Choice and Performance Under Stress: Are Men and Women Different?

Manuela Angelucci[α] and Karina Córdova[β*]

July 1, 2013

Abstract

We test the hypothesis that, while stress worsens entrepreneurial choices and outcomes for all, it does so more for women than men. Although there are no gender differences in the control group, the effects of stress on choice and performance are more negative for women. In particular, experimentally-induced stress causes (i) more long-lasting productivity losses for women and (ii) additional losses for making choices that do not maximize income given one's productivity. The negative treatment effect on women's productivity, choice quality, and earnings is driven by women who either experienced a parental divorce or the death of a close person, or have less educated fathers. The mechanisms that affect choices also differ by gender. Given that women are both subject to more stressors and respond more negatively to stress than men, the differential incidence of and response to stress by gender may be one determinant of why women are under-represented in high-pressure activities, jobs, and professions.

PRELIMINARY - PLEASE DO NOT CITE

[*α]University of Michigan (mangeluc@umich.edu) [β]University of Arizona (cordovak@email.arizona.edu)

This paper is based on previous work developed with Anna Breman. We are grateful to Lisa Ordóñez for fruitful discussions, support, and her great feedback, to Richard Kiser from the Economic Science Laboratory at the University of Arizona for his invaluable help, and to seminar participants at the University of Michigan, UCL, NYU, and the Central European University for their comments.

1 Introduction

Entrepreneurship rates are lower among women than men in most of the world (Kelley, Singer, and Herrington, 2012; Koellinger, Minniti and Schade, 2013). In the U.S., women account for only 2.5% of the top 5 earning executives (Bertrand and Hallock, 2001). More generally, women tend to be under-represented in high-pressure activities, jobs, and professions (e.g., women migrate less, and are less likely to be in tenure-track positions in academia). Several explanations have been proposed to account for these differences, ranging from discrimination (Wolfinger, Mason, and Goulden, 2008), to preferences (Barbulescu and Bidwell, 2012), to women being less risk taking (e.g., Croson and Gneezy, 2009), competitive (e.g., Niederle and Vesterlund, 2007, 2011), and overconfident (e.g., Barber and Odean, 2001; Cech, Rubineau, Silbey, and Seron, 2011) than men.

This paper proposes an additional explanation, related to stress. Stress is the organism's response to stimuli that threaten one's physical integrity, social connectedness, sense of self, and resources (Kemeny, 2003), with heterogeneous causes and effects. Our argument is that, if women are both more exposed to stress and more negatively affected by stress than men, stress might cause women to participate in entrepreneurial activities and high-pressure professions at lower rates than men. This difference arises through two mechanisms. First, conditional on baseline characteristics, women are less likely to choose higher-return, higher-pressure activities than men.[1] Second, conditional on having made this type of choice, women may perform worse than men.

There is evidence that women suffer from higher exogenous stressors than men: e.g. as children, they experience higher parental divorce and poverty rates (Dahl and Moretti, 2008); as adults, they have more parenting and caregiving responsibilities (Bianchi, 2010, 2011; Anxo, Flood, Mencarini, Pailh, Solaz, and Tanturri, 2007). These differences are consistent with the higher self-reorted levels of stress: for example, in the U.S., 58 (16)% of men and 79 (23)% of women reported experiencing moderate to high (extreme) levels of stress in 2011 (APA, 2012); perceived job insecurity is also higher for women (Mauno and Kinnunen, 2002). So, even if the effects of stress on choice and performance did not differ by gender, the overall effect of stress would be higher for women because they are exposed to more stressors. In addition, there are documented biological (Taylor, Klein, Lewis, Gruenewald, Gurung, and Updegraff, 2000; Lee and Harley, 2012) and neurological (Wang, Korczykowski, Rao, Fan, Pluta, Gur, McEwen, and Detre, 2007) gender differences in the responses to stress, as well as evidence that the hormonal changes caused by the menstrual cycle affect women's physiological response to stress (Kirschbaum, Kudielka, Gaab, Schommer, and Hellhammer, 1999) and their bidding behavior (Chen, Katuščák, and Ozdenoren, 2013; Pearson and Schipper, 2013).

To understand these mechanisms, we design and perform a series of laboratory experiments

[1] Ors, Palomino, and Peyrache (2012) show that the ex ante mechanism is present also outside the lab: among applicants for a prestigious MBA program, women are more likely to underperform in the high-stake MBA admission exams, compared to men, despite having performed better than men in low-stake exams.

where we elicit stress and track both productivity and entrepreneurial choices, which we define as choosing a payment scheme with higher returns but a small, positive probability of zero earnings rather than a payment with lower, but certain returns. In this setting, described below, we test the hypotheses that (1) stress negatively affects productivity and entrepreneurial choice, (2) the effect of stress on choice is not income maximizing, (3) this non-maximizing effect works through changes in beliefs, mistakes, and risk attitudes, and (4) all these effects vary by gender. Our data match the broader evidence that women are more likely to be exposed to stressors and report higher levels of stress than men and that women's productivity and entrepreneurial choices are more negatively affected by experimentally induced stress than men's, resulting in lower monetary gains.

In our experiment, people are compensated for their choices and performance in one of four tasks. The first two tasks are two rounds of 20 multiple-choice, SAT-like questions, with 5 minutes to answer them. We call them quiz$_1$ and quiz$_2$. Participants are paid piece-rate in both cases and may earn between 0 and \$20 in quiz$_1$ and between \$0 and \$30 in quiz$_2$. The third task consists of picking their preferred choice out of 6 paired lotteries (à la Holt and Laury, 2002), for which people may earn between \$1 and \$38.5. The fourth task consists of recalling the answers to two multiple choice questions that people answered after watching a video clip at the beginning of the session.

We have two different treatments, emotional load and uncertainty. People in the control group watch a neutral video clip. Subjects under emotional load watch an anxiety-inducing video clip and know that, at the end of the session, one of them will be selected to summarize the clip in front of the other subjects. Participants under uncertainty watch the neutral video clip and are *not* told their quiz$_1$ productivity right after completing this task.

The sessions start by collecting information to estimate participants' baseline stress level, and a saliva sample to measure cortisol (one of the hormones that increases when the body reacts to a stressor) for a subset of 92 subjects in the control and emotional load arms. Then each person watches a video clip and answers questions about it. The first productivity task, quiz$_1$, follows and, after completing it, people in the control group and in the emotional load treatment, but not those in the uncertainty treatment, find out what their number of correct answers was. Afterwards, participants report their expected number of correct answers for quiz$_2$ and choose how to be paid for quiz$_2$. The options are whether to be paid \$1 per correct answers, as in quiz$_1$, or \$1.5, provided they have a minimum number of correct answers, x. If they have fewer than x correct answers, they are paid \$0. We offer this 'upgrade' 4 times, each time lowering the threshold from 12, to 10, 8, and 6 correct answers and giving people the option of choosing the new 'upgrade' - the one with the lower threshold - or sticking with whichever payment they had just selected, regardless of whether it is \$1 with no threshold or \$1.5 with the higher threshold. For every upgrade offered, people report their subjective probability of having x correct answers. Participants, then, solve quiz$_2$, and right after indicate their preferred choices out of 6 paired lotteries. We then collect a second saliva sample in selected sessions. For the last paid task, people have to recall the two multiple choice answers provided after watching the video clip. At this stage, one person from the emotional load group summarizes the video clip she saw in front of the other participants. The session is over after we collect some socioeconomic information (age, gender, college year, GPA, parental education, and negative life events; in addition, we ask to report consumption of alcohol, caffeine, tobacco, and use of various medicines in the sessions in which we collect saliva samples).

We find that, when under stress, both men and women become less productive and make worse choices. However, consistent with our conjectures, the duration or the intensity of the effects is larger for women. Emotional load reduces quiz_1 productivity by 10% for both men and women. However, while the treatments continue to reduce women's quiz_2 productivity by 10%, they no longer affect men's quiz_2 productivity, which returns to the level of the control group. The treatments cause a reduction in the upgrade likelihood - our measure of entrepreneurial choice - of similar magnitudes for men and women, despite the fact that stress did not cause a drop in men's quiz_2 productivity. This finding suggests that, at least for men, stress causes excessive drops in upgrade choices. Indeed, when we look at the treatment effect on the monetary returns on quiz_2, we find that the pooled treatments reduce women's and men's returns by \$2.8 and \$0.5 respectively, with larger losses from uncertainty than from emotional load and the difference between the effects for men and women being statistically significant. About \$1 and \$0.6 (with the difference not statistically significant), that is, 37% and 100% of these losses for women and men respectively, are caused by making the 'wrong' upgrade choice given one's productivity, and the rest by a productivity drop. Most of the losses from making the wrong choice, \$0.63 for women and \$0.45 for men, come from choosing not to upgrade when in fact they would have earned more by upgrading. The remaining \$0.39 and \$0.13 are caused by upgrading too much given one's productivity, and consequently earning \$0. In sum, stress causes women to lose more money than men.

When we look at heterogeneity by negative life events (parental divorce or death of a close person in the previous 5 years) and socioeconomic background (proxied by having a father with at least a college degree), we find that the negative treatment effect on women's productivity, choice quality, and earnings is driven by women who either experienced the above negative life events or have less educated fathers. The effects are not statistically significant for the remaining women, and are also mainly statistically insignificant and do not vary systematically by family background for men.[2]

When we look at heterogeneity by self-reported stress, we find that that the (negative) marginal effects of the experimentally induced stress decrease with baseline stress for men but not so much for women. That is, the stress induced in the lab worsens productivity, choice quality, and earnings for men with below median baseline self-reported stress, but not for men with above median baseline self-reported stress.[3] Conversely, the negative effect of lab-induced stress worsens productivity,

[2]Self-reported stress and these measures of socioeconomic background are uncorrelated with each other for both men and women.

[3]This is the reverse of The Yerkes-Dodson Law (see Lupien, Maheu, Tu, Fiocco, and Schramek, 2007; Diamond, Campbell, Park, Halonen, and Zoladz, 2007) and is consistent with Meglino (1977) and Arsenault and Dolan (1983).

choice quality, and earnings for all women, regardless of their baseline stress. That is, a larger fraction of women than men is negatively affected by our treatments.

In sum, when we use two different proxies for current or past stress, we find that the marginal effect of the treatments interact differently with predetermined individual characteristics, resulting into women being more hurt by the treatments than men.

Lastly, we analyze three different mechanisms through which stress might affect entrepreneurial choice and performance (beliefs, mistakes, and risk attitudes) and identify gender differences. We find that the reasons for excessive upgrades are a surge in optimism for some women, and a surge in optimism, mistakes, and risk-loving preferences for some men. The reasons for upgrading too little, on the other hand, are a surge in pessimism for some women and an increase in mistakes and risk aversion for men. In sum, policies targeted at increasing the accuracy of beliefs will benefit both men and women, while policies targeted at dealing with risk preferences and correcting mistakes will primarily help men. We discuss the interaction of these effects in more detail and discuss potential policy implications that could incorporate the insights of this research.

2 Treatment, Experimental Design, Steps, and Data

2.1 Treatments and Experimental Design

Observational or self-reported data are usually inadequate for estimating the effects of stress because of reverse causality. Decision-making problems, low productivity, or low income may in fact cause stress. To overcome this problem, we design a series of laboratory experiments in which we randomly assign treatment conditions and measure the short-run effects of different aspects of stress on decision-making and productivity. We induce stress through the use of a movie clip and the withholding of information. In addition to measuring stress through the typical self-report measures, we also measured cortisol levels objectively by administering saliva tests to a subsample of participants.

We have two treatments, emotional load, EL, and uncertainty, U, and a control group. Everybody watches one of two video clips. The control group watches a five minute fragment of a BBC documentary about Monument valley, the same setting as the emotional load video clip. The emotional load group watches a four minute clip from the movie "Vertical Limit" (Columbia Pictures, 2000), which depicts a dramatic rock climbing accident. The use of films for the elicitation of emotions is a common and effective practice in psychology experiments (Gross and Levenson, 1995) and this particular video clip has been used in previous research (e.g. Frederickson and Branigan, 2005; Brooks and Schweitzer, 2011). Moreover, we announce that at the end of each session, one person from the emotional load group has to summarize the video clip in front of the other subjects. Public speaking is also a commonly employed stressor in the lab (e.g. Dickerson and Kemeny, 2004). Besides increasing stress, public speaking increases the salience of the video clip.

The uncertainty group watches the documentary and is *not* told the quiz_1 productivity. Uncertainty is one of the most commonly reported stressors in the U.S. (APA 2012). Moreover, there is experimental evidence that negative feedback is preferred to uncertainty about performance, especially among those dealing with high anxiety (Hirsh and Inzlicht, 2008), and descriptive evidence that uncertainty increases anxiety (Flory and Lang 2010). Individuals value a risky prospect less than its worst possible realization (Gneezy, List and Wu, 2006, who actually call this finding the "uncertainty effect"), and uncertainty can affect performance negatively when applying for a job (e.g. Godlonton, 2012). Given the evidence on the different physiological and neurological mechanisms in which men and women react under stress, we expect heterogeneous effects of these two stressors (emotional load and uncertainty) by gender.

Some subjects are exposed to neither treatment, EL=0 and U=0, some to one treatment only, EL=0 and U=1 or EL=1 and U=0, while others to both, EL=1 and U=1. The control group is always the set of subjects exposed to neither treatment, EL=0 and U=0. The treatment groups vary, though. When we pool treatments, the control group is all subjects exposed to at least one treatment, EL, U, or both. When we consider individual treatments separately, though, we select subjects exposed to one treatment only. That is, the treatment group for the EL condition is EL=1 and U=0. Similarly, the treatment group for the EL condition is EL=0 and U=1. Table 1 shows that the size of the control group is 119 (83+36), while the treatment group sizes are 271 (52+92+61+66) for the pooled treatments, 92 for the EL treatment, and 118 (52+66) for the U treatment. We obtain these sample sizes by adding also data from our robustness check experiments. Before explaining that, it is useful to describe the experimental design.

Our experiment is a modified two-by-two design, as shown in Table 2, which lists the sessions and describes which conditions we induce in each of them. In the first 12 sessions, we have both treatment and control groups for the EL condition (video clips and public speaking, i.e. EL=0 and EL=1). However, no subject in the first 6 sessions receives the uncertainty treatment (U=0), while all subjects in the second 6 sessions do (U=1). That is, while the EL treatment varies within session, the U treatment varies across sections only. In addition, we ran 8 more sessions in which we manipulate the cognitive load (CL) by having subjects memorize either 2- (CL=0) or 8-digit numbers (CL=1). The CL treatment varies within session. Two of these sessions have no uncertainty treatment (U=0), while all subjects in the remaining sections receive the uncertainty treatment (U=1). While we use these sessions primarily for robustness checks, as we explain in Section 10, we add the CL=0 subgroup to our main data set, in the following way: subjects with CL=0 and U=0 are part of the control group, T=0; subjects with CL=0 and U=1 are part of our pooled treatment group, T=1, and of our uncertainty treatment group, U=1.

2.2 Experimental Steps

The following list shows the steps of the experimental design.[4]

1. Random assignment of participants to treatment or control group.
2. General explanation of tasks.
3. Collect saliva sample in selected sessions (to measure baseline cortisol).
4. Questionnaire on self-reported stress.
5. Show video clip.
6. Multiple-choice questions about the clip.
7. Quiz$_1$ (20 questions, 5 minutes, \$1 per correct answer)
8. Disclose or ask participants to guess own number of correct answers from quiz$_1$.
9. Ask about expected number of correct answers for quiz$_2$.
10. Upgrade options: participants choose payment scheme for quiz$_2$, 4 sequential offers.
11. During each upgrade, ask probability of having at least a specific number of correct answers.
12. Quiz$_2$ (20 questions, 5 minutes, \$1 or \$1.5 per correct answer)
13. Choice of 6 paired Holt-Laury lotteries.
14. Collect second saliva sample in selected sessions (to measure endline cortisol).
15. Recall answers to questions about the clip.
16. Public speaking task for one randomly picked participant in the EL treatment.
17. SES and life events data collection.

All sessions are performed at the Economic Science Laboratory of the University of Arizona. All subjects are undergraduate students, recruited through the electronic laboratory's system, who volunteer to participate. Upon arrival to the laboratory, we give participants a disclosure form with general information about the experiment and collect all personal belongings to prevent the subjects from taking notes or pictures and from using calculators. We distribute tags to randomly assign each subject to the treatment or control group's workstation (the tags were also the bag checks).

[4]There are two features included in some experimental sessions, but not listed here since they had no impact in our results: we changed the order for the Holt-Laury lotteries choices in 5 sessions (done right before the first quiz instead of doing it after the second quiz), and added a task involving a snack choice.

Before the beginning of each session, a researcher gives a general explanation of the experimental tasks, instructions are also displayed on the computer screens. Participants are guaranteed a $5 show-up fee, plus the chance to earn up to an additional $38.5. We pay people based on their choices and performance in one of four tasks, chosen at random at the end of the experiment.

