DO COMPETITIVE WORKPLACES DETER FEMALE WORKERS? A LARGE-SCALE NATURAL FIELD EXPERIMENT ON JOB-ENTRY DECISIONS

Jeffrey A. Flory, *Claremont McKenna College*
Andreas Leibbrandt, *Monash University*
John A. List, *University of Chicago and Monash University*

22 July 2014

Abstract: An important line of research using laboratory experiments has provided a new potential reason for gender imbalances in labor markets: men are more competitively inclined than women. Whether, and to what extent, gender differences in attitudes toward competition lead to differences in naturally-occurring labor markets remains an open question. To examine this, we run a natural field experiment on job-entry decisions where we randomize almost 9,000 job-seekers into different compensation regimes. By varying the role that individual competition plays in setting the wage and the gender composition, we examine whether a competitive compensation regime, by itself, can cause differential job entry. The data highlight the power of the compensation regime in that women disproportionately shy away from competitive work settings. Yet, there are important factors that attenuate the gender differences, including whether the job is performed in teams, whether the position has overt gender associations, and the age of the job-seekers. We also find that the effect is most pronounced in labor markets with attractive alternative employment options. Furthermore, our results suggest that preferences over uncertainty can be just as important as preferences over competition per se in driving job-entry choices.

Key Words: competitiveness, field experiment, gender, labor market.
JEL codes: C93, J23, J33.

1. Introduction

Whether measured by wages, unemployment rates, or career advancement, persistent gender imbalances exist in labor markets (see, e.g., Goldin, 1990; Altonji and Blank, 1999; Blau and Kahn, 2000; Bertrand and Hallock, 2001; Azmat et al, 2006).[1] The continued existence of significant gender differences is disturbing not only from an equity perspective, but also for its potential impacts on economic growth and even fertility (Galor and Weil, 1996). Several hypotheses have been proposed to explain why such gender inequities exist, with the weight of attention being paid to gender differences in human capital (Blau and Kahn, 2000), discrimination and stereotypes against women (Spencer et al., 1999; Goldin and Rouse, 2000, Reuben et al, 2010), and differences in time taken out of the labor force (Phipps et al., 2001; Hotchkiss and Pitts, 2007). More recently, an experimental literature has emerged that reports another potential source for the observed gender imbalance: women shy away from competitive workplaces whereas men covet, and even thrive in, competitive environments (Gneezy et al., 2003; Niederle and Vesterlund, 2007).

The empirical foundation for this insight arises from gender and competition laboratory experiments. As a brief background, such experiments typically proceed as follows. First, the experimenter recruits a group of students to participate in an experiment. Second, once situated, the experimenter introduces a task—solving mazes, completing math problems, throwing a ball in a bucket, etc.—and asks each subject to choose their preferred compensation regime. Subjects typically have a choice between either a piece rate or a tournament incentive scheme. For example, under the former, the subject is paid $1 per successful attempt. In the latter, the subject is paid $3 per successful attempt if she outperforms an anonymous partner, and zero otherwise. A stylized result that has emerged from these experiments is that men tend to prefer the competitive environment over the non-competitive environment whereas the opposite is true for women, even in tasks where women are more able.

[1] Blau and Kahn (2000) report that according to one index of occupational segregation by sex, 54% of all women in the workforce in 1997 would have to change jobs to equalize the occupational distribution of men and women. They note that in 1999, women were still much more concentrated than men in administrative support and service occupations (accounted for by 41% of the female labor force vs. 15% that of men). They also report that weekly earnings of female full-time workers, though they rose from 61% of men's comparable earnings to 76.5% over 1978 to 1999, appear to have reached a maximum in the mid-1990s at around three-quarter's the amount of male earnings. Bertrand and Hallock (2001) report that among the five highest-paid positions at each of a large sample of US firms, only 2.5% are women, and that they earn 45% less than the men.

