I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

Gender Differences in Reservation Wages in Search Experiments

Andrew McGee
Peter McGee

DISCUSSION PAPER SERIES

Gender Differences in Reservation Wages in Search Experiments

Andrew McGee
University of Alberta and IZA

Peter McGee
University of Arkansas

NOVEMBER 2023

ABSTRACT

Gender Differences in Reservation Wages in Search Experiments*

Women report setting lower reservation wages than men in survey data. We show that women set reservation wages that are 14 to 18 percent lower than men's in laboratory search experiments that control for factors not fully observed in surveys such as offer distributions and outside options. This gender gap—which exists even controlling for overconfidence, preferences, personality, and intelligence—leads women to spend less time searching than men while accepting lower wages. Women—but not men—set reservation wages that are too low relative to theoretically optimal values given their risk preferences early in search, reducing their earnings.

Corresponding author:
Andrew McGee
Department of Economics
University of Alberta
8-14 Tory Building
Edmonton
AB T6G 2H4
Canada

E-mail: mcgee1@ualberta.ca

1. Introduction

Studies find that women set significantly lower reservation wages than men when searching for jobs in Britain (Brown et al., 2011), Germany (Caliendo et al., 2017), Poland (Cukrowska-Torzewska, 2021), France (Le Barbachon et al., 2021), and several developing countries (Khan and Majid, 2020). Gender gaps in reservation wages are a potentially important source of wage gaps: Caliendo et al. (2017) find that the gender gap in reemployment wages essentially disappears after accounting for reservation wages.

This evidence, however, has important limitations. First, self-reported reservation wages in surveys are non-binding and possibly "cheap talk."[1] Second, the factors influencing both self-reported reservation wages and wages (e.g., family considerations, the wage distributions faced by men and women, expectations concerning discrimination, and productivity) are rarely fully observed. Finally, survey evidence that women set lower reservation wages than men cannot answer potentially the most important question: are women's lower reservation wages sub-optimal? In reality, women's reservation wages may represent a rational response to the economic conditions they face.

We investigate gender differences in reservation wages using data from two finite-horizon laboratory search experiments. In both experiments, subjects engaged in search episodes lasting 21 periods. In every period prior to accepting an offer, subjects set binding reservation wages before learning whether they received offers from a known wage distribution. If an offer greater than or equal to the reservation wage was received, the offer was automatically accepted

---

[1] Unlike the other studies above, Le Barbachon et al. (2021) use administrative data in which job seekers indicate their reservation wages when registering as unemployed. Their responses influenced the job postings sent to them by case workers and were thus less likely to be "cheap talk."

and the search ended. The subject received that amount for all remaining periods. If no offer was received or the offer was below the reservation wage, the process repeated in the next period.

The experiments differed primarily in the processes determining whether offers were received and the uncertainties faced by subjects. In the *Real Effort* experiment, subjects completed coding tasks each period that influenced the probability of receiving an offer in one of two treatments. In the *Uncertainty* treatment, subjects knew only that each successfully completed task "may influence" the likelihood that an offer would be received; in fact, each correctly completed task increased the likelihood of receiving an offer by four percentage points. In the *Certainty* treatment, subjects knew that correctly completed tasks increased the likelihood of receiving an offer by four percentage points. Subjects in the *Real Effort* experiment completed a single search episode.

Subjects in the *No Effort* experiment, which replicated the *Probability* treatment from Cox and Oaxaca (1989, 1992), completed ten search episodes. In the search episodes, offers were simply received in each period with a known probability of 0.25. Prior to learning whether offers were received, subjects indicated binding reservation wages in all periods before an offer was accepted. Subjects did not expend real effort. Eliminating the real effort component of search facilitates identifying the optimal path of reservation wages while also removing any channel through which overconfidence—a potential contributing factor to the gender gap in reservation wages investigated in Cortés et al. (2023)—might influence reservation wages.[2] In both experiments, a large number of subject characteristics such as cognitive ability and risk preferences were measured.

---

[2] Krueger and Mueller (2016) found that reservation wages were too high and declined too slowly within search spells among unemployed job seekers relative to a calibrated search model and speculated that this might result from job seekers being persistently overconfident about their prospects.

In the experiments, men and women searched from the same, known offer distributions, outside options were the same, and women had no reason to anticipate discrimination given the clearly described processes governing the experiments. That is, our experiments controlled for many of the factors not observed in survey data. Nevertheless, we find that women set binding reservation wages that are 14 to 18 percent less than men on average in both experiments. Moreover, women set reservation wages that are as much as 30 log-points lower than men's after controlling for personality traits, risk preferences, patience, overconfidence, cognitive ability and understanding of probability. As a consequence of this gender gap, women spend less time searching but also accept lower wages.

The optimal risk-neutral path of reservation wages is known in the *No Effort* experiment. Non-risk averse women who should set reservation wages *at or above* this risk-neutral benchmark set reservation wages that are significantly below the optimal reservation wages at the outset of search episodes. This leads women to accept offers that are too low early in the experiment, which has a profound effect on their earnings. By contrast, we fail to reject the hypothesis that men set reservation wages consistent with the optimal reservation wage path early in search episodes given their risk preferences.

Our study makes three primary contributions. First, we demonstrate that the gender gap in reservation wages found in field research exists even in settings in which the many confounders in field research are controlled and even after controlling for a very large vector of individual characteristics. Second, the gender gap in reservation wages is observed across different search environments as we remove elements of uncertainty and complexity in the offer generating process moving from the *Uncertainty* treatment in the *Real Effort* experiment to the *Certainty* treatment to the *No Effort* experiment. Doing so establishes that neither uncertainty

about the role of effort in search nor overconfidence in one's search effort can account for the observed gender gap. Finally, our most important and novel contribution is to demonstrate that early in search episodes women in the *No Effort* experiment set reservation wages that are too low given their risk preferences. Demonstrating that a gender gap in reservation wages exists when the characteristics of the search environment are controlled is not sufficient to conclude that women should set higher reservation wages as men might be setting sub-optimally high reservation wages. Our finding that women's reservation wages are sub-optimally low early in search, however, implies that active labor market policies encouraging women to hold out for higher wage offers when seeking employment have the potential to be welfare-enhancing.