For the sessions in which we collect saliva samples to measure cortisol, we utilize an special oral swab and give clear instructions to indicate the right procedure to follow (see Appendix for further details). Measuring cortisol with this method is safe and non-invasive; moreover, salivary cortisol has a strong correlation with cortisol measured by other means.[5] Participants in this subsample were asked to refrain from consuming several drinks or foods before the experiment.[6] Recruited individuals unable to fulfill these requirements were not allowed to participate in these sessions. We analyze cortisol as the way to objectively measure the body's reaction to stress and assess the effectiveness of our treatment. In addition, to measure self-reported stress, we have participants answer the Perceived Stress Scale 10-Item Questionnaire (Cohen, Kamarck, and Mermelstein, 1983; Cohen and Williamson, 1988). These are questions regarding feelings of confidence, control and anger in the previous four weeks. The answers are coded to sum a total score that can range from 0 to 40. This is a relative measure of stress as then each individual score is compared to the median of the sample. The questionnaire is available in the Appendix.

We measure productivity through two rounds of 20 multiple-choice questions, selected from the Scholastic Aptitude Test (SAT) mathematics section out of questions that can be solved without complicated calculations.[7]. We choose to carry out a quiz because it resembles the everyday activities that students do, and particularly the SAT questions are related to knowledge that every college student should possess. Productivity is measured by the number of correct answers answered in five minutes. Each subject clicked in their own pace to get the next question. Each correct answer have a value of $1, for a total compensation of up to $20 if the participant were to be paid for this task.

Before responding to the second quiz, we offer the participants an opportunity to upgrade their payment scheme conditional on their expected performance. They are offered four sequential upgrades: instead of getting $1 per correct answer (the risk-free payment), they can get $1.5 if they have a minimum of x correct answers (x=12,10,8,6), in order to make up to $30. It is worth noting

[5]See, for example, Granger, Kivlighan, Fortunato, Harmon, Hibel, Schwartz, and Whembolua (2007) and Levine, Zagoory-Sharon, Feldman, Lewis, and Weller (2007). Additionaly, note that their concerns about compliance of the adequate conditions for saliva collection were addressed in our protocol. In fact, none of our saliva samples were neither contaminated, nor insufficient for the analysis of cortisol.

[6]Specifically, they were asked to avoid alcohol for 12 hours before the experiment, a major meal 60 minutes before the experiment, dairy products 20 minutes before, and sugary, acidic or highly in caffeine foods or drinks immediately before the experiment. This is standard procedure as indicated by Salimetrics and documented by Granger, Blair, Willoughby, Kivlighan, Hibel, Weigand, and Fortunato (2007), Granger, Hibel, Fortunato, and Kapelewski (2009), among many others.

[7]The SAT is a standardized test for college admissions in the United States and consists of three sections: critical reading, mathematics and writing.

that if a participant decided to upgrade at x=10, they can keep upgrading at the following options (x=8,6). This is in fact how we measure entrepreneurial choice: as the likelihood of choosing the risky payment, whose return depends on one's performance. Additionally, participants are asked to estimate the probability of having at least x correct answers (where x=12,10,8,6). All these variables allow us to evaluate the level of confusion that could have lead to non-maximizing actions.

The second treatment, or the second source of stress that we elicit in this experiment, is uncertainty, and we do so with variations in the level of information available to participants. For subjects in 8 out of 20 sessions we provide feedback on their performance in their first quiz. The reasoning is that having knowledge of how good or bad they did, subjects with information can adjust their perceptions on their own performance and/or set and reach more realistic goals. In the remaining 12 sessions we do not provide this information and, in fact, ask participants to estimate their future productivity, giving us an idea of their actual perceptions on performance. This allows us to compare their actual productivity in the second quiz to such perceptions, evaluate if they are optimistic or pessimistic, and if their choices are income-maximizing.

In order to measure if anxiety and uncertainty, as induced in this experiment, can affect attitudes towards risk, we also include a task where subjects had to choose one option in each of six pairs of lotteries following Holt and Laury (2002).[8] If the participant is to be paid for this task, she would roll an 6-sided die to define for which of the six alternatives she would be compensated for. This task pays a minimum of \$1.5 and a maximum of \$38.5.

The last activity is the recollection of the responses they provided about the clip after they watched it (for those in the EL group) or the number showed to them at the beginning (for those in the CL group). This is the memorization or recollection task, and participants can get \$15 if remember their answers or numbers correctly.

Finally, participants respond a questionnaire of basic socioeconomic data (age, sex, college year, GPA and parental education). We also ask a few questions on life events to be able to control for negative events that could be associated as sources of long-term effect, like experiencing parents' divorce, the death of a close relative, financial crisis or an accident.[9] Additionally, for the subsample from which we measure cortisol, we ask questions related to the consumption of alcohol, cigarettes, caffeine-loaded drinks, and several medications, in order to control for any factors that could alter cortisol levels found in the saliva samples.

Despite the multiple tasks and the level of detail of the experiment, it lasts on average around 28-38 minutes to finish from the moment the participant walks in the lab to when final payments are handed. The show-up fee is the same for all the participants: \$5. The rest of the payment

[8]This modified version of paired lotteries was thought as a brief version of the original set of ten paired lotteries in order to reduce the duration of the experiment. A pilot analysis showed no significant difference in risk aversion when we compared responses of individuals evaluating the original and the modified version of the Holt and Laury (2002) lotteries.

[9]This mimics Oswald *et al* (2009) who consider these events as a proxy of Nature-driven allocation of happiness.

is randomly assigned. Each participant rolls a die to determine for which task they would be paid: performance on quiz 1 or 2, lottery choices or memorization task. The expected profits for a participant range from a minimum of \$5 to a maximum of \$43.5.

2.3 Data

Our pooled data are 390 observations, of which 119 in the control group. The breakdown by gender is 51 females and 68 males in the control group and 130 females and 184 males in the pooled treatments. When we consider the two treatments separately, however, we have 38 females and 54 males in the emotional load condition and 52 females and 66 males in the uncertainty condition.

Table 3 shows that the covariates are balanced across treatment arms (columns 1-3). The few exceptions are in age when we pooled all treatments: individuals in the treatment group are 0.4 years older than those in the control; and in GPA and father's education when we look at participants in the uncertainty treatment: those under U treatment have higher GPAs (0.4 points) and slightly less educated fathers. When we regress the treatment dummy on the predetermined covariates, we cannot reject the F test of joint significance of the covariates.

The characteristics of males and females do not differ systematically, except that females are 0.4-0.5 years younger, have a higher GPA than men (around 3-6%, depending on the treatment type), and present a higher self-reported stress (14-19% higher than men). In our estimations we control for the relevant covariates.

Our data show that (1) people with higher baseline cortisol are less productive and that (2) people with higher self-reported stress are less productive, less optimistic, and less likely to upgrade (results available upon request).

3 Methodology

We are interested in studying the effect of stress on entrepreneurial choice and outcomes and in identifying some mechanisms through which this occurs. For this purpose, for each outcome of interest, we show the distribution of outcomes for the control and treatment groups and estimate the Average Treatment on the Treated (ATT) effects.

We produce variables' kernel densities using the Epanechnikov kernel density function and the optimal bandwidth (the bandwidth that would minimize the mean integrated squared error if the underlying distribution were Gaussian). We estimate the ATT effects on the various outcomes Y from the following regressions:

$$Y_i = \alpha_1 + \beta_1 T_i + \beta_{1W} T_i W_i + \gamma_1 X_i + \epsilon_{1i} \quad (1)$$

$$Y_i = \alpha_2 + \beta_{2EL} EL_i + \beta_{2U} U_i + \beta_{2ELW} EL_i W_i + \beta_{2UW} U_i W_i + \delta_2 U_i EL_i + \delta_2 U_i EL_i W_i + \gamma_2 X_i + \epsilon_{2i}. \quad (2)$$

The variables T, EL, and U are dummy variables for the pooled treatments (T) and the separate treatments (EL and U). The dummy W is one for women and zero otherwise; the term X includes the following predetermined variables: a gender dummy, age, college year dummies, a dummy for whether one's father has at least a college degree, baseline self-reported stress, a dummy for whether one's parents divorced or a close relative died in the previous 5 years, as well as dummies for subjects who experienced technical difficulties during the experiment.[10] The parameters β_1 and $\beta_1 + \beta_{1W}$ identify the average effects of the pooled treatments for men and women, while the parameters β_{2EL}, $\beta_{2EL} + \beta_{2ELW}$, β_{2U}, and $\beta_{2U} + \beta_{2UW}$ identify the average effects of emotional load (EL) and uncertainty (U) for men and women. These parameters are identified under random assignment and absent spillover effects from the treatment to the control group.

To estimate separate ATT effects by subgroups, we further interact the three treatment dummies (and their interactions with each other and with the gender dummy) with a subgroup dummy. The coefficient of the interaction tests the hypothesis that the effects differ by subgroup. We estimate all the regressions by OLS with robust standard errors. When we pool outcomes, and therefore have multiple observations per subject (e.g. when we consider all upgrades jointly), we cluster the standard errors by subject.

4 High Emotional Load and Cortisol

To test that the EL condition induces stress, we collected saliva samples for 92 subjects, 39 females and 53 males, and measured the change in cortisol per subject. Figure 1 shows the densities of the change in cortisol by gender and treatment. The EL condition seems to increase cortisol for males, but not for females. Table 4 confirms this finding. While the ATT on changes in cortisol for males is 0.074 (s.e. = 0.026), compared with a mean change of -0.035 in the control group, for females it is only -0.016 (s.e. = 0.023), compared with a mean change of -0.030 in the control group. The difference between the two effects is statistically significant and consistent with findings of higher cortisol responses in young men than young women from at least 11 different studies, reviewed in Kudielka and Kirschbaum (2005).

The lack of a significant change in cortisol for females does not rule out that the condition caused stress for them. Indeed, there are well-known gender differences in the neurobiological response to stress (e.g. Wang *et al*, 2007). In particular, while there is no gender difference in total plasma cortisol stress responses, the effect of stress on free salivary cortisol is higher for men than for women who use oral contraceptives or in the follicolar menstrual phase, but not different from women in the luteal phase (Kirschbaum *et al*, 1999). That is, especially for women, even if there is an increase in cortisol production in response to a stressor, it might not be detected in saliva samples depending on the presence of other hormones.

[10] There were temporary technical problems with the computer screens in three cases. In one session, there were problems with the timer, and the 30 subjects involved had slightly more than 5 minutes to answer quiz 2. Omitting these observations from the analysis does not change the results.

5 Stress and productivity

5.1 Effect of stress on correct answers

Stress may make people less entrepreneurial because they become less productive and correctly anticipate it. To measure the effect of stress on productivity, we look at the treatment effects on the number of correct answers from quiz$_1$ and 2. Figure 2 shows the distribution of correct answers from quiz$_1$ (top panel) and quiz$_2$ (bottom panel) by gender and treatments. The top panel shows that high emotional load reduces quiz$_1$ productivity for both men and women.

Comparing the densities of the control group's quiz$_1$ and 2 shows a performance increase for both genders. This is consistent with both learning and with an increase in effort induced by the higher stakes. Indeed, the average number of correct answers in the control group increases from 8.61 in quiz$_1$ to 12.42 in quiz$_2$. The improvement from quiz$_1$ to quiz$_2$ does not differ systematically for people who do or do not upgrade (from unreported regressions) both unconditionally and within treatment arm, suggesting that the improvement in performance may be primarily driven by learning.

Comparing quiz$_2$'s performance for the control and treatment groups shows that the negative effect of stress on performance is more marked for women than men, and more marked under uncertainty than high emotional load.

The OLS estimates of the ATTs from Table 5 confirm these findings, showing that high emotional load causes productivity to drop by 10%, a statistically significant effect, and that the effect size is the same for both genders. In sum, men's and women's immediate reaction to high emotional load is to become equally less productive. However, the effect of stress on quiz$_2$ productivity varies by gender. The pooled treatments significantly reduce quiz$_2$ performance for women, who experience a 10% drop, that is, a drop of similar magnitude as the effect of stress on quiz$_1$, but not for men, whose estimated ATT is roughly 0. The effect does not vary by treatment type for both gender, although, for women, the distribution of correct answers in the control group first-order stochastically dominates the one under high emotional load, but not the one under uncertainty.

In sum, while high emotional load decreases quiz$_1$ performance by 10% for both genders, both high emotional load and uncertainty continue to decrease quiz 2 performance by 10% for women, but have no significant effect for men. Either the effect of the stressors is inherently more short-lived for men than for women, or men have better coping skills to "bounce back," while women do not, or both. The available evidence on the differential effects of stress by gender shows a similar differential effect by gender. For example, women have stronger and more persistent physiological changes to marital conflict than men (Kiecolt-Glaser, Newton, Cacioppo, MacCallum, Glaser, and Malarkey, 1996; Kiecolt-Glaser, Glaser, Cacioppo, Malarkey, 1998). There are conjectures that a strong initial neural and hormonal response, followed by a quick recovery, may be adaptive (Linden, Earle, Gerin, and Christenfeld, 1997; Sapolsky, Romero, and Munck, 2000) and the reverse dysfunctional (Sapolsky *et al.*, 2000).

5.2 Does stress increase anyone's productivity?

Psychologists conjecture that moderate levels of stress might improve performance for certain tasks (Yerkes and Dodson, 1908). However, the empirical evidence is mixed (see, e.g. the review in Kavanagh, 2005). We test whether stress improves some subject's performance in the several ways. First, we test the hypothesis of equality of variances using the Levene's test. The variances of an outcome are identical for the treatment and control groups under the null hypothesis of homogeneous treatment effects. Therefore, if stress reduces productivity for some, and it increases productivity for others, we reject this null hypothesis.[11] Second, we test for First Order Stochastic Dominance (FOSD) using one-sided Kolmogorov-Smirnov tests. Consider two marginal distribution, $F(T)$ and $F(C)$. If we fail to reject the null hypothesis that $F(T) > F(C)$ but we can reject the hypothesis that $F(T) < F(C)$, we conclude that $F(C)$ first-order statistically dominates $F(T)$. Some forms of heterogeneity - for example, if low productivity people are hurt by stress while high productivity people benefit from it, or the reverse - are inconsistent with FOSD.

Third, we test for heterogeneity by baseline variables - e.g. whether the treatment effect is a function of baseline stress, GPA, age, college year, or paternal education. Fourth, for the uncertainty group, for which we have a baseline performance measure, we can also test the hypothesis that uncertainty increases productivity both at the extensive and intensive margins.

The results, reported in Table 5, consistently reject these forms of heterogeneity: first, the Kolmogorov-Smirnov tests show that the distribution of correct answers in the control group first-order statistically dominates the one in the treatment group for men in quiz$_1$ and women in quiz$_2$; second, we never reject the null of equality of variances; third, we do not find evidence of positive effects of stress on productivity by baseline characteristics (with the exception of men with high baseline stress, for whom quiz$_2$ productivity increases; we return to this later); fourth, we do not find evidence that uncertainty increases either the likelihood of improving one's productivity, or the productivity level.

This general lack of a positive effect of stress on productivity for subsets of our subjects is consistent with the original hypothesis that this beneficial effect applies only to simple, and not to complex tasks (Yerkes and Dodson, 1908). It has useful policy implications, as, in this setting, reducing stress would not hurt anyone's productivity.

[11] Define Y_0 and Y_1 as the potential outcomes in the absence and presence of the treatment. If the treatment effects are constant and amount to T, then $Y_1 = (Y_0 + T)$. In this case the variances of Y_0 and Y_1 are identical, i.e. $Var(Y_0) = Var(Y_1)$. Note, however, that one could have heterogeneous effects even if $Var(Y_0) = Var(Y_1)$, if the covariance of the treatment effects, T, with Y0 is negative and such that $Cov(Y_0, T) = 1/2[Var(Y_0) + Var(T)]$.

5.3 Effect of stress on attempts and mistakes

Does stress make people less productive because they exert less effort? To test this hypothesis, we look at number of attempts and mistakes.[12]

Figures 3 and 4 and Table 6 show that the treatments cause people, and especially women, to attempt fewer questions (especially women under uncertainty) and make more mistakes (especially men under emotional load), suggesting that stress does not seem to cause a drop in effort (in which case, people would take less time to answer each question), but it likely decreases productivity by causing a cognitive impairment, as people make more mistakes despite taking more time to answer each question.

6 Stress and entrepreneurial choice

To test whether stress induces a change in entrepreneurial choice, we compare the choice of upgrades under the different condition. Figure 5 shows the proportion of people choosing to upgrade by treatment and gender. Comparing the proportion of people ever upgrading (the red bar in the histogram) shows that the treatments overall decrease the upgrade likelihood. The largest effect, a 16% reduction, occurs for females under EL, while the smallest one, a 6% drop, occurs for men under uncertainty.

Looking at the individual upgrade choices (the black and grey bars in the histogram) shows some further differences by gender and treatment. While for women the effects are larger than for men (in absolute level), and strongest for the highest upgrade (the one with a cutoff of 12 correct answers for a positive payment) and under high emotional load, and more heterogeneous, varying from a 63% decrease under high emotional load for the highest upgrade to a 6.5% decrease under high emotional load for the lowest upgrade, for men they are smaller, strongest for the second highest upgrade, and for the lowest upgrade (the one with a cutoff of 6 correct answers for a positive payment) and under uncertainty, and less heterogeneous, varying from a 10% *increase* for the highest upgrade under high emotional load to an 18% decrease for the second highest upgrade under uncertainty.