These experiments provide key insights into the relationship between gender and choices over competition. Yet, to date there remains little evidence on the extent to which these findings predict behavior in naturally-occurring labor markets, and what the economic consequences are. This is what our paper offers. Stepping back from the burgeoning literature in laboratory experiments (Gneezy et al., 2003; Niederle and Vesterlund, 2007; Gneezy et al., 2009; Cason et al., 2010, Dohmen and Falk, 2011; Gupta et al., 2011; Balafoutas and Sutter, 2012; Dargnies, forthcoming), we advance this line of research in a new direction by analyzing agents in the process of making economic choices in naturally occurring labor markets. Acknowledging that this approach relinquishes a degree of control over some factors more easily manipulated or measured in the lab, we believe the insights it provides on questions testable only through field data constitute a crucial complement to a literature developed under the tight controls of the laboratory. Through experimental manipulations in the field, we are able to examine *directly* whether the persisting gender gap in labor markets can be at least partly attributed to responses to compensation regimes characterized by competition and earnings uncertainty, two elements that increasingly distinguish opportunities for professional advancement (Lemieux et al., 2009).

To test whether men and women respond differently to employment contracts characterized by competition and uncertainty, we conduct a natural field experiment on job-entry decisions in sixteen major US cities.[2] This is the first known attempt to analyze whether competition incentives affect men's and women's labor choices differently using a controlled natural economic setting where subjects are unaware they are part of an experiment. We posted employment advertisements to an internet job-board in cities with different market wages and randomized interested job-seekers into *different compensation regimes* for the *same* job. The set of possible compensation regimes was identical in all cities. Job-seekers were randomly assigned into treatments offering fixed-wage compensation, compensation depending mildly on individual relative performance, compensation depending heavily on individual relative performance, team relative performance, or on elements of uncertainty. Thereafter, each job seeker decided whether to apply formally for the position.

[2] By "natural field experiment," we refer to the definition articulated in Harrison and List (2004): an experiment with a non-standard subject pool (i.e. not students), with field context in the task and information set, and where the subjects are naturally undertaking the tasks as part of their normal economic goings-on without knowing they are in an experiment. Much of the discussion on laboratory experiments can be equally applied to artefactual field experiments (Harrison and List), also sometimes referred to as "lab-in-the-field" experiments.

In addition, we advertised two different versions of the job. One ad presented a version of the position with male connotations, while the other removed these connotations. Comparing the application patterns for these two versions of the position renders it possible to clarify the relevance of task-dependence and gender-job associations. In their seminal paper, Akerlof and Kranton (2000) argue that gender-job associations are drivers of gender-specific employment patterns and predict that women sort into employments whose requirements match construed female attributes. More recently, evidence is beginning to emerge from the lab that suggests the different responses by men and women to competition may depend partly on the tasks performed. For example, Shurchkov (2012) finds that the standard result that men are much more willing to compete in the lab when using a math task disappears when switching to a verbal task suggesting that gender-task stereotypes drive such gender differences. The question of task-dependency of gender differences in competitiveness remains open, but the suggestion is provocative.[3] Whether the effect of relative performance pay on gender composition of workers is sensitive to aspects of an employment position other than the contract, such as gender associations with the position, has important practical implications. Differentiating the position we advertise such that one version has overtly male associations allows us to explore in the field whether minor adjustments in the position, gender associations surrounding the compensation regime, or interactions between these two, affect each gender's response to the contract environment differently.

With 211 to 690 interested job-seekers per city, a total of 6,779 subjects participated in our first natural field experiment. From these individuals, we find some intriguing data patterns. First, we find as the compensation package becomes more heavily reliant on individual relative performance, the applicant pool becomes significantly more male dominated. In the limit, the gender gap in applications more than doubles when a large fraction of the wage (50%) depends on relative performance. This data pattern is remarkably consistent with the literature using laboratory experiments. Yet, a surprising finding is that this result is not driven by any attraction to competitive work environments among men. Rather, we observe that women, and to some

[3] For a more full assessment of this nascent thread of the gender-competition literature that examines the sensitivity of findings to the tasks used in the lab, see Niederle and Vesterlund (2011).

extent even men, are repelled by competitive work environments, but that women have a stronger aversion to them.[4]

