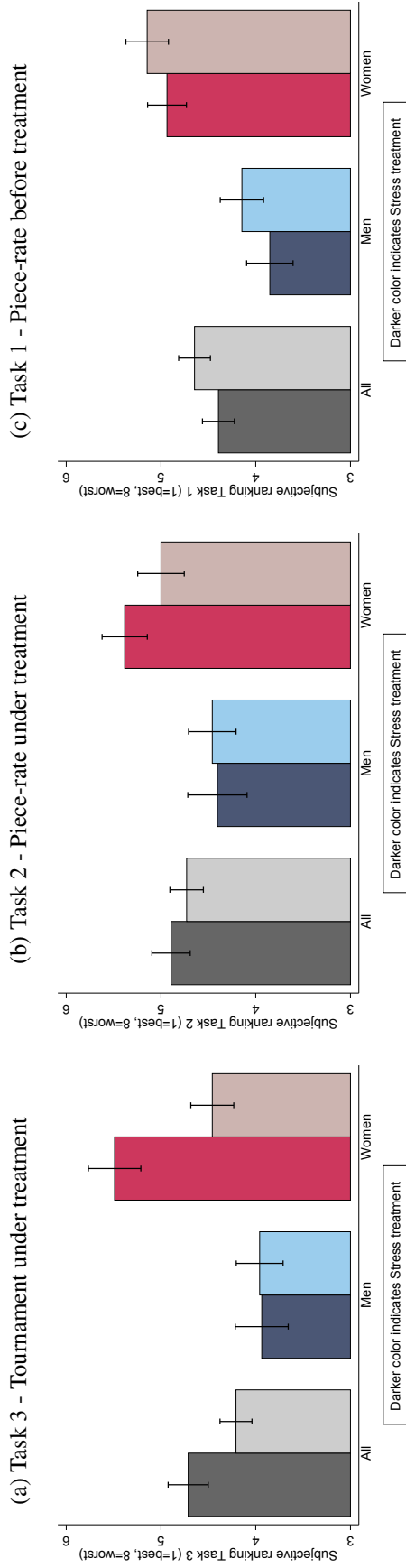
(a) Schema of the room



(b) Working stations of subjects in the room



FIGURE A.2: CONFIDENCE: BELIEFS ABOUT SUBJECT'S PERCEIVED RANK IN THE SESSION, BY TREATMENT AND TASK



Notes: Mean beliefs about subject's rank among the 8 subjects in the session, by task and treatment (i.e. 1=highest confidence, 8=lowest confidence). "Stress treatment" indicates that the subject was exposed to the TSST-G stress procedure. Task 3 was completed under tournament incentives under the stress/control treatment, Task 2 was conducted under piece-rate incentives under treatment and Task 1 was completed under piece-rate incentives before the stress/control manipulation. Confidence questions were non-incentivized and were elicited after Task 6. For details please consult the timeline in Figure 1. Bars indicate mean \pm standard error.

TABLE A.1: SUMMARY STATISTICS, HEART RATE DURING THE EXPERIMENT

<i>Sample</i>	<i>Mean heart rate during the tasks</i>					N
	<i>All</i>	<i>Treatment group</i>		<i>Diff.</i>	<i>Ranksum</i> (p-value)	
		<i>Stress</i>	<i>Control</i>			
(1)	(2)	(3)	(4)	(5)		
<i>Task 1 - Piece-rate before treatment</i>						
All	97.44	97.68	97.21	-0.47	0.712	168
Male	98.77	97.25	100.33	3.09	0.682	89
Female	95.94	98.19	93.85	-4.34	0.280	79
<i>TSST Part 1 - Public speaking task</i>						
All	109.62	120.98	98.53	-22.44	0.000	168
Male	108.70	118.99	98.18	-20.82	0.000	89
Female	110.66	123.33	98.92	-24.41	0.000	76
<i>Task 2 - Piece-rate under treatment</i>						
All	97.69	103.65	91.86	-11.80	0.000	168
Male	98.02	103.09	92.84	-10.25	0.002	89
Female	97.31	104.33	90.81	-13.52	0.000	76
<i>Task 3 - Tournament under treatment</i>						
All	98.32	102.68	94.07	-8.61	0.001	168
Male	99.58	102.97	96.12	-6.85	0.046	89
Female	96.90	102.34	91.86	-10.47	0.004	76
<i>TSST Part 2 - Alphabet task</i>						
All	101.99	108.97	95.18	-13.79	0.000	168
Male	100.37	105.99	94.62	-11.37	0.001	89
Female	103.82	112.49	95.79	-16.71	0.000	76
<i>Task 4 decision</i>						
All	88.86	91.25	86.52	-4.74	0.009	168
Male	88.40	90.66	86.08	-4.58	0.108	89
Female	89.37	91.95	86.98	-4.97	0.041	76

Notes: Mean heart rate during the specified tasks. “Stress treatment” indicates that the subject was exposed to the TSST-G stress procedure. Due to equipment failure, data is missing for 11 observations from the stress treatment and control group each. All differences are tested using a Wilcoxon rank-sum test.