Table 7 presents OLS estimates of these ATTs pooling all upgrades. While the effect of the pooled treatments, a 13.5% decrease, does not differ statistically by gender, the effect of EL is higher (in absolute level) for women than men, -30% versus -4.5%, and the effect of uncertainty is

[12] There are two extreme cases of low effort. In one case, the subject could make no attempt to answer any question, having 0 correct answers, or could answer question randomly, with a positive expect payout of $\sum_i = 1^2 01/x_i$, where x_i is the number of optional answers for each question. This implicitly assumes people can do better than random if they exert effort. Since the modal number of optional answers per question is 4, the expect number of correct answer from random picks is 5. Given that none of the 390 subjects make 0 attempts at answering questions, that the average number of correct answers is 7.7 in the treatment group, and that 97.8% of the people in the treatment group run out of time before attempting to answer all the questions, we conclude that these extreme forms of low effort are not relevant for our sample.

lower for women than for men: -10% versus -17.5% (although the latter difference is not statistically significant).

In unreported regressions, we estimated ATT effects for each upgrade separately. Interestingly, the treatments cause a significant 10% reduction, from 82% of the control group upgrading, also in the likelihood of upgrading when the cutoff is 6 correct answers, despite the fact that only 1.8% of the treatment group had fewer than 6 correct answers in quiz 2. This finding suggests that the subjects underestimate their future performance. We also considered a different dependent dummy variable, which equals one as long as the subject has upgraded at least once.[13] This way, we consider upgrade choices net of potential mistakes. Overall, the pattern of results does not change in any notable way.

In sum, while the treatments reduce the likelihood of making entrepreneurial choices for both men and women, they respond differently to the same treatments. Thus, the mechanisms that are causing these effects likely differ by gender.

7 Stress and "wrong" choices

So far, we established that (1) stress causes a -13.5% decrease in upgrade likelihood, of which the highest effects are for women under high EL (-30%) and men under U (-17.5%), and that (2) productivity drops by 10% only and does not drop for men in quiz 2. Therefore, the observed decrease in entrepreneurial choice may not be fully explained by lower productivity (especially for men). This leaves room for additional explanations for why stress might affect entrepreneurial choice.

Recall that becoming less likely to make entrepreneurial choices, if one correctly anticipates that stress will reduce performance, may be the income maximizing choice. That is, conditional on stress hurting performance, it is not obvious that a drop in the upgrade likelihood is a bad thing. It is therefore important to establish whether and to what extent the effect of stress on entrepreneurial choice is income maximizing or not. For this purpose, Figure 6 and Table 8 report the incidence of "wrong" upgrade choices and of their two subcategories - upgrading too little or too much, by gender and treatment.[14]

[13] For example, suppose that one person starts to upgrade with a cutoff of 10, upgrades again with a cutoff of 8, but then forgets to "re-upgrade" with a cutoff of 6 - that is, this person will earn $0 if she has only 7 correct answers. The dependent variable in Table **??** takes the values of 0, 1, 1, 0 (in descending order), while the alternative dependent variable would be 0, 1, 1, 1.

[14] The implicit assumption for the claim that upgrade choice is wrong is that upgrading does not reduce productivity, i.e. a 'no-choking' condition. This assumption is consistent with evidence from the lab and the field, as well as with our own data. First, choking occurs with very high incentives - maximum individual earnings from a lab experiment equivalent to half of the mean yearly consumer expenditure in the village in Ariely, Gneezy, Lowenstein, and Mazar (2009); women are more likely to underperform in the high-stake MBA admission exams, compared to men, but not in low-stake exams in Ors et al. (2012). Second, in our data, people who upgrade have the same average increase in productivity from quiz_1 to quiz_2 as people who do not upgrade - in both treatment and control groups, and for both

The red bar in Figure 6 and the ATT estimates from Table 8 show that the treatments cause an increase in non-maximizing choices, higher for women than men (the difference in effect size by gender is not significant for the pooled treatments, but it is for the effect of high emotional load). In particular, the pooled treatments almost triple women's rate of suboptimal choices, which increases from 9.8% to 27.6%, while they increase men's rate by 50% only, from 11.7% to 17.6%. This difference in effect sizes is caused by women's strong sensitivity to both high emotional load and uncertainty, their ATT effects are 19.4 and 15.6 percentage point increases, while men are less sensitive to high emotional load than uncertainty, their ATT effects are 5 and 10 percentage point increases.

The yellow bar in Figure 6 and the ATT estimates from Table 8 show that most of the non-maximizing choices consist of upgrading too little, that is of not upgrading when, conditional on one's ex-post performance, one would have earned more by upgrading. For example, women's ATT of the pooled treatments on non-maximizing choices is 17.8 percentage point, and the one for upgrading to little is 13.7 percentage points, or 77%. This share is 113% for men. Again, the effect sizes are larger for women than men because women are more affected by high emotional load, although these differences are never statistically significant.

The blue bar in Figure 6 and the ATT estimates from Table **??** show that a small fraction of the non-maximizing choices consists of upgrading too much, that is of upgrading (or upgrading with too high a cutoff) when, conditional on one's ex-post performance, one would have earned more by not upgrading (or choosing an upgrade with a lower cutoff). These upgrades are 23% and -13% of non-maximizing choices for women and men. Once more, the point estimate of the ATTs is larger for women than men (the one-tailed t-test of difference is significant, but not the 2-tailed one). Note also that, while the point estimates of the effect sizes are larger under uncertainty than under high emotional load, this difference is never statistically significant.

Here are the conclusions so far. (1) Stress causes a drop in upgrade choices for men and women, partly for different reasons. Women become on average less entrepreneurial, partly because stress lowers their performance (and they likely anticipate it) and partly because stress causes a surge in non-maximizing choices. Men become less entrepreneurial despite having only a short-lived performance loss. That is, the mechanisms through which stress affects entrepreneurial choice seem to differ by gender. (2) Stress causes an increase in non income maximizing choices for women and men. While most of the non-maximizing choices consist of not upgrading often enough, stress also causes some women to upgrade too much and end up earning no income as a result. This finding shows that stress has heterogeneous effects on entrepreneurial choice for women, some of whom become"too entrepreneurial". (3) Women are not affected differently by the two stressors, unlike men, who seem to be more sensitive to uncertainty than to high emotional load, but who are nevertheless generally less affected than women. Overall, this results in women being more hurt by high emotional load than men. Conversely, uncertainty causes a significant increase in likely

men and women.

mistakes in the upgrade choice, i.e. an increase in the likelihood that the subjects will not revise the upgrade to the lowest available cutoff.

8 Decomposing the stress-induced income losses

So far, we have shown that stress reduces performance and reduces the likelihood of making entrepreneurial choices. Part of this effect of stress on choice is likely the consequence of the performance drop. However, part of the effect of stress on choice is caused by making the wrong choice, given one's $quiz_2$ productivity. Indeed, stress increases both likelihoods that people do not upgrade, when they in fact should have, and that they upgrade, when they would have earned more money by not upgrading. That is, stress make some people "too entrepreneurial."

One way to quantify the relative weight of these different channels through which stress affects behavior is to look at the treatment effects on the monetary losses from $quiz_2$ and to separate out the parts caused by reduced productivity and by non-maximizing choices. To do so, we exploit the fact that the average difference in $quiz_2$ income between the treatment and the control group, ATT(Y), is the sum of the treatment effect on productivity, ATT(P) and the treatment effect on non-maximizing choices, ATT(C), that is, $ATT(Y) = ATT(P) + ATT(C)$. We can measure ATT(Y) by comparing $quiz_2$ income in the treatment and control groups and ATT(C) by comparing the income losses in the treatment and control groups conditional on productivity. For example, suppose that a person has 10 correct answers in quiz 2 but does not upgrade. Then her losses are \$5, because she earns \$10 rather than the \$15 she would have earned, had she upgraded. Once we have ATT(Y) and ATT(C), it is easy to estimate ATT(P). Moreover, we can further decompose the losses from non-maximizing choices, ATT(C), into losses from upgrading too much, ATT(TMU), and too little, ATT(TLU): $ATT(C) = ATT(TMU) + ATT(TLU)$.

Table 9 presents all the ATT effects just described. The left panel shows that the pooled treatments cause an average loss of \$1.51, and 8.3% drop. Half of this loss comes from non-maximizing choices, i.e for making "wrong" upgrades conditional on one's productivity, while the other half comes from a productivity drop.

These aggregate results mask a substantial amount of heterogeneity by gender. The treatment cause statistically significantly larger losses for women than men (\$2.78, or -16%, vs. \$0.55, or -3%). The causes of these losses also vary by gender: while only about 36% (or \$1) of women's losses are caused by non-maximizing choices and most of the losses are caused by a drop in productivity, all (or \$0.59) of men's losses are caused by non-maximizing choices. For both genders, most of the losses from non-maximizing choices are caused by not upgrading enough (which causes losses of \$0.63 and \$0.45 for women and men) and a smaller share by upgrading too much (which causes losses of \$0.39 and \$0.13 for women and men).

8.1 Stress-induced losses by negative life events, baseline stress, and gender

People who experienced exogenous stressors, such as negative life events, are more vulnerable to additional stress. The literature has shown that women are more likely to experience exogenous stressors (Dahl and Moretti, 2008) and they have a more acute response to the same stressors. It is therefore useful to examine whether the differential response of stress by gender interacts with baseline stressors, as well as with self-reported baseline stress.

We have three different proxies, ranging from long-term to short-term measures of stress: having a father with or without a college degree, which, conditional on subject's GPA, can be considered a proxy for financial stability in the household of origin; having experienced parental divorce or the death of a close person in the previous five years - events that are conceivably beyond the subject's control; self-reported stress in the previous four weeks. We create dummies for above median self-reported stress so that we can split the sample in two, as for the paternal education and life events variables.[15] These variables are uncorrelated with each other (the correlation coefficients are never statistically significant and very small, ranging from -0.161 to 0.039 for the correlations of above median self-reported stress with above median baseline cortisol and negative life events), suggesting that they capture different aspects of stress.

We estimate versions of equations (2) and (2) in which we interact the treatment dummies by the baseline stress dummies. We group divorce, death, and low paternal education into a single dummy variable that equals one for subjects who experienced at least one of the three events, and zero otherwise.[16]

The estimated ATT effects by negative life events, presented in Table 10, show that there are no statistically significant effects of the treatments on women and men who experienced none of the aforementioned events (with the exception of small negative effects on losses from men's 'wrong' upgrade choices) and small productivity losses for men who experienced at least one event. Conversely, the treatment cause large losses for women who experienced at least one event. High emotional load and uncertainty cause these women to lose \$4.3 and \$5. \$1.1 and \$2 of these losses are caused by an increase in the wrong upgrade choice, and the remaining \$3 by a productivity drop.

The estimated ATT effects by baseline self-reported stress, presented in Table 11, show that, the treatments cause equally negative effects for men and women with below median baseline stress and women with above median baseline stress, they cause no statistically significant effects on men with above median baseline stress.

In sum, there are differential responses to stress for men and women depending on baseline

[15] We also observe baseline cortisol for a subsample of 92 subjects. However, because of this small size, we consider these data not reliable, as the comparisons across eight groups - men and women, high and low cortisol, control and treatment - would leave very few observations in each cell.

[16] We found similar patterns of results by having one variable for divorce and death and a second one for low paternal education.

conditions, and, more specifically, more negative impacts of the experimentally induced stress for certain subsets of women.

9 Stress and non-maximizing choices: some mechanisms

The treatments increase non-maximizing choices, mainly inducing people to be less entrepreneurial than what they should be based on their ex-post observed performance. The non-maximizing reduction in upgrades, i.e. upgrading "too little", may be caused by an increase in pessimism (expected performance worse than actual performance), risk aversion, and mistakes (failing to upgrade despite wanting to). Conversely, the non-maximizing increase in upgrades, i.e. upgrading "too much" and ending up with 0 income, may be caused by stress increasing optimism (expected performance better than actual performance), risk-taking behavior, and mistakes (failing to revise upgrade down despite wanting to). Indeed, it is possible that stress may have heterogeneous effects. We explore these potential mechanisms in the following section.

9.1 Stress and non-maximizing choices: pessimistic and optimistic beliefs

Stress may further affect entrepreneurial decisions by changing beliefs. Beliefs can be accurate or inaccurate, and inaccurate beliefs can be optimistic or pessimistic, depending on whether the difference between the expected and realized performance is positive or negative. We ask a series of questions to verify whether people correctly anticipate the effect (or lack of an effect) of stress on their performance, and, if not, whether they underestimate (optimism) or overestimate (pessimism) this effect.

We measure beliefs and their accuracy in two ways. First, before each upgrade for quiz 2, we ask what the subjective probability of having the required number of correct answers to qualify for the upgraded payment is. This gives us four subjective probabilities of having at least 12, 10, 8, or 6 correct answers. Comparing subjective and observed probabilities gives a measure of belief accuracy, optimism, and pessimism.

Second, right before starting quiz$_2$, we ask the subjects how many correct answers from quiz$_2$ they expect to have. Comparing expected and observed performance gives a measure of overall belief accuracy, optimism, and pessimism. Moreover, we can use the expected and observed probabilities to compute an additional, yet indirect, measure of beliefs. This is a weighted average of having 6, 8, 10, and 12 correct answers, where the weights are the difference between the subjective and observed probabilities in each case. This latter variable is a truncated mean, as we do not observe the subjective probabilities of having all the possible correct answers.

Note that, for these particular outcomes, in principle the uncertainty treatment has the additional effect of increasing the variance of beliefs, because people under this condition do not know their quiz 1 performance. Suppose, for simplicity, that people's performance were the same in quiz 1 and 2, and that people knew it. In that case, we'd expect the variance of beliefs to be zero for the control and the high emotional load groups, but positive for the uncertainty group, because its members have to guess their quiz 1 performance. By doing that, they would overestimate the probabilities of having both higher and lower outcomes than their actual ones, ending up having (1) optimistic beliefs in the right tail, (2) pessimistic beliefs in the left tail, and (3) overall less accurate beliefs. In sum, evidence of optimism and pessimism in the tails for the uncertainty group is not necessarily evidence that stress changes beliefs. However, mechanical changes in accuracy of beliefs induced by lack of knowledge of previous performance per se should be mean invariant. Finding changes in expected performance is consistent with the hypothesis that uncertainty-induced stress changes beliefs beyond this mechanical effect. For this purpose, we look at how stress affects the difference between quiz 2 expected and observed probabilities and performance.

Figure 7 shows the difference between subjective and observed (1) probabilities (the first 4 bars) and (2) productivity (the last 2 bars), using both the direct elicitation and the truncated mean computed indirectly by using stated and observed probabilities. People in the control group tend to have pessimistic beliefs. High emotional load either does not change (for men) or actually increases pessimism (for women). As expected, uncertainty increases the variance of beliefs - coarsely measured by the difference between the black and the lighter grey bars - for both men and women. However, this shift is asymmetric and makes people, especially women, more optimistic. This suggests that uncertainty affects beliefs in ways that go above and beyond its mechanical effect through higher belief variance.

The increase in optimism caused by uncertainty is especially marked in the left tail, i.e. for the probabilities of having at least 12 correct answers. Figure 8 suggests that this occurs because the subjects, and especially women, fail to recognize the negative effect of uncertainty on this likelihood. Indeed, the probabilities of having at least 12 and 10 correct answers are the same for women in the control and uncertainty groups, but the performance drops substantially in the uncertainty group, while the probabilities of having at least 8 and 6 correct answers are lower under uncertainty.

Tables 12 and 13 show the ATT effect on optimism for the likelihoods of having at least 12 and 6 correct answers and confirm the findings discussed above: the treatments cause economically and statistically significant changes in belief accuracy for women only (although the difference in impact sizes by gender is not statistically significant); the increase in optimism for the likelihood of having at least 12 correct answers occurs for women only and is induced by uncertainty, while the decrease in optimism for the likelihood of having at least 6 correct answers occurs for women only but is equally affected by high emotional load and uncertainty.

Table 14 shows that both aggregate measures of the difference between subjective and objective productivity significantly increase for women under uncertainty. That is, uncertainty overall increases women's optimism and increases belief accuracy in the direct difference measure by making women less pessimistic.

The takeaways from this exercise are that both types of non-maximizing upgrades may have been affected by changes in beliefs, to some extent. The increased optimism in the likelihood of having at least 12 correct answers may have induced some people to mistakenly upgrade when they should not have done so, while the increased pessimism in the likelihood of having at least 6 correct answers may have induced some women (as this does not happen to men) to not upgrade when they should have done so.[17]

9.2 Stress and non-maximizing choices: confusion and mistakes

Stress may cause a temporary cognitive impairment. Indeed, there is evidence that stress affects both working and declarative memory (Lupien, Gaudreau, Tchiteya, Maheu, Sharma, Nair, Hauger, McEwen, and ,1997; Lupien *et al*, 2007). This is consistent with the treatment group's observed drop in performance and increase in failures to revise the upgrade down to the lower cutoffs. To provide further evidence on this potential mechanism, we study how the treatments affect the incidence of inconsistent choices, beliefs, and or mistakes. To do that, we consider whether people make inconsistent choices in upgrades (e.g. upgrading with a cutoff of 10 but not revising the upgrade down when the cutoff is 8), probabilities (when the reported probability of having at least X correct answers is greater than the probability of having $Y < X$ correct answers), and lotteries (when subjects switch back and forth between the high and low-risk lotteries), as well as whether they fail to recall the answers related to the video clip (a paid task). Figure 9 shows the share of people making inconsistent choices or making recall errors, and the average number of domains in which people make mistakes (multiple inconsistencies in one domain - e.g. multiple inconsistent answers in reported probabilities, count as one inconsistency). The biggest effect sizes in the number of domains in which people make inconsistent choices and mistakes occur under uncertainty, which causes a 66% increase for men (from 0.18 to 0.30) and a 48% increase for women (from 0.27 to 0.40). However, there is heterogeneity in the effects in each domains as well as in the effects by gender. For example, high emotional load causes a 33% decrease in the number of domains in which women are inconsistent or make mistakes (from 0.27 to 0.18). High emotional load also increases men's likelihood of making recall errors from 0 to 0.037, while neither treatment affects women in this domain.