We also find several factors that mitigate the effect of competition-based pay on gender composition of applicants. For example, reducing the percentage of compensation determined by relative performance, hiring into jobs where the work is completed in teams (for supporting laboratory evidence see Dargnies, forthcoming), and hiring into versions of the job that have removed masculine connotations each have an important role in substantially attenuating the observed gender differences. Likewise, we find a link between gender-based differences in competitiveness and the age of job-seekers: older age cohorts show much less of a gender gap than their younger counterparts. Moreover, we also observe that the broader economic environment seems to play a role for entry into competitive workplaces. In particular, we find that the gender gap in applications for competitive workplaces is correlated with prevailing market wages: the gender gap is most prominent in areas with higher local wages. Simply put, women are more likely to walk away from competitive workplaces if there are good outside options – i.e., positions with comparatively high fixed-wages—but not if these outside options are lacking.

Finally, we run a second natural field experiment to shed further light on the underlying mechanisms at work. In this second set of experimental treatments, run across twelve cities with 2,189 different job-seekers, we find that expectations about gender composition of the work environment, a potential driver of the findings in the first experiment, have little effect. That is, neither the gender of co-workers nor the gender of the supervisor affects the gender gap in applications for competitive workplaces. We also find that the gender difference in job-seekers' responses to compensation with a high degree of uncertainty closely resembles the gender difference in responses to compensation heavily reliant on individual relative performance. This lends insight into the relative importance of the primitives at work, and raises the possibility that high wage uncertainty may be just as important as competition per se in affecting job-entry choices and the gender gap.

The remainder of our study is constructed as follows. Section 2 discusses the designs of the natural field experiments. Sections 3-8 describe the various findings. Section 9 concludes.

[4] This finding is consistent with recent work on the negative impacts of competition and inequality in labor markets (see Card et al, 2010 on job satisfaction and Wilkinson and Pickett, 2009 on public health).

2. Experimental Design

To investigate gender differences in job-entry decisions, we conduct two natural field experiments. The first natural field experiment uses a 2 × 6 × 2 design in which we vary the employment advertisement, compensation scheme, and application procedure. The second natural field experiment holds the general reward structure constant, but varies gender composition. The design, presented in Figure 1, allows us to isolate the impact of the contract environment on the *proportion* of initially interested individuals who ultimately apply. Knowing the proportion of interested job-seekers that remain interested upon learning the salary regime requires knowing not only the number of final applicants, but also the number of initially interested. We therefore employ a two-stage experimental method.

As Figure 1 shows, we first advertise the position, without reference to the compensation scheme or the gender composition. Then, after a job-seeker expresses interest in the position, we inform the individual of the compensation scheme or the gender composition, and record whether he or she chooses to apply for the job.[5] This method allows us to cleanly randomize across subjects within each city and within the same time period, in a manner that leaves the normalcy of the field setting undisturbed. It furthermore allows us to gather individual characteristics even on those subjects who expressed interest but chose not to apply after they were informed about the compensation scheme or the gender composition.

We posted the job ads in a total of sixteen major US metropolitan areas.[6] The cities were selected with the twin goals of representing a variety of geographical regions of the US which are characterized by different market wages and maximizing the pool of job-seekers from each area (all cities from which we sampled are among the top 25 most populous cities in the

[5] A simpler approach might have been to post employment ads which included the compensation regime at the outset, then simply compare the ratio of male to female applicants across treatments. However, unable to post differing salary regimes for the same job in the same city at the same time, we would have been forced to either vary the salary regime across cities or post different salary regimes for the same job sequentially within the same city, allowing for potential temporal and spatial confounds. In addition, through this method, we would only observe job-seekers who decide to apply after already knowing the salary scheme; the number and characteristics of job-seekers who would have been interested in the position without knowing the compensation scheme would remain unknown. Inferences based on changes in absolute numbers of final applicants are much more limited. For example, using just the male/female application ratio without knowing the underlying gender ratio of those interested prior to knowing the compensation scheme is problematic. This is because those ratios and other characteristics related to gender such as education might differ across the applicant pools for different employment ads.