TABLE A.2: THE EFFECT OF STRESS AND GENDER ON PERFORMANCE

Dep. Variable	Problems Solved Correctly							
	Task 1		Task 2		Task 2- Task 1		Task 2- Task 1	
	Piece-rate		Piece-rate		Effect of Stress Treatment		Effect of Stress Treatment	
Incentive scheme	All	All	All	All	All	Males	Females	
Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stress treatment	-0.06 (0.39)	-0.24 (0.75)	-0.14 (0.26)	-0.04 (0.33)	-0.13 (0.28)	0.02 (0.40)	0.02 (0.40)	-0.28 (0.36)
Female	-1.38*** (0.43)	-1.56** (0.74)	-0.28 (0.25)	-0.17 (0.31)	0.04 (0.25)	0.19 (0.37)	0.19 (0.37)	
Stress treatment*Female		0.37 (0.86)		-0.21 (0.46)		-0.29 (0.51)		
Solved Task 1 (baseline)			0.77*** (0.05)	0.77*** (0.05)				
Constant	7.41*** (0.51)	7.50*** (0.67)	1.53*** (0.40)	1.48*** (0.43)	-0.18 (0.26)	-0.25 (0.31)	-0.25 (0.31)	-0.06 (0.22)
Observations	190	190	190	190	190	190	95	95
R-squared	0.07	0.07	0.58	0.58	0.00	0.00	0.00	0.01

Notes: OLS. Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the number of addition problems that were correctly completed within the time limit in the specified task. Both Task 1 and Task 2 were completed under piece rate incentives, before and after the stress/control manipulation, respectively. "Stress treatment" is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.3: THE EFFECT OF STRESS, COMPETITION AND GENDER ON PERFORMANCE - ROBUSTNESS TO ADDITIONAL CONTROLS

Dep. Variable	Problems Solved Correctly: Task 3 - Task 2 Effect of Tournament Incentives							
	Not controlling for baseline performance	Baseline cortisol	Baseline heart-rate	Baseline mood	Personality traits (BFI)	Imperfect understan ding	Oral contracep tives	Phase of the menstrual cycle
Sample	All	All	All	All	All	All	All	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stress treatment	-0.18 (0.31)	-0.12 (0.29)	-0.14 (0.34)	-0.16 (0.30)	-0.15 (0.29)	-0.17 (0.30)	-0.17 (0.30)	-0.17 (0.30)
Female	0.31 (0.31)	0.55 (0.37)	0.52 (0.36)	0.49 (0.35)	0.26 (0.34)	0.39 (0.37)	0.48 (0.38)	0.46 (0.43)
Stress Treatment*Female	-1.04** (0.42)	-1.16** (0.43)	-1.09** (0.49)	-1.14** (0.42)	-1.05** (0.42)	-1.03** (0.44)	-1.06** (0.42)	-1.08** (0.42)
Solved Task 1 (baseline)		0.05 (0.06)	0.04 (0.06)	0.05 (0.06)	0.04 (0.06)	0.04 (0.06)	0.05 (0.06)	0.05 (0.05)
Observations	190	190	190	190	190	190	189	187
R-squared	0.06	0.08	0.07	0.07	0.06	0.07	0.07	0.1

Notes: OLS. Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the number of addition problems that were correctly completed within the time limit in the specified task. Both Task 2 and Task 3 were completed under treatment, under piece rate and tournament incentives, respectively. "Stress treatment" is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.4: PERFORMANCE UNDER THE CHOSEN COMPENSATION SCHEME, ACROSS TREATMENTS