Table 15 reports the estimates of the ATT effect for number of domains in which people are inconsistent or make recall errors. We find statistically significant increases for men, especially under uncertainty, and smaller (or negative) and not statistically significant effects for women (although the difference in effect sizes by gender is not statistically significant). We also show estimates of the recall error likelihood for two reasons: first, it is a low-effort, high-return task (people are paid $15 for correctly recalling whether they answered yes or no to two questions asked about 15 minutes earlier); second, it is one of the various examples in which high emotional load causes a cognitive impairment, a result that is missed by looking at the aggregate measure only, from which it appears that uncertainty is the only condition that causes increases in inconsistencies

[17] The optimism variables are strongly positively correlated within subject. Therefore, the treatments likely have heterogeneous effects, inducing some subjects to become more optimistic and others more pessimistic.

and mistakes (high emotional load also causes an increase in women's inconsistencies in reported probabilities). As previewed, the treatments, and especially high emotional load, cause an increase in men's wrong recall from 0 to 2.6 (and 4) percentage points for the pooled (and uncertainty) treatments.

In sum, there is evidence that stress causes men, but not women, to become inconsistent and make more mistakes.

9.3 Stress and non-maximizing choices: risk preferences

Figure 10 shows that stress seems to have opposite effects by gender: while it increases the heterogeneity in risk preferences for men, it decreases it for women, without much apparent difference by treatment type. Table 16 confirm these findings: the treatments statistically significantly increase the share of risk loving and very risk averse men by 5 and 10 percentage points, while they reduce the share of risk neutral and risk averse men by 7.7 and 13 percentage points. Nevertheless, the increase in risk-taking behavior for some men does not cause excessive upgrades (as the treatments do not increase men's likelihood of upgrading too much), while the increase in risk aversion for some may be one of the mechanisms through which stress increases the share of men who upgrade too little.

Conversely, the treatments cause (imprecisely estimated) 10 and 6 percentage point drops in the share of risk neutral and very risk averse women, and increase the share of risk averse women by 15 percentage points. Therefore, it appears that stress is not causing some women to upgrade too much because they have become more risk loving, while, as for men, it may increase the share of women who upgrade too little by increasing their risk aversion.

These findings are somewhat consistent with the existing literature on the effect of stress on risk-taking behavior. For example, Lighthall, Mather, and Gorlick (2009) find differential response to stress on risk taking by gender, with stress causing women and men to decrease and increase risk taking. However, to our knowledge no study investigates the heterogeneous effects of stress on risk by gender.

Note that estimates of the average treatment effects on risk aversion were never statistically significant. For example, the average treatment effect on the risk aversion parameter, which, in the data, varies between -0.72 to 1.17, is 0.01 and statistically not different from zero (s.e. 0.04). This is one clear example of average impacts missing distributional effects.

Lastly, the effects do not systematically vary by treatment type.

10 Stress or Cognitive Load?

To rule out that the effects shown are not driven by cognitive load, we ran experiments with the same structure, but a different treatment. Rather than emotional load and uncertainty, the treatment group is asked to memorize an 8-digit number, while treatment group memorizes a 2-digit number. The subjects can earn \$15 for the successful recall of the number at the end of the experiment (if the task selected out of the 4 compensated tasks is the number recall). We find that the treatment does affect recall: while everybody correctly recalls the number in the control group, the recall rate is only 49% in the treatment group. We deduce that this memorization task increases the cognitive load. However, this treatment does not reduces productivity.

11 Conclusion and Discussion

This paper has shown that stress reduces productivity and increases the likelihood that people will make choices that do not maximize their income. Our experiment design successfully elicited controlled stress in the lab. The women in our sample are more likely to be exposed to stressors beyond their control and report higher levels of stress. Their entrepreneurial choices and performance are more negatively affected by experimentally induced stress than men's, resulting in lower monetary gains. These gender differences are driven by women who experienced a parental divorce or the death of a close person in the previous 5 years, or whose father does not have a college degree: these women are especially susceptible to the two stressors we induce in the lab, which cause them to lose as much as \$5, or 27%, on the task whose completion is a function of both productivity and choice of compensation.

We also study three different mechanisms through which stress might affect entrepreneurial choice and performance - beliefs, mistakes, and risk attitudes - and find that the reasons for excessive upgrades are a surge in optimism for some women, and a surge in optimism, mistakes, and risk-loving preferences for some men. The reasons for upgrading too little, on the other hand, are a surge in pessimism for some women and an increase in mistakes and risk aversion for men.

An obvious policy implication of these findings is that removing the underlying sources of stress may improve entrepreneurial activities and performance for all, and more so for women than men. However, it is often hard to do so, and removing external stressors would not address women's higher vulnerability to on-the-job stress. If the stressors are not easy to remove or if one wants to understand how to help women deal with both external and on-the-job stress, to design effective policies one should know whether the observed response to stress should be changed or not, and, if so, what mechanisms are causing the response to stress.

There are two relevant conclusions from our experiment. First, stress worsens productivity, especially women's, through a temporary cognitive impairment (more mistakes, longer time required to answer questions). Interventions that help recover from this impairment - from providing training or mentoring programs, to devising a protocol to double-check one's choices, to various psychological and physiological interventions - would be effective for both men and (more so) women. Second, while stress causes all subjects, but especially women, a decrease in the likelihood of making an entrepreneurial choice, not making this choice actually maximizes their earnings for a large group, given their lower productivity under stress. Therefore, policies that indiscriminately favor entrepreneurship might hurt some people. For these people, the optimal policy response can be devised after knowing the mechanisms that might cause the wrong choices. Policies that help people form accurate beliefs might be especially useful for women, while policies that help people make fewer mistakes or affect the perception and the reaction to risk might be especially useful for men.

References

[1] American Psychological Association (2011), *Stress in America^{TM}: Our Health at Risk*, retrieved on August 6th, 2012, from `http://www.apa.org/news/press/releases/stress/2011/final-2011.pdf`

[2] Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg (2001),"The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation", *Journal of Economic Perspectives*, 15(3), 47-68.

[3] Anxo, Dominique, Lennart Flood, Letizia Mencarini, Ariane Pailh, Anne Solaz, and Maria Letizia Tanturri (2007), *Time Allocation between Work and Family over the Life-Cycle: A Comparative Gender Analysis of Italy, France, Sweden and the United States*, Institute for the Study of Labor, IZA Discussion Papers 3193.

[4] Ariely, D., Gneezy, U., Lowenstein, G. and Mazar, N. (2009), "Large Stakes and Big Mistakes," *Review of Economic Studies*, 76, 451469.

[5] Arsenault, André, and Shimon Dolan (1983), "The Role of Personality, Occupation, and Organization in Understanding the Relationship Between Job Stress, Performance, and Absenteeism", *Journal of Occupational Psychology*, 56, 227-240.

[6] Banerjee, Abhijit, and Sendhil Mullanaithan (2008), "Limited Attention and Income Distribution", *American Economic Review*, 98, 489-493.

[7] Barber, Brad and Terrance Odean (2001), "Boys Will Be Boys: Gender, Overconfidence, and common stock investment", *The Quarterly Journal of Economics*, 116 (1), 261-292.

[8] Barbulescu, Roxana and Matthew Bidwell (2012), "Why are there so few women in finance? Evidence from the MBA job search process on the mechanisms of gender segregation", *Organization Science*, published online before print June 15, doi:10.1287/orsc.1120.0757

[9] Berns, Gregory S., David Laibson, and George Loewenstein (2008), "Intertemporal choicetoward an integrative framework", *Trends in Cognitive Science*, 11, 482-488.

[10] Bertrand, Marianne, and Kevin F. Hallock, "The Gender Gap in Top Corporate Jobs," Industrial and Labor Relations Review, LV (2001), 321.

[11] Bianchi, Suzanne M. (2010), *Family Change and Time Allocation in American Families*, presented at the "Focus on Workplace Flexibility" Conference, Washington, D.C. November 30th.

[12] Bianchi, Suzanne M. (2011), "Family Change and Time Allocation in American Families", *The ANNALS of the American Academy of Political and Social Science*, 638 (1), 21-44.

[13] Brooks, A. and Schweitzer, M. (2011), "Can Nervous Nelly Negotiate? How Anxiety Causes Negotiators to Make Low First Offers", The Wharton School, University of Pennsylvania, Working Paper # 2011-04.

[14] Camerer, Colin (1990), "Comment on Noll and Krier, 'Some Implications of Cognitive Psychology for Risk Regulation' ", *The Journal of Legal Studies*, 19, 791-799.

[15] Camerer, Colin, George Loewenstein, and Drazen Prelec (2005), "Neuroeconomics: How Neuroscience Can Inform Economics", *Journal of Economic Literature*, 43, 9-64.

[16] Cech, Erin, Brian Rubineau, Susan Silbey, and Caroll Seron (2011), "Professional Role Confidence and Gendered Persistence in Engineering", *American Sociological Review*, 76 (5), 641-666.

[17] Chabris, Christopher F., David Laibson, Carrie L. Morris, Jonathon P. Schuldt, Dmitry Taubinsky (2008), "Individual laboratory-measured discount rates predict field behavior", *Journal of Risk and Uncertainty*, 37, 237-269

[18] Chen, Yan, Peter Katuščák, and Emre Ozdenoren (2013), "Why can't a woman bid more like a man?", *Games and Economic Behavior*, 77, 181-213.

[19] Cohen, Sheldon, Tom Kamarck, and Robin Mermelstein (1983), "A global measure of perceived stress", *Journal of Health and Social Behavior*, 24, 385-396.

[20] Cohen, Sheldon, and Gail Williamson (1988), "Perceived stress in a probability sample of the United States", in S. Spacapam and S. Oskamp (Eds.), *The social psychology of health: Claremont Symposium on applied social psychology*, Newbury Park, Sage, 31-47.

[21] Croson, Rachel and Uri Gneezy (2009), "Gender Differences in Preferences", *Journal of Economic Literature*, 47 (2), 1-27.

[22] Dahl, Gordon, and Enrico Moretti (2008), "The Demand for Sons", *The Review of Economic Studies*, 75, 1085-1120.

[23] Diamond, David M., Adam M. Campbell, Collin R. Park, Joshua Halonen, and Phillip R. Zoladz (2007), "The Temporal Dynamics Model of Emotional Memory Processing: A Synthesis on the Neurobiological Basis of Stress-Induced Amnesia, Flashbulb and Traumatic Memories, and the Yerkes-Dodson Law", *Neural Plasticity*, vol. 2007, Article ID 60803, doi:10.1155/2007/60803

[24] Dickerson, Sally S., and Margaret E. Kemeny (2004), "Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research", *Psychological Bulletin*, 130, 355-391.

[25] Duflo, Esther (2006), "Field Experiments in Development Economics", in Richard Blundell, Whitney Newey, Torsten Persson (eds.), *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, Cambridge: Cambridge University Press, 2(42), 322-348.

[26] Duflo, Esther (2006), "Poor but Rational?", in Abhijt Banerjee, Dilip Mookherjee and Roland Benabou (eds.), *Understanding Poverty*, New York: Oxford University Press, 367-78.

[27] Duflo, Esther, and Abhijit V. Banerjee (2009), "The Experimental Approach to Development Economics", *Annual Review of Economics*, 1, 151-178.

[28] Dupas, P. and and Robinson, J. (2009), *Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya*, NBER Working Paper No. 14693.

[29] Earl, Peter (1990), "Economics and Psychology: A Survey", *The Economic Journal*, 100(402), 718-755.

[30] Flory, Nicole and Elvira Lang (2010), "Diagnosis Uncertainty Increases Anxiety in Patients", Radiological Society of North America (RSNA), Press Release, retrieved on July 5th, 2012 from `http://www2.rsna.org/timssnet/media/pressreleases/pr_target.cfm?ID=510`

[31] Fredrickson, B. L. and Branigan, C. (2005), "Positive emotions broaden the scope of attention and thought-action repertoires", *Cognition and Emotion*, 19 (3), 313-332.

[32] Fudenberg, Drew (2006), "Advancing beyond 'Advances in Behavioral Economics' ", *Journal of Economic Literature*, 44, 694-711.

[33] Fudenberg, Drew, and David K. Levine (2006), "A Dual-Self Model of Impulse Control", *American Economic Review*, 96, 1449-1476.

[34] Gneezy, Uri, John A. List and George Wu (2006), "The uncertainty effect: When a risky prospect is valued less than its worst possible outcome", *Quarterly Journal of Economics*, 121, 1283-1309.

[35] Gross, James J., and Robert W. Levenson (1995), "Emotion elicitation using films", *Cognition and Emotion*, 9, 87-108.

[36] Hagg, Ingemund, Jan Johanson, and Dick Ramstrom (1970), "The Business Game as a Research Tool: Experiments on the Effects of Received Information on the Decision Process", *The Swedish Journal of Economics*, 72, 111-123.

[37] Harrison, Glenn W., Morten I. Lau, E. Elisabet Rutstrm, and Melonie B. Sullivan, Melonie B. (2005), "Eliciting Risk and Time Preferences Using Field Experiments: Some Methodological Issues", in Jeffrey Carpenter, Glenn W. Harrison and John A. List (eds.), *Field Experiments in Economics (Research in Experimental Economics, Volume 10)*, Bingley: Emerald, 125-218.

[38] Hirsh, Jacob B. and Michael Inzlicht (2008), "The Devil You Know: Neuroticism Predicts Neural Response to Uncertainty", *Psychological Science*, 19 (10), 962-967.

[39] Holt, Charles A., and Laury, Susan K. (2002), "Risk Aversion and Incentive Effects", *American Economic Review*, 92, 1644-1655.

[40] Holt, Charles A., and Laury, Susan K. (2005), "Risk Aversion and Incentive Effects: New Data Without Order Effects", *American Economic Review*, 95, 902-912.

[41] Kavanagh, Jennifer, 2005, *Stress and performance: a review of the literature and its applicability to the military*, RAND Corporation, Santa Monica, CA.

[42] Kelley, D, Singer, S., and Herrington, M. (2012), *The Global Entrepreneurship Monitor - 2011 Global Report*, mimeo, Global Entrepreneurship Monitor.

[43] Kemeny, Margaret E. (2003), "The Psychobiology of Stress", *Current Directions in Psychological Science*, 12, 124-129.

[44] Kiecolt-Glaser, J.K., R. Glaser, J.T. Cacioppo, W.B. Malarkey (1998), "Marital stress: immunologic, neuroendocrine, and autonomic correlates", *Annals of the New York Academy of Sciences*, 840, 656663.

[45] Kiecolt-Glaser, J.K., T. Newton, J.T. Cacioppo, R.C. MacCallum, R. Glaser, W.B. Malarkey (1996), "Marital conflict and endocrine function: are men really more physiologically affected than women?", *Journal of Consulting and Clinical Psychology*, 64, 324-332.

[46] Kirchsteiger, George, Luca Rigotti, and Aldo Rustichini (2006), "Your morals might be your moods", *Journal of Economic Behavior and Organization*, 59, 155-172.

[47] Kirschbaum, Clemens, Brigitte Kudielka, Jens Gaab, Nicole Schommer, and Dirk Hellhammer (1999), "Impact of gender, menstrual cycle phase and oral contraceptives on the activity of the hypothalamicpituitaryadrenal axis", *Psychosomatic Medicine*, 61 (2), 154162.

[48] Kivimäki, Mika, Päivi Leino-Arjas, Ritva Luukkonen, Hilkka Riihimäki, Jussi Vahtera, and Juhani Kirjonen (2002), "Work stress and risk of cardiovascular mortality: prospective cohort study of industrial employees", *British Medical Journal*, 325, 857-860.

[49] Koellinger, Philipp, Maria Minniti and Christian Schade (2013), "Gender Differences in Entrepreneurial Propensity", *Oxford Bulletin of Economics and Statistics*, 75 (2), 213-234.

[50] Kudielka, B.M., and Kirschbaum, C. (2005), "Sex differences in HPA axis responses to stress: a review", *Biological Psychology*, 69, 113132.

[51] Lee, Joohyung and Vincent R. Harley (2012), "The male fight-flight response: A result of SRY regulation of catecholamines?", *Bioessays*, 34, 454457.

[52] Lerner, Jennifer S., Deborah A. Small, and George Loewenstein (2004), "Heart Strings and Purse Strings: Carryover Effects of Emotions on Economic Decisions", *Psychological Science*, 15, 337-341.

[53] Lerner, Jennifer S., Seunghee Han, and Dacher Keltner (2007), "Feelings and consumer decision making: extending the appraisal-tendency framework", *Journal of Consumer Psychology*, 17, 184-187.

[54] Levine, Ari, Orna Zagoory-Sharon, Ruth Feldman, John G. Lewis, and Aron Weller (2007), "Measuring cortisol in human psychobiological studies", *Physiology and Behavior*, 90, 4353.