[6] The internet job-boards are city-specific. The 16 cities included in our study are Atlanta, Boston, Chicago, Washington DC, Dallas, Denver, Houston, Los Angeles, Miami, New York, Philadelphia, Phoenix, Portland, San Diego, San Francisco, and Seattle.

country). The order of the ads was posted randomly across cities from January-April of 2010 (first natural field experiment) and May-July of 2011 as well as June-July 2013 (second natural field experiment). Ads were posted on one of the most highly-frequented internet job boards in the country, and were always specific to a given city where we were seeking applicants. Communication with subjects was through email, and every interaction was carefully scripted (see Appendix A). At the end of the experiment, we offered real jobs to applicants in every city and for each position we created. Thirty applicants were hired in total.

2.1 Natural field experiment I: Employment advertisements and application procedures

The postings advertised openings for a position as an administrative assistant, the most common occupation in the US. We decided to advertise administrative assistant jobs because they are the most common type of position to hold (13% of the workforce) and because there is a significant gender gap in wages for administrative and office support jobs (women earned 84.3% of the median weekly earnings of their male counterparts).[7] In addition, an office support position is also very convenient for a natural field experiment. It is fairly simple to create administrative tasks for a subject hired into the position, and it is also quite amenable to setting up remote tasks that can be performed with an internet connection from different locations.

We used two different versions – one sports ad that had masculine connotations, the other a general ad that has removed those masculine connotations. The advertisements were written to closely mirror other want-ads for similar positions. Both versions were posted under "Admin/Office Jobs", and had three short paragraphs (see Appendix A). The first paragraph identified who we were, where we were located, and said that we were looking for an

[7] In 2008 and 2009, office and administrative support jobs were the most common type of position to hold (Bureau of Labor Statistics *Employment and Earnings*, January 2010, Vol. 57(1). Available online: http://www.bls.gov/opub/ee/empearn201001.pdf). Moreover, data from the U.S. Department of Labor (Data from 2010, the year our experiments were conducted, Report 1031, available at: http://www.bls.gov/cps/cpswom2010.pdf) shows that the gender gap in wages is similar in occupations with low education requirements as compared to those requiring higher education. For occupations that do not require a college degree, women earned 82.7% of the median weekly earnings of their male counterparts. For those requiring a B.A. or higher, women earned 80.1% of the median weekly earnings of their male counterparts. One drawback to using office jobs is that they are disproportionately occupied by women. In 2009, 6.3% of the male labor force and 20.4% of the female labor force held office and administrative support occupations (Bureau of Labor Statistics *Employment and Earnings*, January 2010, Vol. 57(1)). In 2001, 79% of administrative support and clerical positions were occupied by women (Gabriel and Schmidtz, "Gender Differences in Occupational Distributions Among Workers," *Monthly Labor Review*, June 2007, Vol. 130(6), Bureau of Labor Statistics.)

administrative assistant in their area to help with a project.[8] The second paragraph detailed the job-tasks (gathering information on local events, preparing short reports, light correspondence, proofreading, filing, and other typical secretarial and office/clerical duties). The third paragraph was a single sentence requesting those who are interested to email their CV or resume. Requesting the job-seekers' resumes before we inform them of the compensation scheme is not unusual on online job boards, and was necessary in order to gather agent characteristics on the full sample of subjects. The advertisement finished with a sign-off from a current employee of our organization, followed by a few brief generalities.

Our job-advertisement with masculine connotations describes administrative assistant duties in an environment surrounded by sports. The general ad is isomorphic to the sports ad, with the exception that it has no overt gender associations and instead has a general focus.[9] The text of the two ads is identical, except for minor adjustments aimed to influence the gender distribution of interested job-seekers. Specifically, the posting for the ad with masculine connotations mentions a variety of pro-, semi-pro-, and college sports; while that for the general ad mentions nothing about sports.

Note that the only mention of compensation in the job advertisement is that the position is paid on an hourly basis. No information is provided at this point about the duration of the position, number of hours, or opportunities for growth. This enables us to capture the broadest possible sample of job-seekers in the area who are interested in secretarial or clerical positions. While some ads do explicitly detail compensation and duration, we verified that it is common for other employers to leave out specifics on hours and salary, so that our ads remained natural.