Dep. Variable <i>Incentive scheme</i> <i>Sample</i>	<i>Problems Solved Correctly</i> <i>Task 4</i> <i>Chosen Incentives Scheme</i>			
	<i>All</i> (1)	<i>All</i> (2)	<i>Males</i> (3)	<i>Females</i> (4)
Stress treatment	-0.84*** (0.21)	-0.21 (0.29)	-0.19 (0.28)	-1.45*** (0.34)
Female	-0.41 (0.27)	0.22 (0.34)		
Stress Treatment*Female		-1.26** (0.50)		
Solved Task 1 (baseline)	0.81*** (0.04)	0.82*** (0.04)	0.91*** (0.05)	0.63*** (0.12)
Constant	1.88*** (0.36)	1.53*** (0.38)	0.85 (0.50)	2.85*** (0.75)
Observations	190	190	95	95
R-squared	0.61	0.62	0.72	0.39

Notes: OLS. Standard errors are clustered at a session level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is the number of addition problems that were correctly completed within the time limit in the counting portion of Task 4, before which subjects chose their preferred compensation scheme. “Stress treatment” is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.5: BINARY MEASURE OF WILLINGNESS TO COMPETE

<i>Dep. Variable</i>	<i>Probability of investing at least 50 points in the tournament</i>			
	<i>All (1)</i>	<i>All (2)</i>	<i>Males (3)</i>	<i>Females (4)</i>
Stress treatment	-0.22***	-0.15**	-0.15**	-0.24***
	-0.06	-0.07	-0.07	-0.09
Female	-0.30***	-0.22**		
	-0.07	-0.1		
Stress treatment*Female		-0.17		
		-0.12		
Solved Task 1 (baseline)	0.04***	0.05***	0.04**	0.05***
	-0.02	-0.02	-0.02	-0.02
Observations	190	190	95	95

Notes: Probit, marginal effects reported. Standard errors are clustered at the session level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is a dummy indicating that the subject in Task 4 invested at least 50 points into the tournament compensation scheme, where the choice occurred before completing the counting portion of Task 4. “Stress treatment” is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.6: WILLINGNESS TO COMPETE FOR FUTURE PERFORMANCE - ROBUSTNESS TO ADDITIONAL CONTROLS

Dep. Variable	Willingness to Compete							
	points invested into tournament ex ante (Task 4)				Phase of the menstrual cycle			
Additional controls:	Not controlling for baseline performance	Baseline cortisol	Baseline heart-rate	Baseline mood	Personality traits (BFI)	Imperfect understanding	Oral contraceptives	
Sample	All	All	All	All	All	All	All	All
Stress treatment	-7.72** (3.25)	-7.77** (3.20)	-9.87*** (3.41)	-7.62** (3.12)	-7.83** (3.03)	-7.49** (3.17)	-8.39** (3.2)	-7.62** (3.63)
Female	-25.27*** (3.41)	-22.43*** (4.29)	-23.05*** (4.19)	-22.08*** (4.33)	-21.84*** (4.36)	-21.95*** (4.18)	-26.15*** (4.83)	-35.28*** (7.08)
Solved Task 1 (baseline)		2.34** (0.95)	2.39** (0.99)	2.37** (0.95)	2.23** (0.99)	2.29** (0.94)	2.19** (0.92)	2.12** (0.91)
Observations	190	189	179	190	190	190	189	187
R-squared	0.18	0.21	0.23	0.21	0.22	0.21	0.22	0.25

Notes: OLS. Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is investment in the tournament compensation scheme in Task 4, where the choice was made before completing the counting portion of the task. 0 indicates all points invested in the piece-rate scheme, and 100 indicates all points invested in the tournament compensation scheme. "Stress treatment" is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.7: WILLINGNESS TO COMPETE FOR PAST PERFORMANCE BEFORE AND UNDER TREATMENT

Dep. Variable	Willingness to Compete									
	points invested into tournament <i>ex post</i>				points invested into tournament <i>ex post</i>					
	for Task 1 (piece-rate before Treatment)		for Task 2 (piece-rate under Treatment)		for Task 1 (piece-rate before Treatment)		for Task 2 (piece-rate under Treatment)			
Sample	All	Males	Females	All	Males	Females	All	Males	Females	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Stress treatment	1.30 (4.81)	1.89 (6.01)	2.24 (6.32)	1.07 (5.67)	-1.34 (4.37)	1.63 (6.59)	1.87 (6.88)	-4.07 (5.28)		
Female	-7.68* (4.02)	-7.09 (7.16)			-17.82*** (4.40)	-14.85** (6.01)				
Stress Treatment*Female		-1.19 (7.07)				-5.95 (7.94)				
Solved Task 1 (baseline)	4.82*** (1.03)	4.82*** (1.03)	6.25*** (0.98)	1.93 (1.70)	2.60** (1.01)	2.62** (1.02)	3.58** (1.34)	0.67 (1.18)		
Constant	11.64 (10.00)	11.32 (10.90)	0.64 (9.52)	21.43* (11.28)	32.60*** (8.03)	30.97*** (8.65)	23.78** (10.63)	27.71*** (7.03)		
Observations	190	190	95	95	190	190	95	95		
R-squared	0.17	0.17	0.23	0.02	0.12	0.12	0.07	0.01		