[55] Lighthall Nichole R., Mara Mather, and Marisa A. Gorlick (2009), "Acute Stress Increases Sex Differences in Risk Seeking in the Balloon Analogue Risk Task", *PLoS ONE*, 4 (7), e6002, doi:10.1371/journal.pone.0006002

[56] Linden, W., T.L. Earle, W. Gerin, and N. Christenfeld (1997), "Physiological stress reactivity and recovery: conceptual siblings separated at birth?", *Journal of Psychosomatic Research*, 42, 117-135.

[57] Loewenstein, George (1996), "Out of control: Visceral influences on behavior", *Organizational Behavior and Human Decision Processes*, 65, 272-92.

[58] Loewenstein, George (2000), "Emotions in economic theory and economic behavior", *American Economic Review*, 90, 426-432.

[59] Loewenstein, George, Elke U. Weber, Christopher K. Hsee, and Ned Welch (2001), "Risk as feelings", *Psychological Bulletin*, 127, 267-286.

[60] Lupien, S., S. Gaudreau, B. Tchiteya, F. Maheu, S. Sharma, N. Nair, R.L. Hauger, B.S. McEwen, and M.J. Meaney (1997), "Stress-Induced Declarative Memory Impairment in Healthy Elderly Subjects: Relationship to Cortisol Reactivity", *The Journal of Clinical Endocrinology and Metabolism*, 82(7), 2070-2075.

[61] Lupien, Sonia J., Francoise Maheu, Mai Tu, Alexandra Fiocco, and Tania Schramek (2007), "The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition", *Brain and Cognition*, 65 (3), 209-237.

[62] Mauno, Saijo and Ulla Kinnunen (2002), "Perceived job insecurity among dual-earner couples: Do its antecedents vary according to gender, economic sector and the measure used?", *Journal of Occupational and Organizational Psychology*, 75(3), 295314.

[63] Meglino, B. M. (1977), "Stress and Performance: Are They Always Incompatible?", *Supervisory Management*, 22 (3), 2-12.

[64] Muraven, Mark, and Roy F. Baumeister (2000), "Self-regulation and depletion of limited resources: Does self-control resemble a muscle?", *Psychological Bulletin*, 126, 247-259.

[65] National Institutes of Health (2002), *Stress System Malfunction Could Lead to Serious, Life Threatening Disease*, NIH Backgrounder, retrieved from http://www.nih.gov/news/pr/sep2002/nichd-09.htm

[66] Niederle, Muriel and Lise Vesterlund (2007), "Do Women Shy Away From Competition? Do Men Compete Too Much?", *The Quarterly Journal of Economics*, 122 (3), 1067-1101.

[67] Niederle, Muriel and Lise Vesterlund (2011), "Gender and Competition", *Annual Review of Economics*, 3, 601-630.

[68] Ors, Evren, Frédéric Palomino, and Eloïc Peyrache (Forthcoming), "Performance Gender-Gap: Does Competition Matter?", Journal of Labor Economics, 31 (3).

[69] Oswald, Andrew J., Eugenio Proto, and Daniel Sgroi (2009), *Happiness and Productivity*, IZA Discussion Paper No. 4645.

[70] Ozdenoren, Emre, Stephen W. Salant, and Dan Silverman (2012), "Willpower and the Optimal Control of Visceral Urges", *Journal of the European Economic Association*, 10, 342-368.

[71] Pearson, Matthew and Burkhard C. Schipper (2013), "Menstrual Cycle and Competitive Bidding", *Games and Economic Behavior*, 78, 1-20.

[72] Raghunathan, Rajagopal, and Michel T. Pham (1999), "All negative moods are not equal: Motivational influences of anxiety and sadness on decision making", *Organizational Behavior and Human Decision Processes*, 79, 56-77.

[73] Sapolsky, R.M., L.M. Romero, A.U. Munck (2000), "How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions", *Endocrine Review*, 21, 5589.

[74] Shamosh, Noah A., and Jeremy R. Gray (2008), "Delay discounting and intelligence: A meta-analysis", *Intelligence*, 36, 289-305.

[75] Shiv, Baba, George Loewenstein, G., Antoine Bechara, Hanna Damasio, and Antonio R. Damasio (2005), "Investment behavior and the negative side of emotion", *Psychological Science*, 16, 435-439.

[76] Shiv, Baba, and Alexander Fedorikhin (1999), "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making", *Journal of Consumer Research*, 26 (3), 278-292.

[77] Small, Deborah A., and Jennifer S. Lerner (2008), "Emotional policy: Personal sadness and anger shape judgments about a welfare case', *Political Psychology*, 29, 149-168.

[78] Strauss, John and Duncan Thomas (2008), "Health over the life course", in Paul Schultz and John Strauss (eds.), *Handbook of development economics*, Vol. 4, North-Holland, 3375-3474.

[79] Taylor, Shelley E., Jennifer S. Lerner, Rebecca M. Sage, Barbara J. Lehman, and Teresa E. Seeman (2004), "Early environment, emotions, responses to stress, and health", *Journal of Personality*, 72, 1365-1393.

[80] Taylor, Shelley E., Jennifer S. Lerner, David K. Sherman, Rebecca M. Sage, and Nina K. McDowell (2003), "Are self-enhancing cognitions associated with healthy or unhealthy biological profiles?", *Journal of Personality and Social Psychology*, 85, 605-615.

[81] Taylor, Shelley E., Laura C. Klein, Brian P. Lewis, Tara L. Gruenewald, Regan A.R. Gurung, and John Updegraff (2000), "Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight", *Psychological Review*, 107, 411-429.

[82] Thaler, Richard, and Hersh Shefrin (1981), "An Economic Theory of Self-Control", *Journal of Political Economy*, 89, 392-406.

[83] Vohs, Kathleen D., Roy F. Baumeister, Brandon J. Schmeichel, Jean M. Twenge, Noelle M. Nelson, and Dianne M. Tice (2008), "Making choices impairs subsequent self-control: A limited-resource account of decision making, self-regulation, and active initiative", *Journal of Personality and Social Psychology*, 94, 883-898.

[84] Vohs, Kathleen D. and Ronald J. Faber (2007), "Spent Resources: Self-Regulatory Resource Availability Affects Impulse Buying", *Journal of Consumer Research*, 33, 537-547.

[85] Ward, Andrew, and Traci Mann (2000), "Don't mind if I do: Disinhibited eating under cognitive load", *Journal of Personality and Social Psychology*, 78, 753-763.

[86] Wolfinger, Nicholas H., Mary Ann Mason, and Marc Goulden (2008), "Problems in the Pipeline: Gender, Marriage, and Fertility in the Ivory Tower", *The Journal of Higher Education*, 79 (4), 388-405.

[87] Yerkes R.M., J.D. Dodson (1908), "The relation of strength of stimulus to rapidity of habit-formation", *Journal of Comparative Neurology and Psychology*, 18, 459-482.