Within cities, we were careful to post the ads for the sports-version and general-version of the position within a short enough time frame of each other to avoid any possible temporal-based selection issues. Yet we still spaced them two days apart, intentionally allowing for many other ads to be posted to the job board between them, to ensure job-seekers did not mistakenly believe they were expressing interest in both positions when responding to only one advertisement. We

[8] While not all other employers post self-identifying information, many do. We chose to explicitly identify a genuine organization – one with its own web pages, physical address, phone numbers, etc. – to enhance the normalcy of the situation and minimize any risk of suspicion. Having a genuine employee of the organization sign off the email added further insurance, should any job-seekers wish to do a brief internet search to verify the ad's authenticity.

[9] Working as an administrative assistant is generally a female-dominated career and thus it is possible that the general ad by itself is perceived as female-oriented. Indeed, the general version of the ad resulted in a pool of subjects the substantial majority of which were women.

posted each advertisement once in each city – on Monday or Wednesday, approximately 10 am local time – randomly alternating whether the sports ad or general ad came first.

In addition to varying the gender-framing of the position by using two different ads, we manipulate the application procedure. By asking job-seekers to fill out application questionnaires with varying lengths, we change the requested level of investment in the application.[10] The application questionnaires were randomized at the city level. In eight cities, job-seekers had to fill out a long questionnaire with four interview questions, while in the other eight cities the questionnaire was short and contained only one question (see Appendix A).

2.2 Natural field experiment I: Compensation scheme treatments

Within each of the two jobs in a given city, we randomized job-seekers who expressed interest in the position into one of six different treatments. The characteristics of the treatments are summarized in Table 1 and the scripts are listed in Appendix A. Subjects were not given the treatment until after they had already expressed interest in the job, and they received the treatment usually within 24 hours of expressing this interest. The first two treatments use the same compensation scheme, a fixed-wage of $15 an hour, but differ in whether the job is done in teams (T1) or alone (T2). These two treatments serve as our primary control group, against which we compare behavior of job-seekers under the alternative compensation schemes.

The next two treatments T3 and T4 consist of individual tournament-based salary regimes.[11] The difference between the two is that T3 has a relatively low competition-based reward, whereas T4 has a much higher one. We tell the job-seeker she will be matched with another person we are currently hiring into the same position, and that whichever of the two performs better earns a bonus. For the individual tournament-low treatment T3, the base salary is $13.50 an hour and the bonus is equivalent to $3 an hour. The worker thus earns $13.50 an hour or $16.50 an hour, depending on how she performs with respect to her co-coworker. For the

[10] Employers often use self-selection devices to sort out better suited applicants (Spence, 1973; Salop and Salop, 1976). We implemented this commonly used device in our experiment by manipulating the effort needed to complete an application.

[11] Other employers offering similar jobs on this job board also sometimes use compensation schemes involving bonuses. However, as an added measure to preserve the normalcy of the economic environment, we added language to the non-standard compensation treatments to explain the nature of the contract. For example, we included in all treatments the sentence, "We have frequent deadlines, and timely quality information from you is important." The intention here was to provide a natural rationale for why we might want to encourage competition among employees. In addition, for the two team treatments (team-hourly, and team-tournament), we added the sentence "As the work is best done in teams, you will be paired with one co-worker."

individual-high treatment T4, the base salary is $12 an hour, with a bonus equal to $6 an hour for the most productive worker. So the worker earns $12 an hour or $18 an hour, depending on relative performance. Dividing the individual relative performance pay scheme into a high-stakes competition treatment and low-stakes competition treatment helps provide a rough measure of the elasticity of the gender gap in competitiveness, and the sensitivity of each gender's response to differing degrees to which competition affects total compensation.

The fifth treatment T5, group-based competition, places two new hires into a team, and makes their compensation contingent on the group's relative performance with respect to another similarly composed team. The base payment ($12) and bonus ($6) in this compensation scheme are as in T4. We selected the highly competitive incentive for this treatment in order to intensify any effect, and to render possible comparisons to treatment T4.

The final treatment aims to capture the gender difference in attitudes towards uncertainty over job compensation. Here, the base salary for the position is $13.50 per hour, plus a bonus that translates to $3 per hour if their assistance happens to contribute to journal article publications.[12] Just as in the competition treatments, there is uncertainty over the payment and the worker is allowed to believe that she has the ability to influence the outcome through her own labor. In contrast to the competition treatments, however, she is not performing against anyone else.