2. Related Literature

A small number of studies use laboratory experiments to investigate whether individuals adhere to optimal reservation wage policies. Schotter and Braustein (1981) find that reservation wages decrease over time in an infinite horizon treatment, which is inconsistent with the theoretically optimal constant reservation wage. Brown et al. (2011) provide evidence that this is due to non-stationary, subjective search costs (essentially distaste for waiting in the lab). More consistent with optimal search behavior, Cox and Oaxaca (1989, 1992) find that reservation wages decline over search episodes and with risk aversion in finite horizon experiments. These studies do not consider whether search behaviors differ by gender. By contrast, we demonstrate that women (but not men) set sub-optimally low reservation wages given their risk preferences.

Outside the laboratory, a number of studies investigate gender differences in job search. Keith and McWilliams (1999) find that women in the National Longitudinal Survey of Youth 1979 were less likely to engage in search while employed even though the increases in earnings from such moves tend to be high. Likewise, Eriksson and Lagerström (2012) find that women in

Sweden were less likely to search in areas geographically distant from their current location even though the wage gains from search are increasing in the search area. Among unemployed French job seekers, Le Barbachon et al. (2021) find that gender gaps in reservation wages and maximum acceptable commutes are largest among married individuals with children. Fluchtman et al. (2023) show using Danish UI data that women were more likely than men to apply for part-time jobs. These studies attribute the search decisions made by women resulting in lower earnings to domestic responsibilities and gender task specialization. By contrast, we find that women are willing to accept lower wages even when domestic responsibilities do not affect one's ability to earn—suggesting that deep-rooted norms may contribute to gender differences in job search.

Studies have also documented that women are less likely to negotiate and are often treated less generously than men in negotiations (e.g., Babcock and Laschever, 2003; Busse et al., 2017). The consequences of gender disparities in negotiation behavior are non-trivial. Sin et al. (2022) find that one-third of a 13 to 17 percent gender wage gap remaining after controlling for worker characteristics and sorting can be attributed to differences in negotiation behavior. Similarly, Card et al. (2016) find that approximately one-fifth of the gender wage gap in Portugal can be attributed to gender differences in sorting and negotiation behavior. The binding reservation wages in our experiment are analogous to a series of one-off, take-it-or-leave-it negotiations. Women's lower propensity to negotiate may result from expectations about the behavior of women and the backlash associated with behavior at odds with these expectations (e.g., Bowles et al., 2007; Andersen et al., 2015). Even in the absence of a role for gender expectations, potential backlash, and selection into "negotiations," however, women in our experiments set lower acceptable amounts than men. Whether women should negotiate more frequently (i.e., the extensive margin) is unclear (Exley et al., 2020), but our experiment suggests

5

that women who are "negotiating" ask for too little (i.e., the intensive margin), which is often not possible in the negotiations literature due to the strategic considerations involved.

The only other laboratory experiment to examine gender differences in reservation wages of which we are aware is Cortés et al. (2023). In an exit survey for undergraduate students, Cortés et al. find that women accept jobs earlier than men and that the gender earnings gap in accepted wages shrinks as search episodes persist. In a laboratory experiment, Cortés et al. test a search model in which differences in risk preferences and overconfidence drive gender differences in reservation wages. Subjects engaged in a real effort search task but always received an offer from one of two distributions. Better performing subjects received offers from the distribution with higher likelihoods of offers from the right tail of the distribution. Controlling for risk preferences and overconfidence explains approximately a third of the 15 percent gender gap in initial reservation wages in their experiment.

Our experiments make three important contributions relative to Cortés et al. (2023). First, Cortés et al. note that psychological traits contribute to gender differences in labor market outcomes, but they control only for overconfidence and risk preferences. We collect a broad set of individual difference measures and find that personality traits and cognitive ability can each account for as much as a quarter of the gender gap in reservation wages in the *Real Effort* experiment. By contrast, risk preferences and overconfidence have very little effect on the estimated gender gaps in reservation wages in either of our experiments.

Second, Cortés et al.'s model generates comparative statics for risk preferences and overconfidence but cannot map out the optimal reservation wage path because reservation wages depend on unobserved factors such as beliefs, the speed of learning, and effort costs. By contrast,

we can show in our *No Effort* experiment that women's reservation wages are below the optimal reservation wage path given their risk preferences early in search episodes.

Finally, Moore and Healy (2008) distinguish between three types of overconfidence: overplacement, overestimation, and overprecision. Gender gaps have been found for all three types, and each may influence job search.[3] Job seekers may overestimate the likelihood that they will receive job offers given their search efforts, but they may also overrate their desirability to employers relative to the competition. Cortés et al.'s notion of overconfidence is overplacement (i.e., beliefs about where one falls in the ability distribution) because in their design relative performance influences the expected value of offers. By contrast, we measure overestimation (i.e., beliefs about one's performance relative to one's actual performance) because in our *Real Effort* experiment absolute performance influences the extensive margin of receiving an offer. Controlling for overestimation explains at most four percent of the gender difference in reservation wages in our experiments. Our findings together with those of Cortés et al. suggest that overconfidence—whatever its form—cannot account for most of the gender gaps in reservation wages in the experiments.

3. Experimental Details

The *Real Effort* experiment in which subjects completed a single 21-period search episode proceeded as follows.[4] In every period prior to accepting an offer, subjects had 90 seconds to complete coding tasks that influenced the probability of receiving an offer.[5] The instructions

---

[3] For evidence of gender differences in overplacement, overestimation, and overprecision, see Haeckl (2022), Bordalo et al. (2019) and Barber and Odean (2001), respectively.

[4] The *Real Effort* data were originally analyzed in McGee and McGee (2016), who tested hypotheses about the influence of locus of control beliefs on task completion and reservation wages and the moderating effects of uncertainty about the return to effort.

[5] Subjects were given a key matching ten words to ten four-digit numbers. The key was followed by seven questions, each listing a word and five four-digit numbers with the correct answer being the number corresponding to the word in the key. Subjects moved to additional keys and questions after they completed a set of questions.

informed subjects in the *Uncertainty* treatment that successfully completed tasks "may influence" whether an offer would be received. In fact, each correctly completed task increased the likelihood of receiving an offer by four percentage points in both treatments. Subjects in the *Certainty* treatment were informed of this relationship. After completing the tasks, subjects indicated the number of tasks that they believed they correctly completed and indicated a binding reservation wage.[6] Subjects then learned their offer if one was received and whether the offer was accepted. This process repeated each period until an offer had been accepted or until the 20th period concluded without an accepted offer.

Conditional on receiving an offer, wage offers of 5, 10, 15, 20, 35, 40, and 45 experimental currency units (ECUs) were received with probabilities 0.25, 0.25, 0.25, 0.10, 0.05, 0.05, and 0.05, respectively, which was common knowledge. Subjects earned a subsidy of 15 ECUs in every period before accepting an offer, but there was no subsidy in the 21st period.