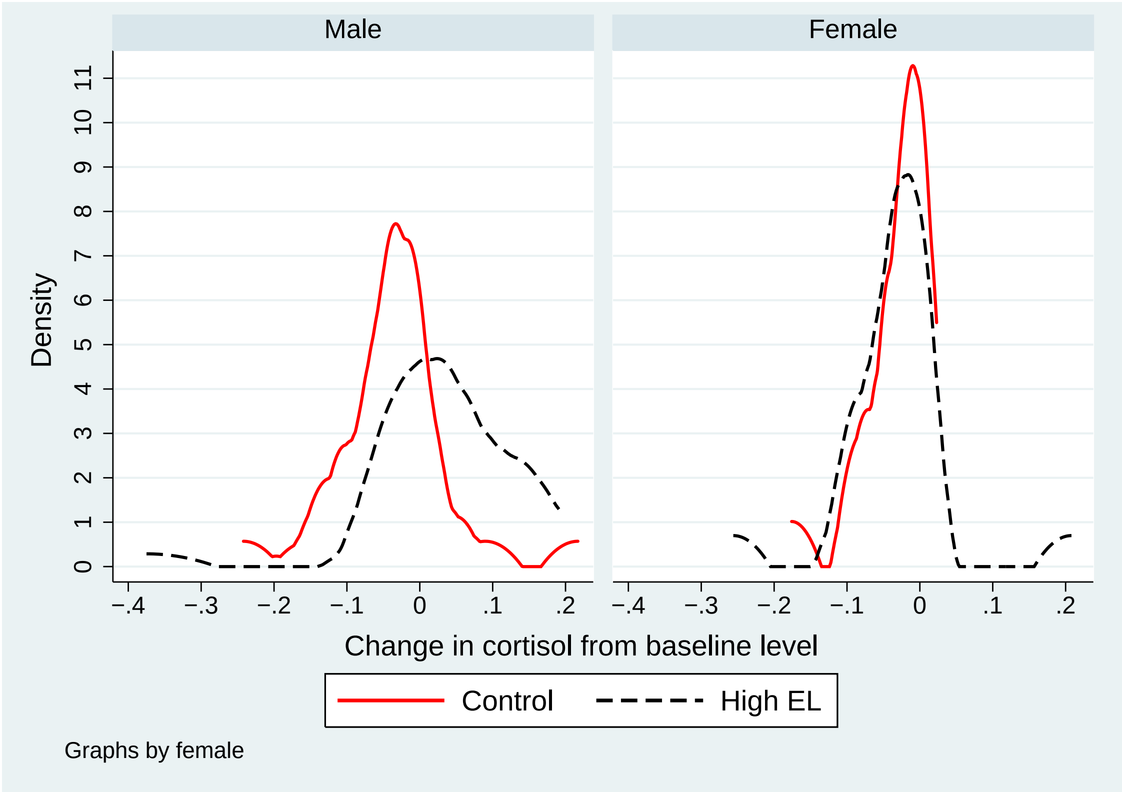

Figure 1: Change in cortisol from baseline levels by gender

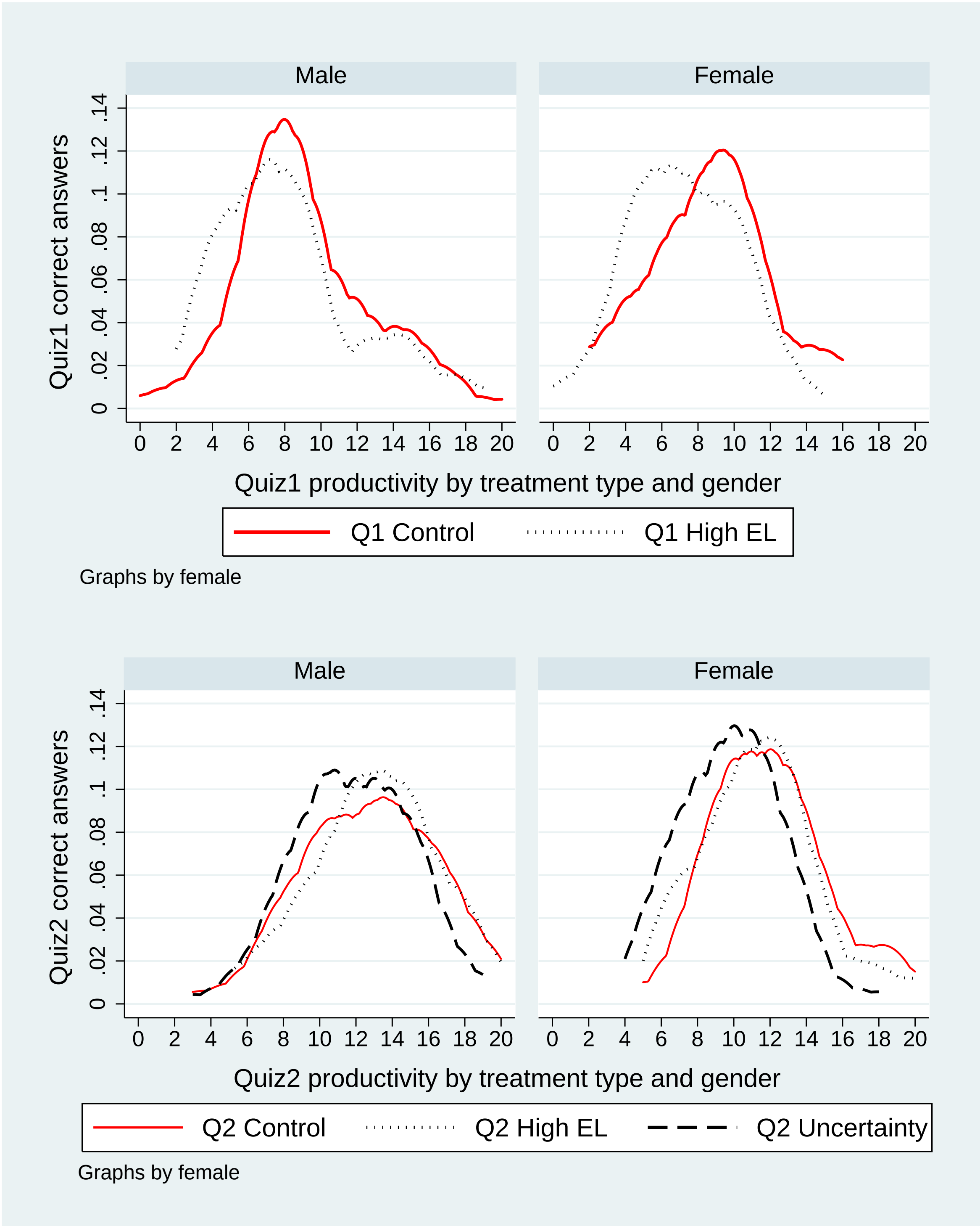

Figure 2: Quiz 1 and 2 correct answers by treatment and gender

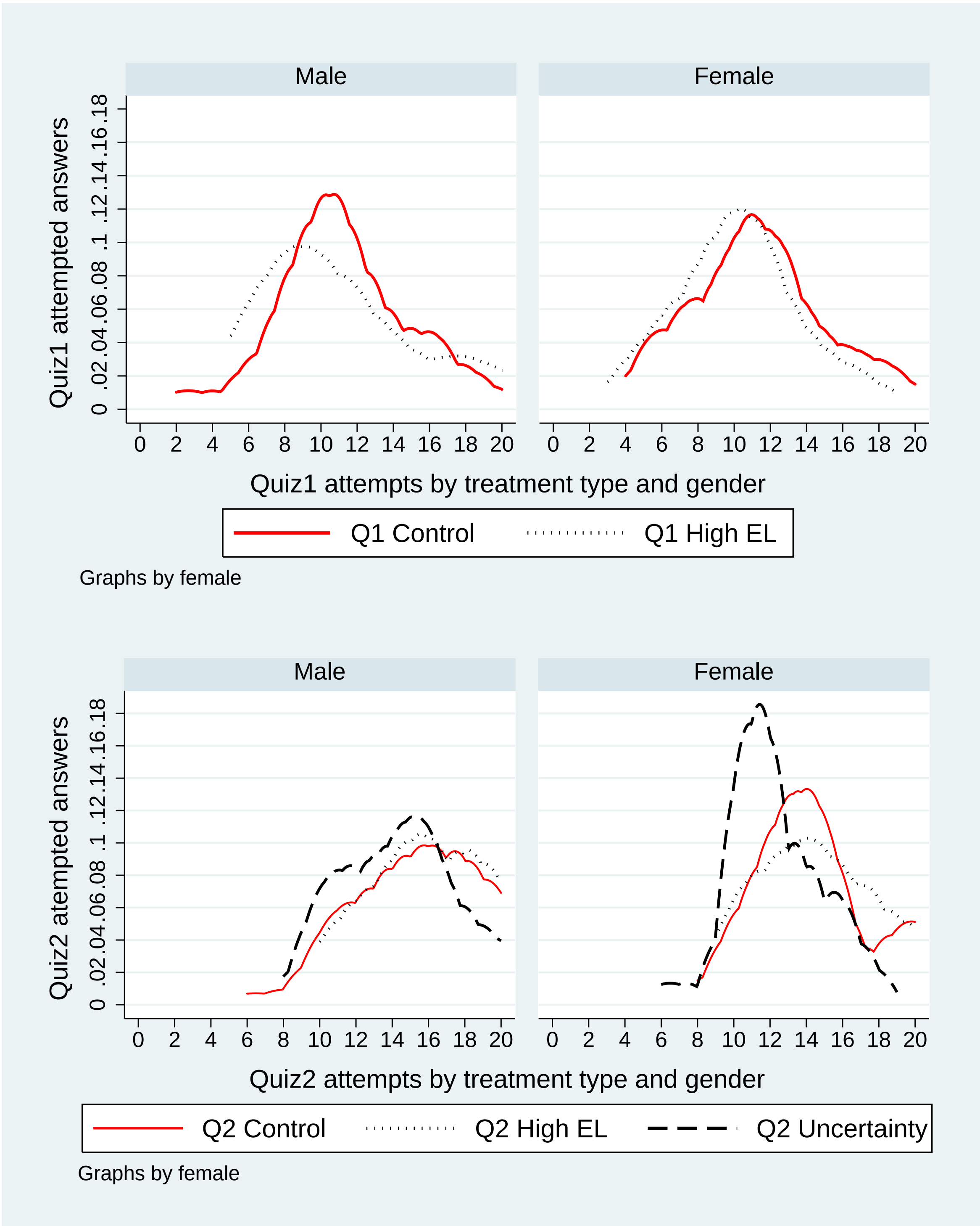

Figure 3: Quiz 1 and 2 attempted answers by treatment and gender

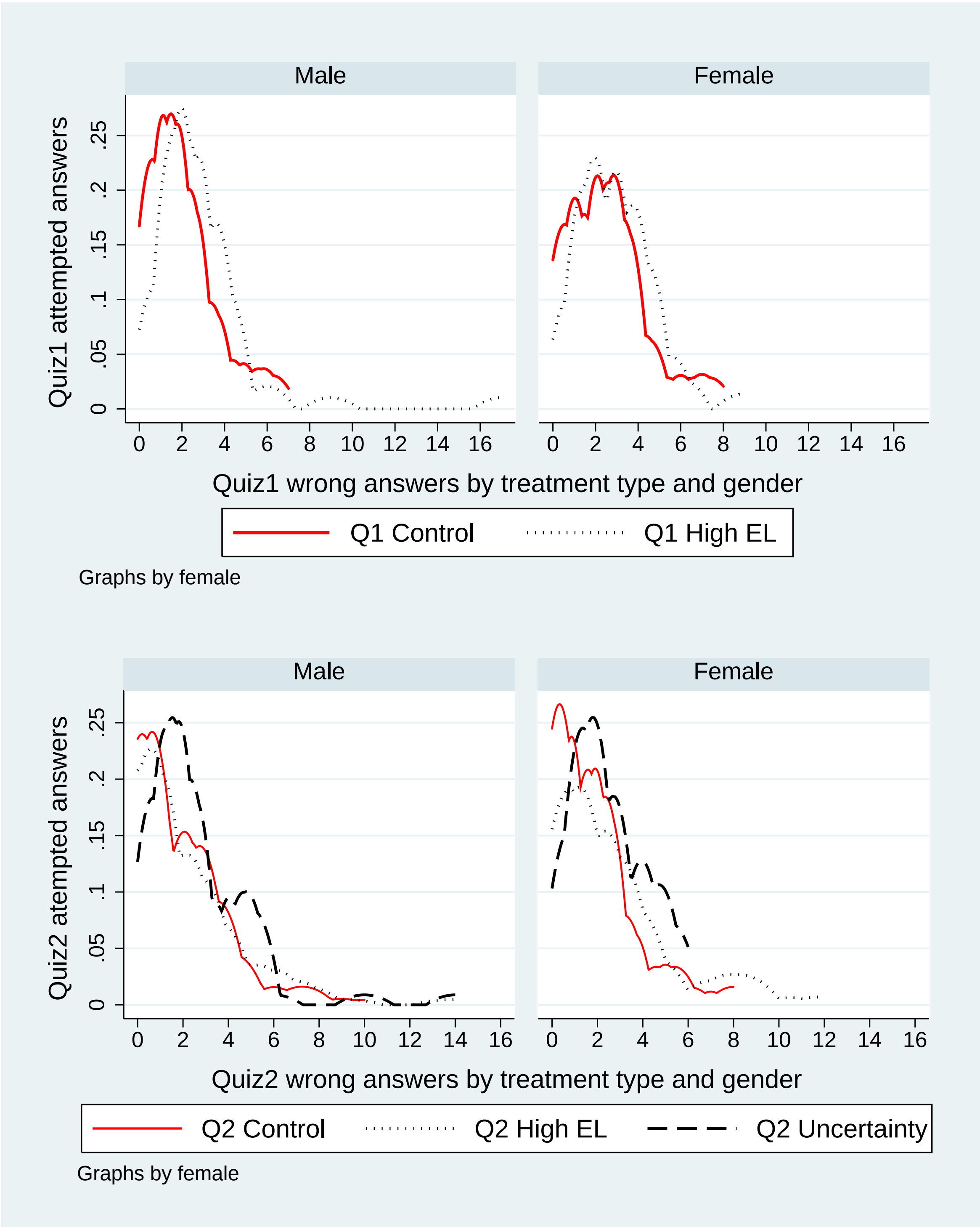

Figure 4: Quiz 1 wrong answers by treatment and gender

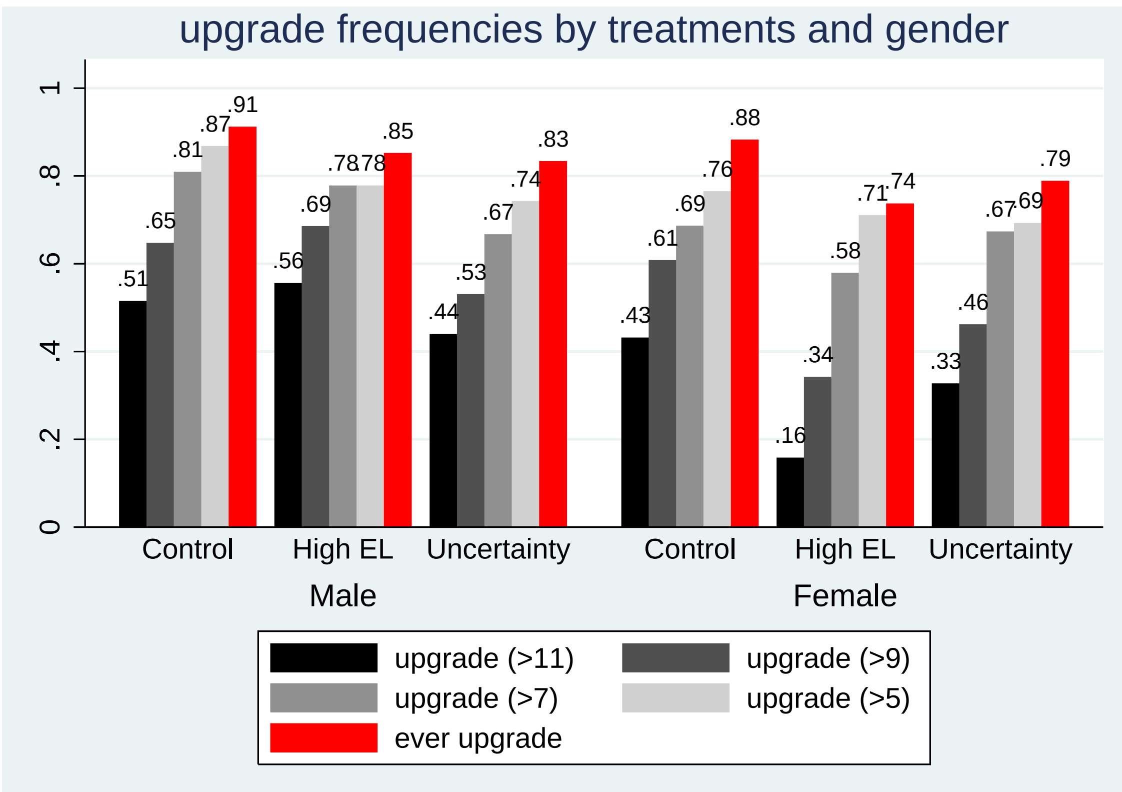

Figure 5: Entrepreneurial choice by treatment and gender

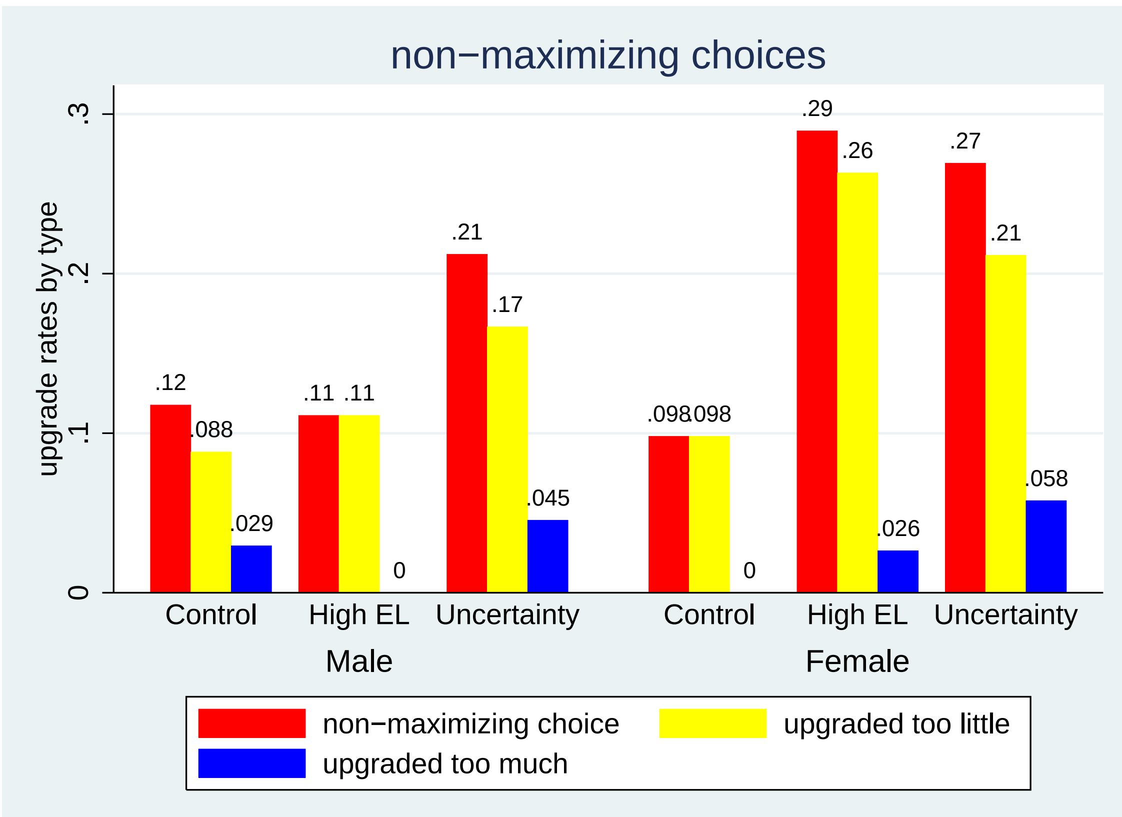

Figure 6: Non income maximizing entrepreneurial choice by treatment and gender

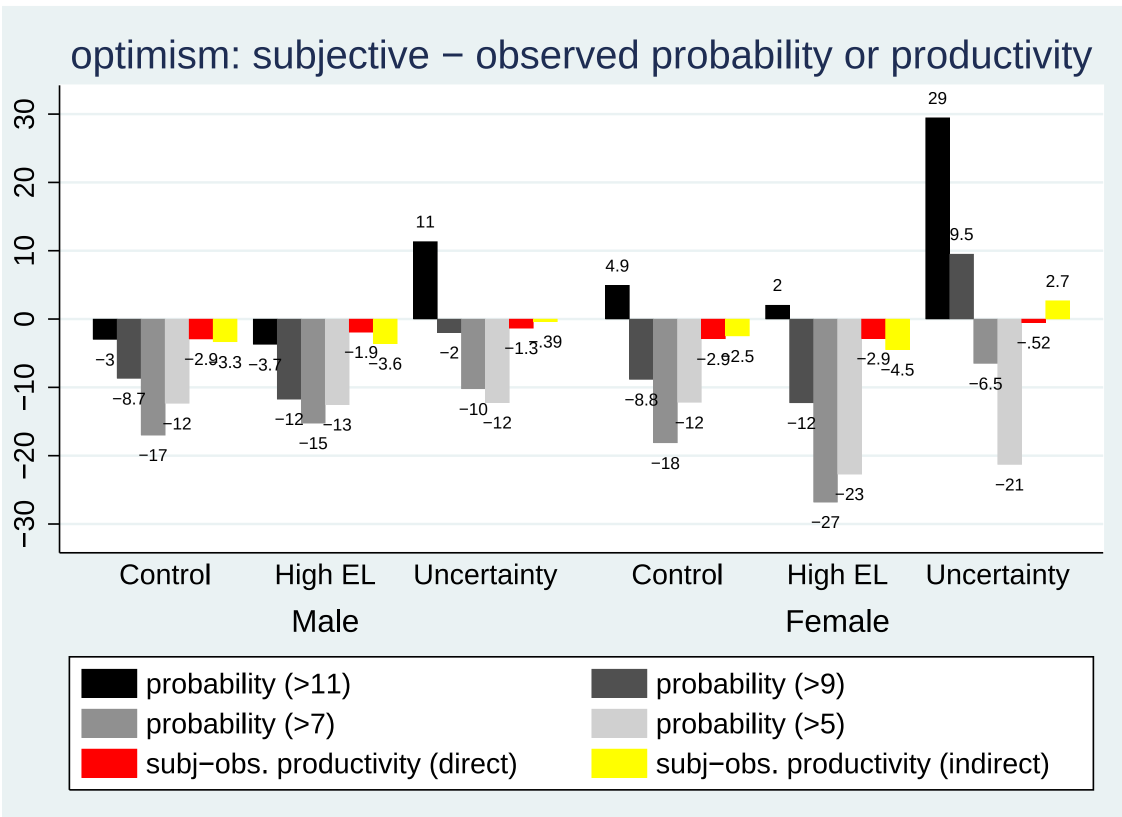

Figure 7: Subjective - observed probabilities and productivity by treatment and gender