Notes: OLS. Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the investment in the tournament compensation for past performance. For Task 5 (columns 1-4) competition is based on performance in Task 1, (i.e before the *stress/control* procedure), and for Task 6 (columns 5-8) on performance in Task 2, (i.e after the *stress/control* procedure). 0 indicates all points invested in the piece-rate scheme; 100 indicates all points invested in the tournament compensation scheme. "Stress treatment" is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.8: WILLINGNESS TO COMPETE FOR FUTURE PERFORMANCE, BY CORTISOL RESPONSE (OLS AND 2SLS)

Dep. Variable	Willingness to Compete (Task 4)					
	points invested into tournament ex ante					
	All		Males		Females	
Sample Estimation Method	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Cortisol response	-2.35 (1.52)	-5.62** (2.31)	-0.90 (2.91)	-5.64 (3.66)	-3.48*** (0.88)	-5.65 (3.58)
Female	-22.39*** (4.28)	-22.92*** (4.29)				
Solved Task 1 (baseline)	2.33** (0.95)	2.32** (0.94)	2.13* (1.13)	2.17* (1.15)	2.68** (1.28)	2.61** (1.25)
Constant	43.49*** (8.33)	45.54*** (7.99)	44.12*** (9.96)	46.62*** (9.26)	19.49** (7.68)	20.85** (7.70)
Observations	189	189	95	95	94	94
R-squared	0.20	0.19	0.04	0.00	0.11	0.09

Notes: Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is investment in the tournament compensation scheme in Task 4, where the choice occurred before completing the counting portion of the task. 0 indicates all points invested in the piece-rate scheme; 100 indicates all points invested in the tournament compensation scheme. "Cortisol response" is the percentage increase cortisol between samples 1 and 2. Sample 1 was collected prior to the TSST-G stress/control procedure; Sample 2 was collected after the second part of the TSST-G. For details regarding the timeline see Figure 1. In Columns 2, 4 and 6, a dummy variable "Stress treatment", indicating exposure to the TSST-G stress procedure, is used as an instrument for "Cortisol response". First-stage results are presented in Appendix Table A.10.

TABLE A.9: TOURNAMENT PERFORMANCE, BY CORTISOL RESPONSE (OLS AND 2SLS)

Dep. Variable	Problems Solved Correctly in Task 3					
	Tournament Incentives					
	All		Males		Females	
Sample Estimation Method	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Cortisol response	-0.18** (0.07)	-0.63*** (0.18)	-0.09 (0.16)	-0.13 (0.20)	-0.28** (0.11)	-1.12*** (0.34)
Female	-0.43 (0.28)	-0.51 (0.30)				
Solved Task 1 (baseline)	0.81*** (0.04)	0.81*** (0.04)	0.91*** (0.05)	0.91*** (0.05)	0.61*** (0.13)	0.59*** (0.15)
Constant	1.56*** (0.33)	1.84*** (0.37)	0.79* (0.43)	0.81* (0.46)	2.38*** (0.79)	2.91*** (0.88)
Observations	189	189	95	95	94	94
R-squared	0.59	0.55	0.72	0.72	0.33	0.07

Notes: Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the number of addition problems that were correctly completed within the time limit in the counting portion of Task 3, which was completed after the treatment and under tournament incentives. ‘Cortisol response’ is the percentage increase cortisol between samples 1 and 2. Sample 1 was collected prior to the TSST-G stress/control procedure; Sample 2 was collected after the second part of the TSST-G. For details regarding the timeline see Figure 1. In Columns 2, 4 and 6, a dummy variable ‘Stress treatment’, indicating exposure to the TSST-G stress procedure, is used as an instrument for ‘Cortisol response’. First-stage results are presented in Appendix Table A.10