Upon completion of the search episode, subjects completed the incentivized Holt and Laury (2002) risk preferences measure consisting of paired lotteries, an unincentivized risk preference self-assessment (Falk et al., 2018 and 2023), and measures of the Big Five personality traits (John et al., 2008), optimism (Scheier et al., 1994), self-esteem (Rosenberg, 1965), locus of control (Rotter, 1966), and cognitive ability. The cognitive ability measure consisted of the three-question Cognitive Reflection Test (CRT) (Frederick, 2005) together with three questions from the Wonderlic test (McKelvie, 1989). Subjects were paid their earnings from the search episode, for the outcome of one randomly selected lottery on the paired lottery risk preference measure, and $0.50 for each correct answer on the cognitive ability test. In total, 347 subjects—189 in the

---

[6] Similar to Cortés et al. (2023), we did not incentivize the elicitation of beliefs both because doing so would have been very time-consuming and because of the potential strategic behavior it might have induced in the experiment.

Uncertainty treatment and 158 in the *Certainty* treatment—from Simon Fraser University

participated in the *Real Effort* experiment in 2014.

In the *No Effort* experiment, subjects completed ten 21-period search episodes.[7] The

experiment proceeded as described in the Introduction. Offers arrived in each period with

probability 0.25. In each period prior to accepting an offer, subjects stated a binding reservation

wage. If an offer was received and was greater than or equal to the reservation wage, it was

accepted, the search ended, and the subject earned that amount in all remaining periods.

Otherwise, the subject moved to the next period and the process repeated. Subjects did not earn a

search subsidy, and no search occurred in the 21st period. Offers were drawn from a known

uniform distribution over the integers from 1 to 10 ECUs.

After the search episodes, subjects completed a survey measuring the Big Five traits, locus

of control beliefs, optimism, self-esteem, cognitive ability, risk preferences, understanding of

probability, patience, altruism, positive and negative reciprocity, trust, competitiveness

(Helmreich and Spence, 1978), greed (Fortin, 2008), and overestimation (Moore and Healy,

2008).[8] For characteristics such as risk preferences (Croson and Gneezy, 2009), competitiveness

(Niederle and Vesterlund, 2007) and overestimation (Bordalo et al., 2019), the literature

demonstrates gender gaps, while we conjectured that other traits such as greed and patience

might influence reservation wages. Some of the instruments differed in the experiments. In the

No Effort experiment, we used the Ten Item Personality Inventory (TIPI) (Gosling et al., 2003)

rather than the 50-item John et al. (2008) instrument. Both samples include a self-assessment of

risk preferences and a paired lottery task, but the paired lottery task in the *No Effort* experiment

came from the experimentally validated Global Preference Survey (GPS) (Falk et al., 2018 and

---

[7] Both experiments were programmed and conducted in z-Tree (Fischbacher, 2007).
[8] The instruments and the instructions for both experiments are provided in the Appendix.

2023) as do the measures of patience, altruism, positive and negative reciprocity, and trust. We note that while the GPS measures were unincentivized, the GPS items were selected based on their ability to predict choices in corresponding, incentivized experiments (Falk et al., 2023).[9] We used the GPS measures to economize on time given the number of characteristics measured.

The Moore and Healy (2008) overestimation instrument requires subjects to assign a probability for themselves to each possible score on a general trivia quiz prior to taking the quiz. Overestimation is measured as the difference between the expected score and the actual score. The understanding of probability measure consisted of four questions, each meant to assess a different concept (independence, joint probability, expected values, and conditional expected values). In both experiments, subjects were asked to sit quietly until all subjects had completed all parts of the experiment to minimize the effects of outside options on search behavior, and subjects were paid at the same time. Subjects in the *No Effort* experiment were paid for one randomly selected search episode, the outcome of the overestimation instrument, and $0.50 per correct answer on the cognitive ability and understanding of probability measures. Between 2019 and 2023, 120 subjects at the University of Arkansas participated in the *No Effort* experiment.

4. Results

4.1 Reservation wages

Table 1 reports the mean reservation wages by experiment, treatment and gender pooling over all periods in Panel A and using only the first period in Panel B.[10] Using all periods,

---

[9] Consistent with Falk et al.'s (2023) findings, the unincentivized risk preference self-assessment and the incentivized paired lottery risk preference measure in the *Real Effort* experiment are positively and significantly correlated.

[10] We leave the analysis in ECUs because of the differences in conversion rates to US$ (8:1) and CD$ (50:1) and exchange rates over time. Average earnings in the *Real Effort* (*No Effort*) experiment were CD$22.15 (US$24.01). The conversion rates were designed to yield roughly equivalent monetary incentives, but the *No Effort* experiment had additional earnings from the measures of overestimation and understanding of probability.

women's average reservation wages are 17.7%, 14.2%, and 18.0% lower than men's in the *Uncertainty* treatment, *Certainty* treatment, and *No Effort* experiment, respectively—differences that are statistically significant and similar in magnitude to those in Cortés et al.'s (2023) experiment. Women, however, were more likely to accept offers early as a consequence of setting lower reservation wages. As such, there are more reservation wages from men in later periods, and the individuals who remain searching tend to be those with higher reservation wages.[11] Using only the first periods of search, women's average reservation wages are 30.3%, 21.3%, and 23.6% lower than men's in the *Uncertainty* treatment, *Certainty* treatment, and *No Effort* experiment, respectively—differences that are significant at least at the 5% level.[12]

Women's reservation wages are also consistently lower than men's over the entire course of a search episode. Panels A to C of Figure 1 display the mean reservation wages by period and gender for each of the treatments. In the first 15 periods of the *Real Effort* treatments, women's average reservation wages are above men's in only one of 30 periods. In the *No Effort* experiment, women's average reservation wages are never above men's in any period.[13] The gender difference also persists as subjects gain experience searching in the *No Effort* experiment. Reservation wages decrease slightly for both genders after the first five search episodes, but the gender gap in reservation wages actually *increases*. In the first five episodes, the average reservation wages for men and women are 6.36 and 5.33, respectively, while in the last five

---

[11] Across both experiments, women make up 49.9% of first period observations, 47.5% of observations in the first five periods, and 43.3% of observations in the last 15 periods.