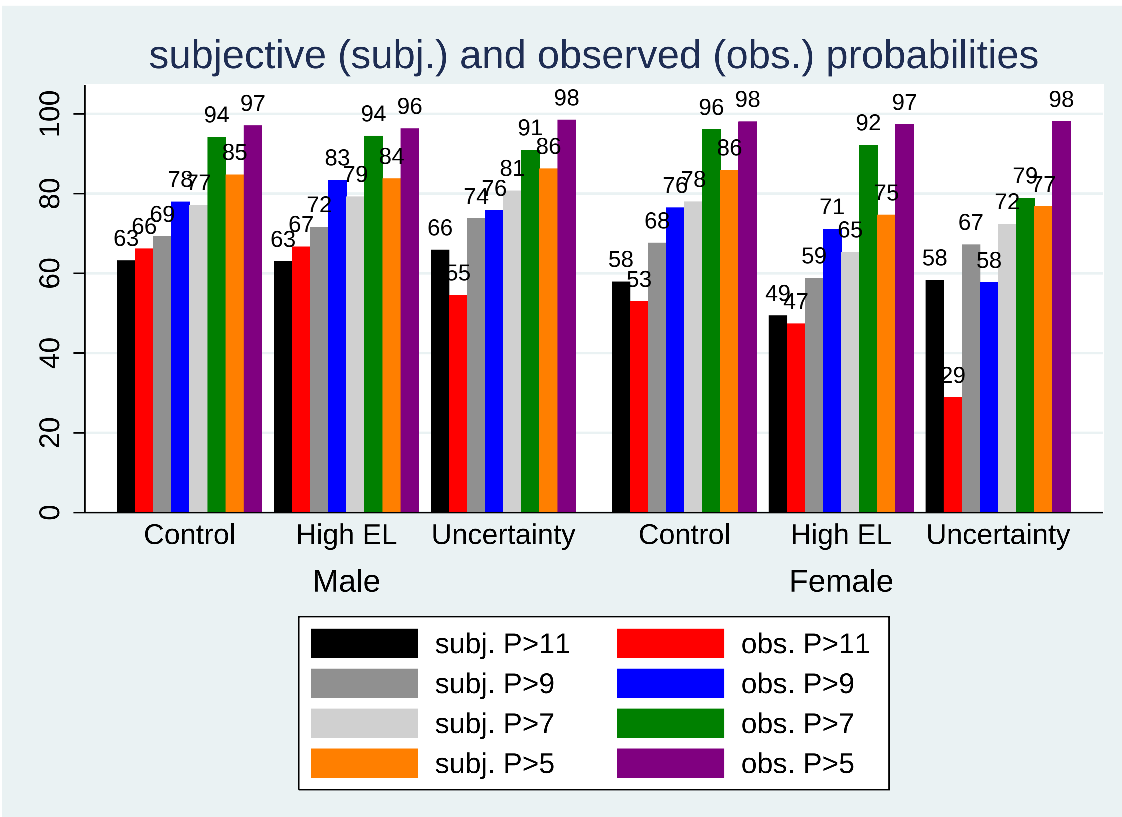

Figure 8: Subjective and observed probabilities by treatment and gender

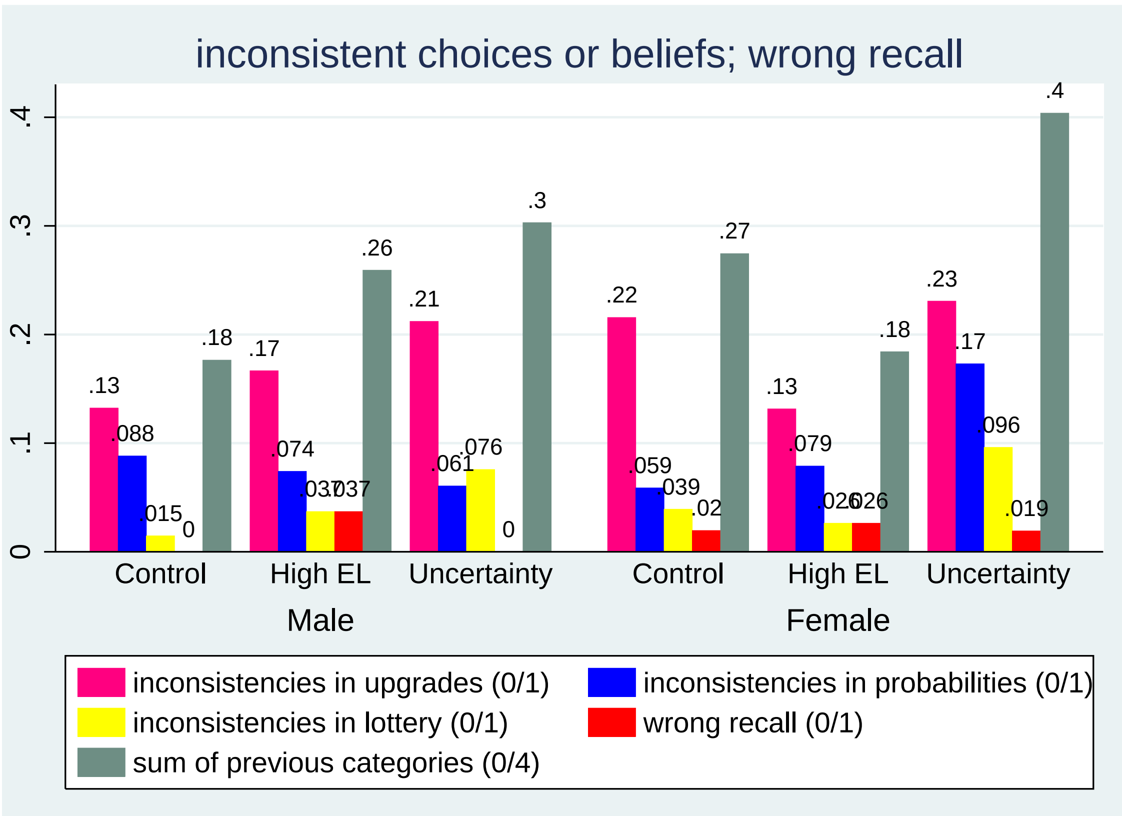

Figure 9: Inconsistencies choices, beliefs; Wrong recall

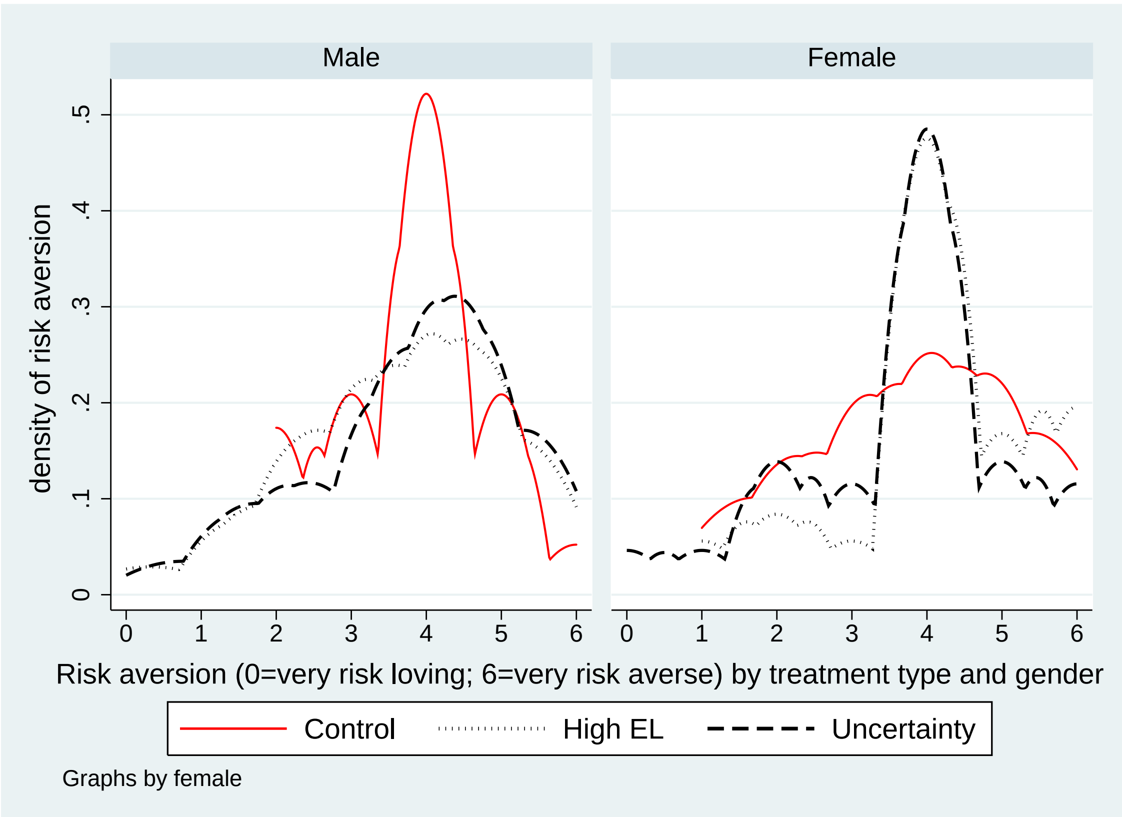

Figure 10: Density of risk attitudes by treatment and gender

Table 1: Experimental design and sample sizes

	U=0	U=1
EL=0	*83*	**52**
EL=1	**92**	**61**
CL=0 (robustness check)	*36*	**66**
CL=1 (robustness check)	37	74

Table 2: Experimental sessions by date and treatment

Type	Date	Pooled Treatments (T)	EL	U	Cortisol	Main data (M) or robustness check (R)
EL	Apr19 2010	T=0,1	EL=0,1	U=0		M
EL	Apr19 2010	T=0,1	EL=0,1	U=0		M
EL	Apr22 2010	T=0,1	EL=0,1	U=0		M
EL	Feb01 2011	T=0,1	EL=0,1	U=0	Y	M
EL	Feb02 2011	T=0,1	EL=0,1	U=0	Y	M
EL	Feb03 2011	T=0,1	EL=0,1	U=0	Y	M
EL	Apr26 2011	T=0,1	EL=0,1	U=1		M
EL	Apr28 2011	T=0,1	EL=0,1	U=1		M
EL	Feb08 2012	T=0,1	EL=0,1	U=1		M
EL	Feb09 2012	T=0,1	EL=0,1	U=1		M
EL	Apr04 2012	T=0,1	EL=0,1	U=1		M
EL	Apr05 2012	T=0,1	EL=0,1	U=1		M
CL	Apr20 2010	T=0	EL=0	U=0		M,R
CL	Apr21 2010	T=0	EL=0	U=0		M,R
CL	Apr27 2011	T=1	EL=0	U=1		M,R
CL	Apr27 2011	T=1	EL=0	U=1		M,R
CL	Feb08 2012	T=1	EL=0	U=1		M,R
CL	Feb09 2012	T=1	EL=0	U=1		M,R
CL	Apr04 2012	T=1	EL=0	U=1		M,R
CL	Apr05 2012	T=1	EL=0	U=1		M,R

Table 3: Difference in means, by treatment condition and gender.

Difference in means by treatment condition and gender										
	Means	**Regression Coefficients: All**			**Regression Coefficients: Men**			**Regression Coefficients: Women**		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		T	EL	U	T	EL	U	T	EL	U
Female	0.426	0.007	-0.006	0.051						
		[0.887]	[0.938]	[0.465]						
Age	20.649	0.016	0.007	0.022	0.013	-0.029	0.026	0.017	0.04	0.013
		[0.031]**	[0.808]	[0.137]	[0.178]	[0.421]	[0.011]**	[0.128]	[0.298]	[0.573]
Gpa	3.332	-0.009	0.013	-0.133	-0.066	-0.09	-0.108	0.008	0.037	-0.237
		[0.671]	[0.622]	[0.058]*	[0.324]	[0.307]	[0.234]	[0.644]	[0.007]***	[0.050]**
College yr: % sophomore	0.215	0.06	0.109	0.069	0.014	0.046	0.009	0.127	0.212	0.114
(freshman=base category)		[0.452]	[0.374]	[0.523]	[0.896]	[0.784]	[0.956]	[0.280]	[0.173]	[0.442]
College yr: % junior	0.336	0.038	0.108	0.055	-0.044	0.047	-0.025	0.134	0.207	0.116
		[0.628]	[0.388]	[0.588]	[0.667]	[0.771]	[0.853]	[0.265]	[0.218]	[0.497]
College yr: % senior	0.284	-0.01	0.096	-0.033	-0.026	0.216	-0.129	0.011	-0.049	0.028
		[0.902]	[0.520]	[0.770]	[0.811]	[0.270]	[0.396]	[0.930]	[0.800]	[0.867]
Father has college degree	0.712	-0.019	-0.012	-0.017	-0.001	-0.01	0.033	-0.027	0.031	-0.065
		[0.740]	[0.900]	[0.832]	[0.994]	[0.935]	[0.767]	[0.739]	[0.803]	[0.568]
Mother has college degree	0.623	-0.001	0.027	-0.05	-0.055	-0.09	-0.134	0.09	0.186	0.091
		[0.990]	[0.753]	[0.508]	[0.433]	[0.415]	[0.197]	[0.281]	[0.136]	[0.420]
Self-reported stress	15.267	-0.004	-0.005	-0.006	-0.013	-0.019	-0.02	0.008	0.01	0.014
		[0.348]	[0.356]	[0.234]	[0.009]***	[0.004]***	[0.002]***	[0.187]	[0.266]	[0.079]*
Parental divorce / death of	0.382	0.046	-0.003	0.099	0.024	-0.077	0.083	0.08	0.106	0.078
close person previous 5 yrs		[0.347]	[0.967]	[0.142]	[0.717]	[0.459]	[0.351]	[0.289]	[0.334]	[0.462]
Constant		0.423	0.223	0.532	0.867	1.583	0.656	0.035	-0.962	0.698
		[0.032]**	[0.711]	[0.170]	[0.008]***	[0.039]**	[0.098]*	[0.905]	[0.187]	[0.301]
Observations		390	211	237	224	122	134	166	89	103
F test of joint coefficient significance [pvalues]		[0.873]	[0.965]	[0.303]	[0.121]	[0.213]	[0.002]	[0.462]	[0.053]	[0.229]

Robust standard errors of the mean differences in parenthesis. ***,**,* = significant at the 1,5,10% level. Self-reported level of stress is the score from the Perceived Stress Scale 10-Item Questionnaire as in Sheldon *et al* (1983). The ranges of the total score are 0-40 points. The estimates in columns (1) to (9) are the coefficients from the treatment dummies in regressions of each outcome on the treatment.

Table 4: Effect of high emotional load on change in cortisol by gender

ATT of High Emotional Load on Change in Cortisol			
			H_0:w=m
	(Women)	(Men)	[p-value]
Effect of High Emotional Load (EL)			
Change in cortisol ATT	-0.016	**0.074**	**[0.015]**
	[0.023]	**[0.026]****	
H_0: F(EL)<F(C) [pvalues]	[0.495]	[0.962]	
H_0: F(EL)>F(C) [pvalues]	[0.957]	**[0.001]**	
N	39	53	
Mean Control group (C)	-0.030	-0.035	[0.743]

OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Kolmogorov-Smirnov test [pvalues] to establish first-order stochastic dominance. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 5: $Quiz_1$ and $Quiz_2$ correct answers by treatment and gender

ATT of High Emotional Load and Uncertainty on Correct Answers $Quiz_1$ and $Quiz_2$						
	(Women)	(Men)	H_0:w=m [p-value]	(Women)	(Men)	H_0:w=m [p-value]
		Quiz$_1$			***Quiz$_2$***	
			Effect of Pooled Treatments (T)			
Productivity ATT	**-1.08**	-0.78	[0.702]	**-1.18**	0.01	**[0.078]***
	[0.57]*	[0.55]		**[0.48]****	[0.47]	
H_0: F(T)>F(C) [pvalues]	[0.930]	[0.989]		[1.000]	[0.966]	
H_0: F(T)<F(C) [pvalues]	[0.413]	[0.129]		**[0.061]**	[0.400]	
H_0: SD(T)=SD(C) [pvalues]	[0.458]	[0.887]		[0.492]	[0.330]	
Positive ATT by: (1) baseline stress;	No	No		No	Yes	
(2) gpa, age, college year, father's education.	No	No		No	No	
N	166	224		166	224	
			Effect of High Emotional Load (EL)			
Productivity ATT	-1.12	-0.81	[0.757]	-0.98	-0.01	[0.262]
	[0.72]	[0.73]		[0.63]	[0.58]	
H_0: F(EL)>F(C) [pvalues]	[0.960]	[0.992]		[0.869]	[0.364]	
H_0: F(EL)<F(C) [pvalues]	[0.406]	**[0.097]**		[0.398]	[0.736]	
H_0: SD(T)=SD(C) [pvalues]	[0.599]	[0.735]		[0.833]	[0.424]	
Positive ATT by: (1) baseline stress;	No	No		No	No	
(2) gpa, age, college year, father's education.	No	No		No	No	
N	89	122		89	122	
			Effect of Uncertainty (U)			
Productivity ATT				**-1.37**	-0.23	[0.159]
				[0.57]**	[0.67]	
H_0: F(U)>F(C) [pvalues]				[1.000]	[0.986]	
H_0: F(U)<F(C) [pvalues]				**[0.050]**	[0.267]	
H_0: SD(T)=SD(C) [pvalues]	[0.757]	[0.915]		[0.374]	[0.268]	
Positive ATT by: (1) baseline stress;	No	No		No	No	
(2) gpa, age, college year, father's education.	No	No		No	No	
Positive ATT on changes (extensive margin)	n.a.	n.a.		No	No	
Positive ATT on changes (intensive margin)	n.a.	n.a.		No	No	
N				103	134	
H_0: ATT(EL)=ATT(U) [pvalues]				[0.535]	[0.711]	
Mean Control group (C)	8.55	9.06	[0.450]	11.94	12.79	[0.182]

OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Kolmogorov-Smirnov test [pvalues] to establish first-order stochastic dominance. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 6: $Quiz_1$ and $Quiz_2$ attempted and wrong answers by treatment and gender

ATT of High Emotional Load and Uncertainty on Attempted Answers $Quiz_1$ and $Quiz_2$						
	(Women)	(Men)	H_0:w=m [p-value]	(Women)	(Men)	H_0:w=m [p-value]
		$Quiz_1$			***$Quiz_2$***	
Effect of Pooled Treatments (T)						
Attempts at Quiz ATT	-0.87	-0.54	[0.691]	**-0.70**	0.04	[0.224]
	[0.61]	[0.56]		**[0.43]****	[0.44]	
Quiz wrong answers ATT	0.47	**0.39**	[0.841]	0.47	0.03	[0.332]
	[0.29]	**[0.24]***		[0.32]	[0.31]	
N	166	224		166	224	
Effect of High Emotional Load (EL)						
Attempts at Quiz ATT	-0.87	-0.17	[0.527]	-0.19	0.18	[0.618]
	[0.80]	[0.76]		[0.54]	[0.51]	
Quiz wrong answers ATT	0.63	**0.91**	[0.619]	0.78	0.19	[0.378]
	[0.40]	**[0.39]****		[0.51]	[0.43]	
N	89	122		89	122	
Effect of Uncertainty (U)						
Attempts at Quiz ATT				**-1.29**	-0.12	[0.123]
				[0.52]**	[0.55]	
Quiz wrong answers ATT				0.08	0.09	[0.972]
				[0.34]	[0.39]	
N				103	134	
H_0: ATT(EL)=ATT(U) [pvalues]						
Attempts				[0.160]	[0.846]	
Wrong answers				[0.160]	[0.846]	
Mean Control group (C)						
Attempts	10.85	11.53	[0.201]	14.43	15.45	**[0.102]**
Wrong answers	2.69	2.33	[0.168]	1.69	1.66	[0.910]

OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Additional controls: age, gender, gpa, college year, paternal education.