TABLE A.10: FIRST-STAGE REGRESSIONS: CORTISOL RESPONSE ACROSS TREATMENTS

Dep. Variable <i>Sample</i>	Cortisol response: (sample 2-1)/sample 1		
	<i>All</i>	<i>Males</i>	<i>Females</i>
	(1)	(2)	(3)
Stress treatment	1.36*** (0.14)	1.42*** (0.18)	1.32*** (0.26)
Female	-0.17 (0.15)		
Solved Task 1 (baseline)	-0.00 (0.03)	0.02 (0.04)	-0.04 (0.04)
Constant	-0.07 (0.21)	-0.25 (0.30)	0.03 (0.27)
Observations	189	95	94
R-squared	0.27	0.32	0.22
Cragg-Donald Wald F stat.	66.44	43.6	26.04

Notes: OLS, standard errors are clustered at a session level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable, “Cortisol response”, is the percentage increase cortisol between samples 1 and 2. Sample 1 was collected prior to the TSST-G stress/control procedure; Sample 2 was collected after the second part of the TSST-G. For details regarding the timeline see Figure 1. “Stress treatment” is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.11: SENSITIVITY ANALYSIS: WILLINGNESS TO COMPETE AND TOURNAMENT PERFORMANCE

Here we regress willingness to compete in Task 4 on i) stress treatment, ii) Task 3 performance and iii) both. Since Task 3 performance is endogenous (potentially affected by treatment), one must interpret the coefficients and standard errors with caution. However, comparing coefficients, standard errors and R-squared values across specifications is instructive. While adding Task 3 performance lowers the stress treatment coefficient and increases the R-squared, adding stress treatment affects neither the coefficient for Task 3 performance nor the R-squared.

Dependent variable	Willingness to Compete									
	points invested into tournament ex ante (Task 4)									
	All			Males			Females			
Sample	(1)	(2)	(4)	(4)	(4)	(5)	(6)	(7)	(8)	(9)
Stress treatment	-7.59** (3.14)	-3.89 (3.13)		-7.99 (5.00)	-7.01 (5.13)	-7.31 (4.58)			-0.85 (4.72)	
Solved Task 3		4.38*** (0.91)	4.62*** (0.91)		5.22** (1.87)	5.34** (1.92)			4.46*** (0.97)	4.53*** (0.84)
Female	-22.06*** (4.14)	-20.25*** (4.36)	-20.13*** (4.41)							
Solved Task 1 (baseline)	2.33** (0.95)	-1.24 (1.15)	-1.43 (1.15)	2.07* (1.15)	-2.68 (2.15)	2.85** (1.34)	-2.74 (2.17)		0.03 (1.59)	-0.02 (1.53)
Constant	45.89*** (8.05)	37.67*** (8.00)	35.33*** (7.93)	48.02*** (9.53)	43.60*** (9.35)	20.55** (8.06)	39.68*** (9.19)		7.87 (9.36)	7.31 (8.65)
Observations	190	190	190	95	95	95	95	95	95	95
R-squared	0.21	0.27	0.27	0.05	0.11	0.09	0.10	0.09	0.21	0.21
Adjusted R-squared	0.20	0.26	0.26	0.03	0.09	0.07	0.08	0.07	0.18	0.19

Notes: OLS. Standard errors are clustered at a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is investment in the tournament compensation scheme in Task 4, where the choice occurred before completing the counting portion of the task. 0 indicates all points invested in the piece-rate scheme, and 100 indicates all points invested in the tournament compensation scheme. "Solved Task 3" captures the number of correctly answered problems in Task 3, which occurred under tournament incentives under treatment. "Stress treatment" is a dummy variable indicating that the subject was exposed to the TSST-G stress procedure.

TABLE A.12: SENSITIVITY ANALYSIS: TOURNAMENT PERFORMANCE AND RELATED CONFIDENCE

Here we regress confidence in Task 3 on i) stress treatment, ii) Task 3 performance and iii) both. Since Task 3 performance is endogenous (potentially affected by treatment), the results should be interpreted with caution. However, comparing specifications shows that while Task 3 performance explains confidence, the stress treatment does not additionally explain confidence.