[12] In Panel A of Table 1, we report the p-values for Wald tests of the equality of the gender coefficients from regressions of reservation wages on an indicator for being a woman while clustering the standard errors at the subject level. In Panel B for the *Real Effort* data, we report p-values for Wilcoxon rank-sum tests of the equality of the means for the two groups; the p-value for the *No Effort* data is for a Wald test similar to those in Panel A to account for having multiple observations from each subject. Using just the first period of the first episode in the *No Effort* data, men set higher reservation wages than women (7.31 vs. 5.47, Wilcoxon rank-sum p = 0.036).

[13] We focus on the first 15 periods because the last five periods have relatively few observations. Women's average reservation wages in the last five periods are greater than men's in only five of 15 total periods across all three treatments.

episodes they are 6.21 and 4.97. This suggests that the gender gaps are not byproducts of confusion or misunderstanding of the experiment as subjects in these episodes have had the opportunity to learn-by-doing.[14]

Reservation wages differ by gender in the experiments despite the fact that men and women search in the same way, receive offers from common, known offer distributions, and have similar outside options—all factors that might differ by gender in survey data in unobserved ways. To investigate the roots of the gender gap, we turn to the individual characteristics described in Section 3. Table 2 reports the means by gender for these measures by treatment along with p-values for tests of the equality of the means. Among the characteristics measured in both the *Real* and *No Effort* experiments, women were more agreeable, less open and emotionally stable, had lower self-esteem, were less willing to take risks, and had lower cognitive ability across all treatments—though these differences are not statistically significant in every treatment.[15] In the *Real Effort* experiment, men believed that they completed more coding tasks on average than women—a difference that is significant at the 10% level in the *Uncertainty* treatment—even though there were no significant differences in actual performance, which is consistent with men overestimating their performance. Among the characteristics measured only in the *No Effort* experiment, women exhibited less negative reciprocity and were less competitive, but otherwise we find no significant gender differences.

To examine the extent to which these characteristics can explain the gender gaps in reservation wages, we regress the log-reservation wage on an indicator for being a woman, period dummies, and different sets of controls. The *Real Effort* regressions include an indicator

---

[14] Appendix Figure A1 depicts the average reservation wages by gender across the first five search (top panel) and the last five search episodes (bottom panel).

[15] That men perform better than women on the cognitive reflection test is well-documented (e.g., Zhang et al., 2016) and not unique to our sample.

for the *Uncertainty* treatment, while the *No Effort* regressions include episode indicators. We report estimates in Table 3 from separate regressions for the *Real Effort* and *No Effort* experiments using all subject-period observations in Panel A and using only observations from the first period of search in Panel B. While the estimates in Panels A and B exhibit mostly similar patterns, we focus below on the estimates in Panel B given that using only observations from the first period in which all subjects search addresses the potentially non-random selection of subjects who continue searching after the first period.

Without controls other than the treatment and episode indicators in Column 1, women set reservation wages that are 26.7 (33.3) log-points lower than men's in the first periods of the *Real Effort* (*No Effort*) experiment. Column 2 adds to the controls measures of the Big Five personality traits, optimism, locus of control, and self-esteem.[16] In the *Real Effort* (*No Effort*) experiment, controlling for personality reduces the estimated gender gap in reservation wages by 25 (6) percent. The reduction in the gender gap in reservation wages in the *Real Effort* experiment is driven primarily by the inclusion of the Big Five traits, which though not individually significant are jointly significant at the 5% level. More agreeable and conscientious subjects set lower reservation wages, while more open subjects set higher reservation wages. By contrast, including a measure of risk preferences in the controls in Column 3 has no appreciable effect on the gender gap in either experiment.[17]

The *Real Effort* experiment introduced a channel through which overestimation may influence reservation wages as individuals believing offers to be more likely in each period conditional on their effort should hold out for higher offers. In Column 4 for the *Real Effort*

---

[16] All of the coefficient estimates for the regressions in Table 3 are reported in Appendix Table 1.
[17] In Table 3, we use self-assessed willingness to take risks rather than the paired-lottery measures to control for risk preferences because we observe statistically significant gender differences in the former but not the latter. The measures, however, are significantly positively correlated overall, for each gender, and in both experiments.

experiment only, we include among the controls the number of tasks correctly completed and subjects' stated beliefs about the probability of receiving an offer. Accounting for overconfidence in this manner, however, reduces the estimated gender gap in reservation wages in the *Real Effort* experiment by a mere four percent.

Controlling for cognitive ability in Column 5 reduces the estimated gender gap by a further 24 percent in the *Real Effort* experiment. While cognitive ability accounts for a large share of the gender gap in reservation wages in the *Real Effort* experiment, the estimated gender gap in Column 5 remains a very substantial 14.8 log-points (p = 0.078). Moreover, controlling for cognitive ability in the *No Effort* experiment actually increases the estimated gender gap in reservation wages. Finally, Column 6 includes as controls all of the additional measures collected in the *No Effort* experiment. One-standard deviation increases in self-assessed altruism and competitiveness decrease reservation wages by 7.1 ($p = 0.075$) and 7.8 ($p = 0.094$) log-points, respectively, but otherwise the additional controls are statistically insignificant.[18] In this specification, women set reservation wages that are an estimated 38.3 log-points lower than those of men.

All of the estimates in Table 3 are robust to alternative specifications (reported in the Appendix) including the use of the paired lottery task instead of the self-assessed risk preferences as a control, the inclusion of an interaction between the indicators for being a woman and the *Uncertainty* treatment in the *Real Effort* data, alternative methods of controlling for overestimation in the *Real Effort* data, and additional demographics such as age and race. Overall, we conclude that the extensive vector of individual characteristics explains relatively little of the observed gender gap in reservation wages, which is itself observed in every treatment

---

[18] This includes the understanding of probability measure, which implies that a better understanding of probability had no effect on reservation wages despite the stochastic process determining whether offers were received.

regardless of the uncertainties faced by subjects and the processes determining whether offers are received.