Table 7: Pooled upgrade likelihood by treatment and gender

ATT of High Emotional Load and Uncertainty on the Pooled Upgrade Likelihood			
	(Women)	(Men)	H_0:w=m [p-value]
Effect of Pooled Treatments (T)			
Upgrade ATT	**-0.104**	**-0.083**	[0.778]
	[0.057]***	[0.049]	
N	664	896	
Effect of High Emotional Load (EL)			
Upgrade ATT	**-0.186**	-0.032	**[0.109]**
	[0.072]**	[0.061]	
N	356	488	
Effect of Uncertainty (U)			
Upgrade ATT	-0.063	**-0.124**	[0.497]
	[0.069]	**[0.061]****	
N	412	536	
H_0: ATT(EL)=ATT(U)	[0.110]	[0.168]	
Mean Control group (C)	0.622	0.709	**[0.048]**

Pooled OLS of upgrades 1 to 4. Standard errors clustered by individual. 390 subjects. ***,**,* = significant at the 1,5,10% level. EL, C = high emotional load, control groups. Pooled treatments include subjects who received both EL and U. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 8: Non-income maximizing entrepreneurial choice by treatment and gender

ATT of High Emotional Load and Uncertainty on the Likelihood of Making Non-income Maximizing Choice (ex post) due to too little upgrade and failure			
			H_0:w=m
	(Women)	(Men)	[p-value]
Effect of Pooled Treatments (T)			
Non-income maximazing choice ATT	**0.178**	0.059	[0.126]
	[0.059]***	[0.050]	
Too little upgrade ATT	**0.137**	0.067	[0.335]
	[0.057]**	[0.045]	
Too much ATT	**0.041**	-0.007	[0.128]
	[0.018]**	[0.024]	
N	166	224	
Effect of High Emotional Load (EL)			
Non-income maximazing choice ATT	**0.194**	0.050	**[0.074]**
	[0.087]**	[0.058]	
Too little upgrade ATT	**0.165**	0.029	[0.185]
	[0.085]**	**[0.055]**	
Too much ATT	0.029	-0.024	[0.126]
	[0.028]	[0.019]	
N	89	122	
Effect of Uncertainty (U)			
Non-income maximazing choice ATT	**0.156**	0.100	[0.579]
	[0.076]*	[0.066]	
Too little upgrade ATT	0.104	0.089	[0.874]
	[0.072]	[0.060]	
Too much ATT	0.052	0.01	[0.402]
	[0.033]	[0.034]	
N	103	132	
Non-income maximazing choice ATT			
H_0: ATT(EL)=ATT(U)	[0.705]	[0.163]	
Mean Control group (C)	0.098	0.117	[0.734]
Too little upgrade ATT			
H_0: ATT(EL)=ATT(U)	[0.524]	[0.356]	
Mean Control group (C)	0.098	0.088	[0.857]
Too much upgrade ATT			
H_0: ATT(EL)=ATT(U)	[0.609]	[0.174]	
Mean Control group (C)	0	0.029	[0.157]

Note: Failure = Failing to have the minimal number of correct answers for payment>0 (unconditional). Therefore: Too much = 1 if # correct answers<required#. OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Kolmogorov-Smirnov test [pvalues] to establish first-order stochastic dominance. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 9: Losses and Returns on Quiz 2 (US Dlls)

ATT of High Emotional Load and Uncertainty on Losses and Returns on Quiz 2 (US Dlls)									
			H_0:w=m			H_0:w=m			H_0:w=m
	(Women)	(Men)	[p-value]	(Women)	(Men)	[p-value]	(Women)	(Men)	[p-value]
	Pooled Treatments (T)			**High Emotional Load (EL)**			**Uncertainty (U)**		
Return on Quiz 2 ($) ATT	**-2.78**	-0.55	**[0.067]**	**-2.41**	-0.27	[0.17]	**-3.05**	-1.14	[0.21]
	[0.86]***	[0.87]		**[1.18]****	[1.03]		**[1.08]*****	[1.09]	
Of which, loss from	**-1.02**	**-0.59**	[0.300]	**-0.92**	-0.21	[0.290]	**-1.03**	**-0.85**	[0.194]
wrong upgrade,	**[0.33]*****	**[0.29]****		**[0.43]****	[0.33]		**[0.48]****	**[0.41]****	
the sum of:									
(1) loss from	**-0.63**	**-0.45**	[0.627]	**-0.73**	-0.31	[0.426]	-0.46	**-0.54**	[0.866]
too little upgrade	**[0.28]****	**[0.24]***		**[0.41]***	[0.31]		[0.34]	**[0.31]***	
and									
(2) loss from	**-0.39**	-0.13	[0.283]	-0.18	0.11	[0.116]	-0.57	-0.31	[0.775]
too much upgrade	**[0.19]****	[0.17]		[0.15]	[0.09]		[0.37]	[0.30]	
N	166	224		89	122		103	134	
Mean Control group									
Return	17.39	18.58	[0.277]	17.39	18.58	[0.277]	17.39	18.58	[0.277]
Loss (wrong upgrade)	-0.47	-0.54	[0.790]	-0.47	-0.54	[0.790]	-0.47	-0.54	[0.790]
Loss (too litle upgrade)	-0.47	-0.42	[0.868]						
Loss (too much upg.)	0.00	-0.12	[0.170]	0.00	-0.12	[0.170]	0.00	-0.12	[0.170]

OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 10: Return on quiz2 and losses from wrong upgrades. Heterogeneity by life events and gender

ATT of High Emotional Load and Uncertainty on the Return on Quiz 2 and Losses, Heterogenous Effect by Life Events and Gender						
	(Women)	(Men)	H_0:w=m [p-value]	(Women)	(Men)	H_0:w=m [p-value]
	No Negative Life Events			***Divorce, Death, Low Parental Education***		
	Effect of Pooled Treatments (T)					
Return on	-0.281	0.188	[0.769]	**-4.602**	-1.105	**[0.052]**
Quiz 2 ($) ATT	[1.126]	[1.139]		**[1.191]*****	[1.328]	
Loss from suboptimal	-0.222	**-0.852**	[0.327]	**-1.635**	-0.258	**[0.020]**
upgrade ATT	[0.524]	**[0.375]****		**[0.408]*****	[0.432]	
N	68	107		98	117	
	Effect of High Emotional Load (EL)					
Return on	-0.098	0.822	[0.665]	**-4.306**	-1.15	[0.169]
Quiz 2 ($) ATT	[1.616]	[1.390]		**[1.642]*****	[1.572]	
Loss from suboptimal	-0.679	-0.721	[0.964]	**-1.133**	0.49	**[0.013]**
upgrade ATT	[0.737]	[0.505]		**[0.439]****	[0.406]	
N	40	64		49	58	
	Effect of Uncertainty (U)					
Return on	0.068	-1.263	[0.519]	**-5.065**	-1.017	**[0.058]**
Quiz 2 ($) ATT	[1.459]	[1.457]		**[1.431]*****	[1.577]	
Loss from suboptimal	0.606	**-1.062**	**[0.038]**	**-2.013**	-0.596	**[0.095]**
upgrade ATT	[0.514]	**[0.628]***		**[0.630]*****	[0.585]	
N	39	61		64	73	
Mean T=0 ($):						
Return	15.52	18.2	**[0.055]**	18.81	19	[0.908]
Upgrade loss	-0.909	-0.357	[0.245]	-0.137	-0.742	**[0.084]**
H_0: ATT(EL)=ATT(U)						
Return	[0.927]	[0.176]		[0.663]	[0.931]	
Upgrade loss	**[0.060]**	[0.647]		[0.248]	**[0.056]**	

Note: We obtain similar results when we consider life events and low paternal education separately. OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 11: Return on quiz2 and losses from wrong upgrades. Heterogeneity by baseline self-reported stress and gender

ATT of High Emotional Load and Uncertainty on the Return on Quiz 2 and Losses, Heterogenous Effect by Life Events and Gender

	(Women)	(Men)	H_0:w=m [p-value]	(Women)	(Men)	H_0:w=m [p-value]
	Below median stress			***Above median stress***		
	Effect of Pooled Treatments (T)					
Return on	-3.26	-2.507	[0.646]	-2.803	1.29	[0.016]
Quiz 2 ($) ATT	[1.310]**	[1.027]**		[1.140]**	[1.266]	
Loss from suboptimal	-1.128	-1.157	[0.959]	-0.975	-0.046	[0.133]
upgrade ATT	[0.442]***	[0.331]***		[0.454]**	[0.437]	
N	56	116		108	107	
	Effect of high Emotional Load video clip (EL)					
Return on	-2.735	-1.612	[0.627]	-2.002	1.151	[0.142]
Quiz 2 ($) ATT	[2.004]	[1.211]		[1.459]	[0.1606]	
Loss from suboptimal	-1.802	-0.634	[0.143]	-0.326	0.213	[0.476]
upgrade ATT	[0.726]**	[0.339]*		[0.526]	[0.555]	
N	33	60		54	61	
	Effect of uncertainty about task 1 productivity (U)					
Return on	-4.637	-3.078	[0.407]	-2.159	1.005	[0.158]
Quiz 2 ($) ATT	[1.458]***	[1.235]**		[1.389]	[1.735]	
Loss from suboptimal	-1.107	-1.445	[0.678]	-0.955	-0.192	[0.368]
upgrade ATT	[0.602]*	[0.545]***		[0.647]	[0.559]	
N	34	72		68	61	
Mean T=0 ($):						
Return	18.157	21.051	[0.063]	17.048	16.447	[0.686]
Upgrade loss	-0.236	0	[0.320]	-0.629	-0.973	[0.434]
H0: ATT(EL)=ATT(U)						
Return	[0.368]	[0.246]		[0.920]	[0.939]	
Upgrade loss	[0.426]	[0.189]		[0.390]	[0.532]	

Note: We obtain similar results when we consider life events and low paternal education separately. OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 12: Inconsistencies and mistakes by treatment and gender: Optimism for Highest Upgrade

ATT of High Emotional Load and Uncertainty on Optimism for Highest Upgrade			
	(Women)	(Men)	H_0:w=m [p-value]
	Effect of Pooled Treatments (T)		
Optimism *(x≥12)* ATT	**15.41**	0.19	[0.152]
	[8.16]*	[6.77]	
H_0: F(T)<F(C) [pvalues]	[0.782]	[0.788]	
H_0: F(T)>F(C) [pvalues]	**[0.017]**	[0.484]	
N	166	224	
	Effect of High Emotional Load (EL)		
Optimism *(x≥12)* ATT	0.14	-1.95	[0.873]
	[9.86]	[8.48	
H_0: F(EL)<F(C) [pvalues]	[0.187]	[0.739]	
H_0: F(EL)>F(C) [pvalues]	[0.781]	[0.848]	
N	89	122	
	Effect of Uncertainty (U)		
Optimism *(x≥12)* ATT	**20.08**	5.9	[0.271]
	[9.94]**	[8.36]	
H_0: F(U)<F(C) [pvalues]	[0.928]	[0.981]	
H_0: F(U)>F(C) [pvalues]	**[0.005]**	[0.144]	
N	103	134	
H_0: ATT(EL)=ATT(U)	**[0.052]**	[0.389]	
Mean Control group (C)	4.94	-2.98	[0.381]

Note: Optimism for highest upgrade (x≥12): stated - observed probability. OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Kolmogorov-Smirnov test [pvalues] to establish first-order stochastic dominance. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 13: Inconsistencies and mistakes by treatment and gender: Optimism for Lowest Upgrade

ATT of High Emotional Load and Uncertainty on Optimism for Lowest Upgrade			
	(Women)	(Men)	H_0:w=m [p-value]
Effect of Pooled Treatments (T)			
Optimism *(x≥6)* ATT	**-10.57**	-2.76	[0.191]
	[4.49]**	[3.73]	
H_0: F(T)>F(C) [pvalues]	[0.999]	[0.876]	
H_0: F(T)<F(C) [pvalues]	**[0.004]**	[0.515]	
N	166	224	
Effect of High Emotional Load (EL)			
Optimism *(x≥6)* ATT	**-11.15**	-0.44	[0.181]
	[6.55]*	[4.24]	
H_0: F(EL)>F(C) [pvalues]	[0.998]	[0.944]	
H_0: F(EL)<F(C) [pvalues]	**[0.004]**	[0.958]	
N	89	122	
Effect of Uncertainty (U)			
Optimism *(x≥6)* ATT	**-9.29**	-1.91	[0.284]
	[5.32]*	[4.36]	
H_0: F(U)>F(C) [pvalues]	[1.000]	[0.544]	
H_0: F(U)<F(C) [pvalues]	**[0.052]**	[0.720]	
N	103	134	
H_0: ATT(EL)=ATT(U)	[0.784]	[0.701]	
Mean Control group (C)	-12.19	-12.33	[0.976]

Note: Optimism for lowest upgrade (x≥6): stated - observed probability. OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Kolmogorov-Smirnov test [pvalues] to establish first-order stochastic dominance. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 14: Inconsistencies and mistakes by treatment and gender: Differences in subjective and observed productivity

ATT of High Emotional Load and Uncertainty on Subjective - Observed Productivity						
			H_0:w=m			H_0:w=m
	(Women)	(Men)	[p-value]	(Women)	(Men)	[p-value]
Subjective - Observed Productivity						
	Direct			**Indirect and Truncated**		
Effect of Pooled Treatments (T)						
Subj - Obs Productivity ATT	1.224	0.262	[0.371]	2.47	-0.351	[0.255]
	[0.901]	[0.704]		[1.866]	[1.562]	
H_0: F(T)>F(C) [pvalues]	[0.992]	[0.883]		[0.997]	[0.997]	
H_0: F(T)<F(C) [pvalues]	[0.189]	[0.637]		[0.167]	**[0.059]**	
N	117	155		166	224	
Effect of High Emotional Load (EL)						
Subj - Obs Productivity ATT	0.53	0.504	[0.987]	-1.118	-0.723	[0.896]
	[0.119]	[1.098]		[2.248]	[1.957]	
H_0: F(EL)>F(C) [pvalues]	**[0.082]**	[0.739]		[0.941]	[0.889]	
H_0: F(EL)<F(C) [pvalues]	[0.942]	[0.882]		[0.711]	[0.192]	
N	54	40		89	122	
Effect of Uncertainty (U)						
Subj - Obs Productivity ATT	**1.669**	0.129	[0.283]	**4.28**	0.86	[0.263]
	[0.999]*	[0.804]		**[2.351]***	[1.928]	
H_0: F(U)>F(C) [pvalues]	[0.982]	[0.864]		[1.000]	[0.997]	
H_0: F(U)<F(C) [pvalues]	**[0.075]**	[0.263]		**[0.080]**	**[0.098]**	
N	92	70		103	134	
H_0: ATT(EL)=ATT(U)	[0.254]	[0.731]		**[0.026]**	[0.436]	
Mean Control group (C)	-2.888	-2.923	[0.976]	-2.472	-3.326	[0.678]

OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. Kolmogorov-Smirnov test [pvalues] to establish first-order stochastic dominance. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 15: Inconsistencies and mistakes by treatment and gender: Confused Choices and Wrong Recollection

ATT of High Emotional Load and Uncertainty on Likelihood of Making Confused Choices and Wrong Recall			
	(Women)	(Men)	H_0:w=m [p-value]
Inconsistent Choices/Beliefs and Question Recall			
Effect of Pooled Treatments (T)			
Inconsistencies ATT	-0.037	**0.096**	[0.148]
	[0.073]	**[0.056]***	
Wrong Recall ATT	-0.003	**0.026**	[0.320]
	[0.023]	**[0.013]***	
N	166	224	
Effect of High Emotional Load (EL)			
Inconsistencies ATT	-0.109	0.06	[0.128]
	[0.087]	[0.069]	
Wrong Recall ATT	0.006	0.04	[0.495]
	[0.037]	[0.029]	
N	89	122	
Effect of Uncertainty (U)			
Inconsistencies ATT	0.019	**0.162**	[0.210]
	[0.087]	**[0.029]****	
Wrong Recall ATT	-0.002	-0.003	[0.991]
	[0.025]	[0.007]	
N	103	134	
Inconsistencies:			
H_0: ATT(EL)=ATT(U)	[0.138]	[0.183]	
Mean Control group (C)	0.254	0.147	[0.155]
Wrong Recall:			
H_0: ATT(EL)=ATT(U)	[0.789]	[0.158]	
Mean Control group (C)	0.027	0	[0.156]

Note: This regression uses the dummy variable *confused*, which =1 if person makes inconsistent upgrade choices, or states inconsistent probabilities, or switches back and forth in lottery choices. OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.

Table 16: Risk attitudes by treatment and gender

ATT of High Emotional Load and Uncertainty on Risk Aversion						
			H_0:w=m			H_0:w=m
	(Women)	(Men)	[p-value]	(Women)	(Men)	[p-value]
	Risk aversion: risk loving			***Risk aversion: risk neutral***		
Effect of Pooled Treatments (T)						
Risk Aversion ATT	-0.009	**0.053**	[0.140]	-0.100	-0.077	[0.817]
	[0.037]	**[0.020]*****		[0.075]	[0.067]	
N	166	224		166	224	
Effect of High Emotional Load (EL)						
Risk Aversion ATT	-0.01	**0.056**	[0.190]	**-0.156**	0.020	[0.163]
	[0.041]	**[0.033]***		**[0.090]***	[0.088]	
Effect of Uncertainty (U)						
Risk Aversion ATT	0.014	**0.054**	[0.498]	-0.075	-0.101	[0.828]
	[0.047]	**[0.032]***		[0.089]	[0.077]	
Mean T=0	0.040	0.000	**[0.080]**	0.200	0.201	[0.976]
N	141	188		141	188	
H_0: ATT(EL)=ATT(U)	[0.569]	[0.975]		[0.352]	[0.148]	
	Risk aversion: risk averse			***Risk aversion: very risk averse***		
Effect of Pooled Treatments (T)						
Risk Aversion ATT	**0.155**	**-0.129**	**[0.009]**	-0.065	**0.107**	[0.089]
	[0.080]*	**[0.072]***		[0.079]	**[0.065]***	
N	166	224		166	224	
Effect of High Emotional Load (EL)						
Risk Aversion ATT	**0.204**	**-0.171**	**[0.008]**	-0.017	0.055	[0.584]
	[0.109]*	**[0.087]****		[0.103]	[0.081]	
Effect of Uncertainty (U)						
Risk Aversion ATT	0.126	-0.121	**[0.055]**	-0.119	0.079	**[0.095]**
	[0.095]	[0.087]		[0.089]	[0.079]	
Mean T=0	0.186	0.275	[0.158]	0.226	0.137	[0.132]
N	141	188		141	188	
H_0: ATT(EL)=ATT(U)	[0.496]	[0.579]		[0.301]	[0.779]	

Note: Coefficient of risk aversion from Holt-Laury lotteries. OLS. Robust standard errors. ***,**,* = significant at the 1,5,10% level. EL, C = high emotional load, control groups. Additional controls: age, gender, gpa, college year, baseline stress, parental divorce or death of close person in the previous 5 years, paternal college, dummies for screen and timer problems.