Dependent variable	Confidence in Task 3								
	All			Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stress treatment	0.49** (0.20)	0.03 (0.17)		-0.08 (0.31)	-0.18 (0.28)		1.05*** (0.29)	0.26 (0.29)	
Solved Task 3		-0.55*** (0.05)	-0.55*** (0.04)		-0.54*** (0.08)	-0.53*** (0.08)		-0.55*** (0.05)	-0.57*** (0.04)
Female	0.72*** (0.25)	0.49** (0.23)	0.49** (0.23)						
Solved Task 1 (baseline)	0.22*** (0.04)	0.23*** (0.05)	0.23*** (0.05)	-0.25*** (0.04)	0.24** (0.09)	0.23** (0.09)	-0.16** (0.07)	0.18** (0.08)	0.20** (0.08)
Constant	5.31*** (0.34)	6.33*** (0.34)	6.35*** (0.32)	5.84*** (0.40)	6.30*** (0.38)	6.19*** (0.36)	5.42*** (0.46)	6.98*** (0.54)	7.15*** (0.48)
Observations	190	190	190	95	95	95	95	95	95
R-squared	0.18	0.46	0.46	0.17	0.41	0.41	0.12	0.42	0.42
Adjusted R-squared	0.17	0.45	0.45	0.15	0.39	0.40	0.10	0.41	0.41

Notes: OLS. Standard errors are clustered on a session level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is confidence in Task 3, measured as subject's subjective ranking among 8 participants in the session. 1 thus indicates highest confidence and 8 the lowest confidence. "Solved Task 3" captures the number of correctly answered problems in Task 3, which occurred under tournament incentives under treatment. "Stress treatment" indicates that the subject was exposed to the TSST-G stress procedure.

TABLE A.13: HEART-RATE RESPONSE TO STRESS/CONTROL TREATMENT VS. RESPONSE TO TOURNAMENT INCENTIVES

<i>Sample</i>	<i>Heart-rate response to specified tasks</i>					<i>N</i>
	<i>All</i>	<i>Treatment group</i>		<i>Diff.</i>	<i>Ranksum</i> (p-value)	
		<i>Stress</i>	<i>Control</i>			
(1)	(2)	(3)	(4)	(5)		
<i>Response to stress/control treatment (TSST Part 1 - Task 1)</i>						
All	12.18	23.30	1.33	-21.97	0.000	168
Male	9.93	21.75	-2.16	-23.90	0.000	89
Female	14.72	25.13	5.07	-20.07	0.000	76
<i>Response to tournament incentives (Task 3 - Task 2)</i>						
All	0.64	-0.97	2.21	3.18	0.000	168
Male	1.57	-0.11	3.28	3.40	0.001	89
Female	-0.41	-1.99	1.06	3.05	0.009	76

Notes: Differences between mean heart rates during the specified tasks. “Stress treatment” indicates that the subject was exposed to the TSST-G stress procedure. Due to equipment failure in we have missing data for 11 observations from the stress treatment and control group, each.

Abstrakt

Mnoho zásadních kariérních událostí, jako jsou například zkoušky a pracovní pohovory, obnáší soutěž s ostatními a stres. Rozdílné reakce mezi muži a ženami na tyto dva faktory by mohly přispět k vysvětlení genderových nerovností pozorovaných na pracovním trhu. Za účelem prozkoumání této otázky jsme provedli laboratorní experiment, ve kterém manipulujeme úroveň psychosociálního stresu pomocí procedury Trier Social Stress Test. Úspěšnost manipulace demonstrujeme pomocí hladiny kortizolu ve slinách. Probandi v experimentu řeší sérii početních úloh, přičemž jsou za ně vypláceni buď formou turnaje, nebo paušální sazby, a dále měříme jejich ochotu soutěžit. Z našich výsledků plyne, že ženy pod zvýšeným stresem mají nižší výkonnost v početních úlohách než ženy v kontrolní skupině, ale pouze pokud jsou vyplácené formou turnaje. U platby formou paušální sazby rozdíl ve výkonnosti žen nenacházíme. U mužů v našem vzorku stres výkonnost nijak neovlivňuje. Stres dále celkově snižuje ochotu soutěžit, což je u žen možné vysvětlit jejich nižší výkonností pod stresem při výplatě pomocí turnaje. Naše výsledky tak napomáhají vysvětlit předchozí poznatky o genderových rozdílech v reakci na soutěživé prostředí.

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

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

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

(c) Jana Cahlíková, Lubomír Cingl a Ian Levely, 2017

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

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

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

Editor: Jan Zápál

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

ISBN 978-80-7343-396-3 (Univerzita Karlova, Centrum pro ekonomický výzkum
a doktorské studium)
ISBN 978-80-7344-425-9 (Národohospodářský ústav AV ČR, v. v. i.)