The existence of the gender gap in reservation wages, however, does not tell us whether women set reservation wages that are too low, men set reservation wages that are too high, or both. To investigate this issue, Panel C of Figure 1 plots the optimal reservation wage path over the course of a search episode for risk-neutral subjects in the *No Effort* experiment along with the mean reservation wages of men and women in each period. Reservation wages for men decrease over the course of a search episode but are well above the optimal path throughout, while women start below the optimal path, then hew closely to it for several periods before finishing well above it.[19]

Not all subjects are risk-neutral, so the deviations from the optimal reservation wage path in Panel C are not necessarily sub-optimal. Subjects who are risk-neutral or risk-seeking on the paired lottery task should set reservation wages at or above the risk-neutral optimum, while risk-averse subjects should set reservation wages below it. Panel D of Figure 1 plots the average reservation wages by gender and risk preferences against the optimal path using the paired lottery task to classify subjects' risk preferences. The lottery task consisted of choosing between a lottery with a 50% chance of receiving $300 and a 50% chance of receiving $0 and a safe payment, where the value of the safe payment increases from $0 to $300 across choices. A risk-averse subject should switch to the safe choice at a value less than the certainty equivalent of the lottery ($150). Risk-neutral subjects should switch to the safe payment when it is equal to the certainty equivalent, and risk-seekers should switch to the safe choice when it is greater than the

---

[19] The high reservation wages later in search episodes likely reflect sample selection as individuals who set higher reservation wages are more likely to still be searching later in episodes. This dynamic motivates our focus throughout the paper on the early part of search episodes.

certainty equivalent. Using this categorization scheme, 27 women and 27 men (45% of the sample for each gender) were either risk-neutral or risk-seeking (the distributions over these categories differ slightly by gender). We use the paired lotteries risk preference measure in this context because the risk preference typology is well-defined for this measure.

Women who were risk-neutral or risk-seeking set reservation wages that are significantly lower than the risk-neutral optimal reservation wage of 6 in period 1 (5.07, p = 0.008) and period 2 (5.02, p = 0.009).[20] By contrast, men who were risk-averse or risk-neutral (i.e., those who should set reservation wages at or below the risk-neutral reservation wage) have average reservation wages that are not significantly different from the optimal reservation wage in period 1 (6.22, p = 0.385) or period 2 (6.24, p = 0.292). At the beginning of the search episode, women—but not men—set reservation wages that are too low given their risk preferences. In terms of the importance for earnings of setting sub-optimal reservation wages early in search episodes, non-risk averse women whose search ends in the first two periods after accepting an offer less than the optimal reservation wage earned on average 33.7% less over the search episode than they would have had they accepted the risk-neutral optimal reservation wage in the period in which they accepted an offer.

4.2 Search duration, accepted wages, and earnings

A gender difference in reservation wages matters in the laboratory and the labor market only to the extent that it influences earnings and search durations. Higher reservation wages should result in higher accepted wages and longer search durations, but these dynamics have offsetting effects on earnings in the experiments. Table 4 reports averages by treatment and

---

[20] As a robustness check, we instead identify 27 women whose self-assessed willingness to take risks was greater than or equal to 7 (out of 10). These women set reservation wages significantly below the risk-neutral benchmark values in period 1 (5.17, p = 0.017) and in period 2 (5.10, p = 0.027).

gender for search duration, accepted wages, and earnings as well as p-values for tests of gender differences in the means. The evidence on search durations and accepted wages is consistent with our expectations given the gender differences in reservation wages. Men in the *Real Effort* (*No Effort*) experiment spent on average about two (one) more periods searching per episode. Likewise, men accepted wages that are 9 to 13 percent higher than women on average conditional on having accepted a wage (i.e., omitting the 30 (52) search episodes in the *Real Effort* (*No Effort*) data that end without a searcher accepting a wage). Finally, men earned more than women in every treatment: 4.0 percent and 5.4 percent more in the *Real Effort* treatments and 3.4 percent more in the *No Effort* experiment. These earnings differences are not statistically significant, but we note that the p-values in Table 4 are for two-sided tests of the equality of the gender-specific means. [21] Using a one-sided hypothesis test concerning earnings in the *Real Effort* experiment pooling the treatments and a 10% significance level, we would conclude that men earned more than women.[22]

5. Conclusion

In survey data, women indicate that they have lower reservation wages than men, and this gender gap appears to contribute to the gender wage gap. Evidence from surveys, however, must contend with the fact that these self-reported reservation wages are non-binding (and thus possibly cheap talk) but may be correlated with unobserved factors influencing wages. Our

---

[21] The sample size in the *No Effort* experiment was based on the desire to have sufficient power to test the gender equality of reservation wages and to test the equality of these reservation wages with the optimal reservation wage path early in search episodes. The *Real Effort* experiment was not designed to test the gender equality of earnings. The lack of power to detect gender differences in earnings in both experiments stems from the high variance in wage offers. In the *Real Effort* experiment, this high variance in combination with the search subsidy reflected our desire to encourage high reservation wages in order to observe more periods of active search. In the *No Effort* experiment, we elected to replicate a well-known design with a known optimal reservation wage path.

[22] In a survey question at the end of the *Uncertainty* treatment, 70% of subjects indicated that they believed that correctly completing a coding task increased the probability of receiving an offer by a lot or a little. That is, most subjects' beliefs were basically accurate, which resulted in mean reservation wages and coding tasks completed very similar to those in the *Certainty* treatment. This motivates our pooling of the *Real Effort* treatments in Table 4.

experimental results complement survey evidence by providing evidence of a gender gap in reservation wages when these reservation wages are binding and in a controlled environment in which the gender gap cannot be explained by the wage distributions, outside options, productivity or search differences, anticipated discrimination, or the many factors that likely influence reservation wages in actual labor markets. In addition, we show that the gender gap in reservation wages cannot be explained by a large vector of characteristics and preferences including the Big Five personality traits, optimism, self-esteem, locus of control, cognitive ability, understanding of probability, patience, altruism, positive and negative reciprocity, trust, competitiveness, greed, risk preferences and overconfidence.

Most importantly, our study demonstrates that women set reservation wages that are too low relative to theoretically optimal values given their risk preferences—a conclusion studies of reservation wages using survey evidence cannot reach without strong assumptions. In the lab, this suboptimal behavior results in lower earnings. In the labor market, the welfare consequences of sub-optimally low reservation wages could be profound over the course of a career. Our results suggest that active labor market policies encouraging women to hold out for higher wage offers when looking for work could be welfare-enhancing.

References

Andersen, S., Ertac, S., Gneezy, U., List, J.A. and Maximiano, S., 2018. On the cultural basis of gender differences in negotiation. *Experimental Economics*, 21, pp.757-778.

Babcock, L. and Laschever, S., 2003. *Women don't ask: Negotiation and the gender divide*. Princeton University Press, Princeton, N.J.

Barber, B., and Odean, T., 2001. Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116(1), pp.261–292.

Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A., 2019. Beliefs about gender. *American Economic Review*, 109 (3), pp.739-73.

Bowles, H.R., Babcock, L. and Lai, L., 2007. Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes*, 103(1), pp.84-103.

Brown, M., Flinn, C.J. and Schotter, A., 2011. Real-time search in the laboratory and the market. *American Economic Review*, 101(2), pp.948-974.

Brown, S., Roberts, J. and Taylor, K., 2011. The gender reservation wage gap: Evidence from British panel data. *Economics Letters*, 113(1), pp.88-91.

Caliendo, M., Lee, W. and Mahlstedt, R., 2017. The gender wage gap and the role of reservation wages: New evidence for unemployed workers. *Journal of Economic Behavior & Organization*, 136, pp.161-173.

Card, D., Cardoso, A. and Kline, P., 2016. Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, 131(2), pp.633-686.

Cortés, P., Pan, J., Reuben, E., Pilossoph, L. and Zafar, B., 2023. Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab. *The Quarterly Journal of Economics*, forthcoming.

Cox, J. and Oaxaca, R., 1989. Laboratory Experiments with a Finite Horizon Job-Search Model. *Journal of Risk and Uncertainty*, 2, pp.301-330.

Cox, J. and Oaxaca, R., 1992. Direct tests of the reservation wage property. *The Economic Journal*, 102(415), pp.1423-1432.

Croson, R. and Gneezy, U., 2009. Gender Differences in Preferences. Journal of Economic Literature, 47(2), pp. 448-474.

Cukrowska-Torzewska, E., 2021. Gender gap in reservation wages and the choice of education field. *Applied Economics Letters*, 30(3), pp.319-323.

Eriksson, S. and Lagerström, J., 2012. The labor market consequences of gender differences in job search. *Journal of Labor Research*, 33, pp.303-327.

Exley, C., Niederle, M., and Vesterlund, L., 2020. Knowing when to ask: The cost of leaning in. *Journal of Political Economy,* 128(3), pp.816–854.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U., 2018. Global evidence on economic preferences. *Quarterly Journal of Economics,* 133(4), pp.1645–1692.

Falk, A., Becker, A., Dohmen, T., Huffman, D., and Sunde, U., 2023. The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science*, 69(4), pp. 1935–1950.

Fischbacher, U. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10**,** 171–178 (2007).

Fluchtmann, J., Glenny, A., Harmon, N. and Maibom, J., 2023. The Gender Application Gap: Do men and women apply for the same jobs? *Working Paper*

Fortin, N., 2008. The gender wage gap among young adults in the United States: The importance of money versus people. *Journal of Human Resources*, 43(4), pp.884-918.

Frederick, S., 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), pp.25-42.

Gosling, S., Rentfrow, P. and Swann Jr, W., 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), pp.504-528.

Haeckl, S., 2022. Image concerns in ex-ante self-assessments–Gender differences and behavioral consequences. *Labour Economics*, 76, 102166.

Helmreich, R. and Spence, J., 1979. The work and family orientation scale (W0F0). *JSAS Catalog of Selected Documents in Psychology*.

Holt, C. and Laury, S., 2002. Risk Aversion and Incentive Effects. *American Economic Review*, 92(5), 1644-1655.

John, O., Naumann, L., and Soto, C., 2008. Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In: Oliver P. John, R. W. Robbins, and L. A. Pervin (Eds.), Handbook of Personality, Theory, and Research, N. Y. Guilford Press, pp. 114-158

Keith, K. and McWilliams, A., 1999. The returns to mobility and job search by gender. *ILR Review*, 52(3), pp.460-477.

Khan, B. and Majid, M., 2020. A note on the gender reservation wage gap in developing countries. *Scottish Journal of Political Economy*, *67*(5), pp.462-468.

Krueger, A. and Mueller, A., 2016. A Contribution to the Empirics of Reservation Wages. *American Economic Journal: Economic Policy*, 8(1), pp. 142-179.

Le Barbanchon, T., Rathelot, R. and Roulet, A., 2021. Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, *136*(1), pp.381-426.

McGee, A. and McGee, P., 2016. Search, effort, and locus of control. *Journal of Economic Behavior and Organization,* 126 (Part A), pp.89-101.

McKelvie, S., 1989. The Wonderlic Personnel Test: Reliability and validity in an academic setting. *Psychological Reports*, 65(1), pp.161-162.

Moore, D. and Healy, P., 2008. The trouble with overconfidence. *Psychological Review*, 115 (2), p.502.

Niederle, M. and Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122 (3), pp.1067-1101.

Rosenberg, M., 1965. *Society and the Adolescent Self-Image*. Princeton University Press, Princeton, N.J.

Rotter, J., 1966. Generalized Expectancies for Internal Versus External Control of Reinforcement. *Psychological Monographs General and Applied*, 80 (1, Whole No. 609).

Scheier, M., Carver, C., & Bridges, M. 1994. "Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test." *Journal of Personality and Social Psychology*, 67, 1063-1078.

Schotter, A. and Braunstein, Y., 1981. Economic search: an experimental study. *Economic Inquiry*, 19(1), pp.1-25.

Sin, I., Stillman, S. and Fabling, R., 2022. What Drives the Gender Wage Gap? Examining the Roles of Sorting, Productivity Differences, Bargaining, and Discrimination. *Review of Economics and Statistics*, 104(4), pp.636-651.

Zhang, D., Highhouse, S. and Rada, T., 2016. Explaining sex differences on the Cognitive Reflection Test. *Personality and Individual Differences*, 101, pp.425-427.
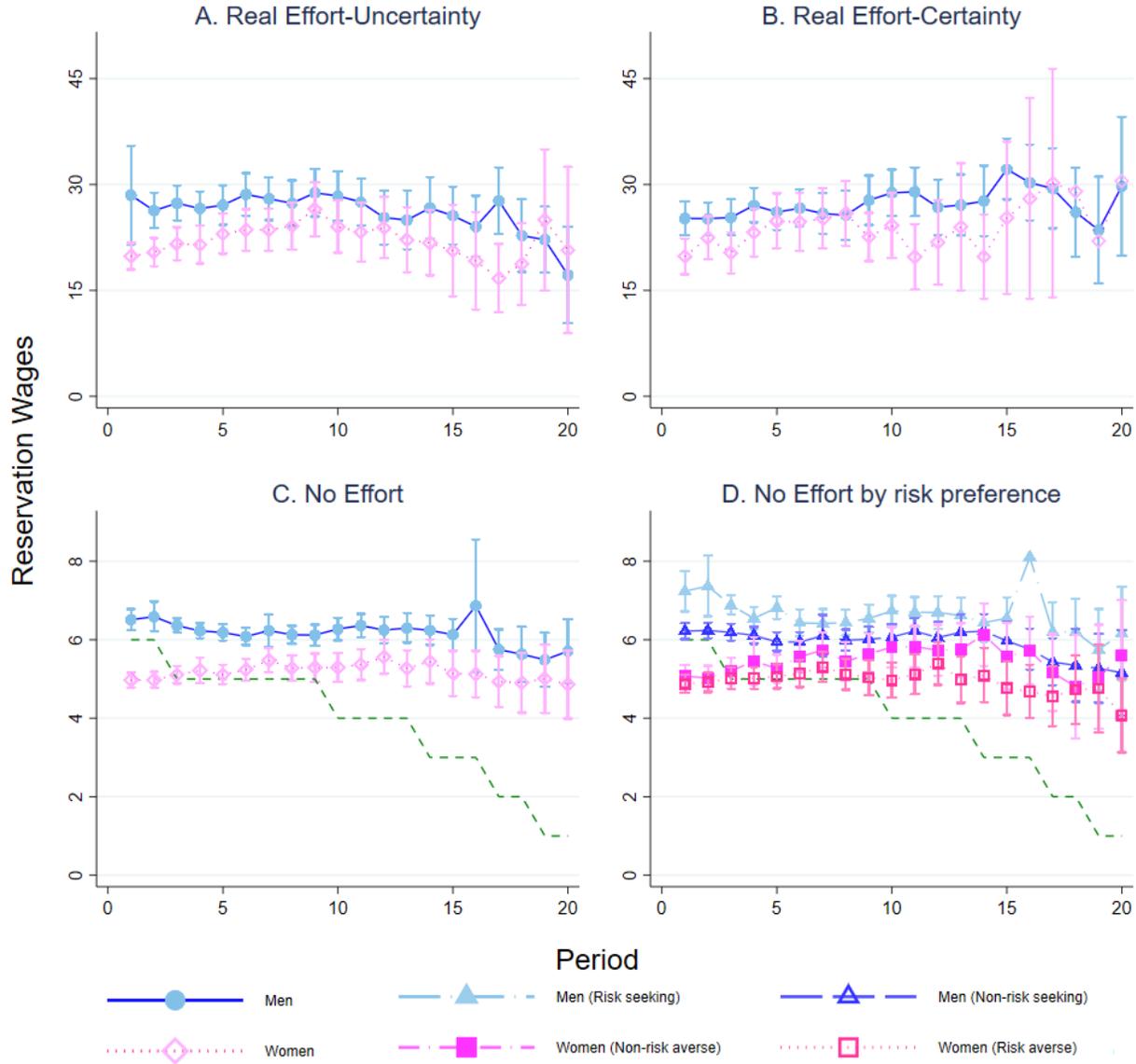
Figure 1: Average reservation wages by period

Notes: Reservation wages are measured in Experimental Currency Units. The green dashed line in the *No Effort* panel is the risk-neutral optimal reservation wage path. The bars around the markers represent 95% confidence intervals. We omit confidence intervals for women in periods 18-20 in *Real Effort Certainty* to preserve the scale of the figure. These confidence intervals are [4.0, 54.0], [-17.4, 61.4], [-153.7, 214.7]. In these periods there are 4, 3, and 2 women searching, respectively. One man set a reservation wage of 80 in the period 16 in the 5th episode. The mean for non-risk averse men in that round excluding that reservation wage is 6.54. In Panel D, we omit the confidence interval for the mean in this period, [4.4, 11.3], to preserve the scale of the figures.

Table 1: Reservation wages by gender and treatment

Real Effort Uncertainty		Real Effort Certainty		No Effort	
(1)	(2)	(3)	(4)	(5)	(6)
A. Using all subject-period observations					
Women	Men	Women	Men	Women	Men
22.00	26.74	22.81	26.60	5.16	6.29
(10.14)	(14.36)	(10.46)	(10.23)	(2.49)	(2.94)
N = 733	N = 753	N = 419	N = 752	N = 4,085	N = 4628
p = 0.002		p = 0.027		p < 0.001	
B. Using the first period observations only					
Women	Men	Women	Men	Women	Men
19.85	28.49	19.82	25.20	4.97	6.51
(9.77)	(32.27)	(10.53)	(11.32)	(2.35)	(3.38)
N = 105	N = 84	N = 68	N = 90	N = 600	N = 600
p = 0.001		p = 0.010		p < 0.001	

Notes: For the tests in Panel A, we regress reservation wages on a female dummy and cluster the standard errors at the subject-level. The reported p-values are for the Wald test of the equality of the gender coefficients. In Panel B, the p-value reported for the *No Effort* treatment is from the same test as in Panel A to account for having multiple observations from the same subject. For the *Real Effort* treatments, the p-values in Panel B are for Wilcoxon rank-sum tests of the equality of the means.

Table 2: Means subject characteristics by gender and treatment

	Real Effort Uncertainty			Real Effort Certainty			No Effort		
	Women N=105	Men N=84	p	Women N = 68	Men N=90	p	Women N = 60	Men N=60	p
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Agreeable.	38.0	37.0	0.26	39.5	37.8	0.07	9.3	8.6	0.15
Extroversion	31.2	32.1	0.27	31.4	32.1	0.53	8.6	8.3	0.51
Conscientious	34.0	33.4	0.47	34.9	33.8	0.14	11.1	10.7	0.37
Openness	32.4	34.1	0.01	34.1	35.1	0.11	10.1	10.6	0.32
Emotional Stability	28.5	31.9	0.01	28.7	31.8	0.01	8.6	9.4	0.12
Optimism	20.3	21.2	0.18	21.7	21.5	0.79	20.0	20.6	0.47
Self-esteem	35.0	37.4	0.02	36.2	37.9	0.08	34.7	37.4	0.03
Locus of control	11.1	11.3	0.69	11.6	12.0	0.15	11.8	12.3	0.12
Risk (p.l.)	5.9	6.0	0.59	5.8	5.6	0.71	12.8	12.2	0.98
Risk (s.a.)	5.7	6.0	0.09	5.9	6.5	0.10	5.7	6.6	0.04
Cognitive ability	2.0	2.6	0.01	1.8	2.7	0.00	2.2	3.0	0.01
Correct items	11.9	10.9	0.14	12.0	12.1	0.84			
Beliefs	12.3	15.8	0.08	13.6	14.0	0.52			
Overconfidence							0.7	0.8	0.80
Probability understanding							1.4	1.7	0.12
Patience (s.a.)							6.2	6.8	0.20
Patience (p.c)							14.5	12.0	0.30
Trust (send.)							11.1	11.6	0.67
Trust (rec.)							0.5	0.5	0.48
Trust (s.a.)							5.7	5.4	0.67
Altruism (s.a.)							6.2	5.8	0.54
Altruism (d.)							179.6	148.9	0.30
(-) reciprocity (s.a)							5.0	5.5	0.22
(-) reciprocity (MAO)							37.8	42.1	0.08
(+) reciprocity (Gift)							20.8	20.2	0.49
Competitive.							17.7	19.6	0.00
Greed							1.0	0.9	0.63

Notes: Standard deviations are given in parentheses. The p-values are for Wilcoxon rank-sum tests of the equality of the means. All of the measures are increasing in the associated trait; for risk preferences, the measures are increasing in willingness to take risks. The Big 5 instruments were John et al. (2008) in the *Real Effort* and Gosling et al. (2003) in the *No Effort*. The self-assessment ("s.a.") of risk preferences asks subjects to rate their willingness to take risks on a scale from 0 to 10. The paired lottery task ("p.l.") asks subject to choose between a sure option and a risky option. The Holt and Laury (2002) instrument has 10 paired lotteries in *Real Effort*

treatment while the Falk et al. (2016, 2018) instrument has 31 paired lotteries in the *No Effort* treatment. The paired choice task ("p.c.") measuring patience asks subjects to make 25 choices between an amount of money paid out today versus a larger amount of money paid out in the future. The cognitive ability measure consists of the three question CRT (Frederick, 2005) and three Wonderlic questions (McKelvie, 1989). In the *No Effort*, overconfidence is measured using Moore and Healy's (2008) overestimation measure for one intermediate difficulty geography quiz from their original study. The measures collected only in the *No Effort* treatment come from the GPS (Falk et al., 2016 and 2018) except for competitiveness (*Work and Family Orientation Questionnaire*; Helmreich and Spence, 1978), greed (Fortin, 2008), and probabilistic understanding. The self-assessments ask subjects to rate themselves on a scale from 0-10 on their willingness to wait for monetary gains, to be altruistic to someone, to punish someone, etc. The two measures of trust come from playing the roles of sender and receiver in a hypothetical trust game, where the latter is the average fraction returned across four hypothetical amounts received (5, 10, 15, and 20). Negative ("-") reciprocity is the subject's minimum acceptable offer (MAO) in a hypothetical ultimatum game. Altruism (Donation) and Positive Reciprocity (Gift) are dollar values of responses in hypothetical situations. Probabilistic understanding is measured by a four-question quiz asking about statistical independence, joint probability, conditional probability, and expected value. All of the instruments are provided in the Appendix.

Table 3: Marginal effect of being a women on reservation wages

	(1)	(2)	(3)	(4)	(5)	(6)
	A. Using all subject-period observations					
Real effort	-0.195***	-0.178***	-0.178***	-0.188***	-0.163***	
[N =2,657]	(0.054)	(0.059)	(0.059)	(0.056)	(0.059)	
No effort	-0.268***	-0.253***	-0.237***		-0.251***	-0.292***
[N =8,713]	(0.062)	(0.061)	(0.064)		(0.063)	(0.063)
	B. Using first period observations only					
Real effort	-0.267***	-0.201**	-0.204**	-0.196**	-0.148*	
[N = 347]	(0.074)	(0.081)	(0.083)	(0.081)	(0.084)	
No effort	-0.333***	-0.312***	-0.303***		-0.328***	-0.383***
[N =1,200]	(0.074)	(0.073)	(0.073)		(0.068)	(0.075)
Controls:						
Personality		✓	✓	✓	✓	✓
Risk		✓	✓	✓	✓	✓
Beliefs and Task Performance (*RE* only)				✓	✓	
Cognitive					✓	✓
All (*NE* only)						✓

Notes: The table reports the marginal effect of being a woman from regressions of the log of reservation wages on various sets of controls. All specifications include period indicators within a search episode. The *Real Effort* specifications include an indicator for the *Uncertainty* treatment, and the *No Effort* specifications also include indicators for each search episode. Standard errors clustered at the subject level are reported in parentheses in Panel A and for the *No Effort* data in Panel B, and robust standard errors are reported in parentheses in Panel B for the *Real Effort* data. All individual difference measures are standardized within the treatment (*Real Effort* or *No Effort*) to have mean 0 and standard deviation of 1. The risk preference measure is a self-assessed willingness to take risks. The personality controls are the Big 5, optimism, self-esteem, locus of control, and overconfidence. The cognitive ability measure is the Cognitive Reflection Test and three Wonderlic questions. The other controls in the last column ("All") include probabilistic understanding, a self-assessment of patience, trust measured from the perspective of the sender and receiver in a trust game, self-assessed altruism, a self-assessment of negative reciprocity, altruism, negative reciprocity as measured by a hypothetical ultimatum game, positive reciprocity, competitiveness, greed, patience as measured by a paired choice list, and overestimation. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 4: Search duration, accepted wages, and earnings

Real Effort Uncertainty		Real Effort Certainty		No Effort	
(1)		(2)		(3)	

A. Search duration (# of periods)

Women	Men	Women	Men	Women	Men
N = 105	N = 84	N = 68	N = 90	N = 600	N = 600
6.98	8.96	6.19	8.39	6.81	7.71
(5.75)	(7.08)	(5.47)	(6.19)	(5.66)	(5.87)
p = 0.039		p = 0.019		p = 0.028	
Pooling Uncertainty and Certainty: p = 0.001					

B. Accepted wages (ECUs)

Women	Men	Women	Men	Women	Men
N = 99	N = 70	N = 66	N = 82	N = 578	N = 570
24.19	27.64	24.39	27.99	6.80	7.46
(12.41)	(11.48)	(12.20)	(12.17)	(2.26)	(1.92)
p = 0.064		p = 0.077		p < 0.001	
Pooling Uncertainty and Certainty: p = 0.005					

C. Earnings (ECUs)

Women	Men	Women	Men	Women	Men
N = 105	N = 84	N = 68	N = 90	N = 600	N = 600
435.90	453.21	445.07	468.89	95.24	98.60
(186.97)	(181.25)	(199.14)	(181.99)	(50.98)	(52.31)
p = 0.521		p = 0.441		p = 0.300	
Pooling Uncertainty and Certainty: p = 0.138					

Notes: For *Real Effort* treatments, the p-values are for two-sided Wilcoxon rank-sum of the equality of the means. For the *No Effort* experiment, we regress the outcome on a female dummy and cluster the standard errors at the subject-level. The reported p-values are for two-sided tests Wald test of the equality of the gender coefficients. Accepted wages are conditional on having accepted a wage in a search episode.