The treatments were communicated via emails. All job-seekers expressing interest in the job received an email. All the emails contained exactly the same text, except the few lines describing compensation (see Appendix A). We thanked the job-seeker for their interest in the position, let them know everyone who contacted us about the job was receiving a general initial email from us with basic information, and included the treatment-specific script. At this point, subjects were still unaware of the job's duration and number of hours per week.[13]

[12] We were careful to emphasize the clerical role of the position to subjects, and to clarify that they would merely be providing administrative assistance. This was to ensure that job-seekers did not think they were applying for a position where they might be involved in the actual writing of publications. We should note it is conceivable that the uncertainty treatment T6 may introduce an additional meaning to the work not present in the other treatments, if subjects in the other treatments did not understand that their work led to the potential publication of articles. The marginal adjustment to the meaning of the task may complicate the interpretation of the findings in treatment T6 if there is a gender/task/treatment interaction. However, we have addressed this shortcoming in the second natural field experiment in which we run a new uncertainty treatment.

[13] They are informed of these details at the time we offer the job, at which point they decide whether to accept it. We were careful in the emails to create an environment open to questions, apologizing for any questions about the job

2.3 Natural field experiment II: Gender composition treatments

The second natural field experiment digs deeper into the mechanisms underlying the findings from the first experiment. The procedure was identical to that of the first experiment, except that we posted only the male job ad, and used different treatments. For the first five treatments, the compensation structure is identical to that of T4 in the first experiment (high-stakes individual competition), except that we manipulate the gender structure of the contract. While 2T4 of the second experiment is identical to T4 of the first experiment and thus includes no information about gender, 2T4a and 2T4b provide information on the gender of the competitor (male in 2T4a, female in 2T4b), and 2T4c and 2T4d provide information on the gender of the supervisor (male in 2T4c, female in 2T4d).

These new treatments allow us to test whether the effects of competitive workplaces identified in the first natural field experiment are driven by expectations about the gender structure of the contract. If job-seekers believe the coworker or boss will be male, for example, this may affect each gender's willingness to apply under the competitive regime differently. Particularly for the job with more masculine connotations, it is possible job-seekers expect the competitor and supervisor to be male. With these five additional treatments, we test whether findings from the first experiment are driven by the expected gender composition of the environment, or instead point to a different explanation (such as a general preference for competition per se).

The final treatment, 2T6, introduces heavy wage uncertainty and removes explicit competition (but does not include information on gender). The bonus in 2T6 is identical to that of 2T4 ($6), but is not earned by out-performing a coworker. Just as in T6 of the first experiment, it is received if the worker's assistance ends up contributing to journal article publications. Comparison of 2T6 and 2T4 sheds further light on the role played by uncertainty in the effects of relative performance-based pay on gender composition of applicants. The characteristics of the treatments are summarized in Table 1 and the scripts are listed in Appendix B.

---

they may have asked to which we have not responded, and welcoming further inquiries. We could not interact with subjects before they received the salary regime information and made their choice about applying for the job without risking heterogeneous treatment. Yet by creating an environment welcome to inquiries, we minimize the possibility that job-seekers feel jilted or ignored, which might have a (gender-dependent) effect on their propensity to apply.

2.4 The response variable and job-seeker characteristics

The key variable of interest is the individual decision of whether or not to apply, once the compensation method is known. Every individual who contacted us to express interest in the job constitutes a 'subject' in our experiment. In order to actually apply, however, the interested job-seeker had to fill out the interview questionnaire and send it back to us.[14] We therefore classify all subjects who returned the questionnaire to us as having applied, and those who did not return the questionnaire as having not applied.

Using internet-based job applications presents a minor challenge in determining the gender of subjects, since it cannot be visually observed. In addition, directly asking a subject's gender (i.e. via email) has important legal implications, and might seem unnatural and disrupt the normalcy of the field setting. Further, directly asking for gender could have altered the subject's decision of whether or not to apply.

In order to determine gender, we use each subject's first name and employ a three-tier method. The vast majority of names are assigned gender based on probabilities derived from the Social Security Administration (SSA) database on name popularity by gender and birth year.[15] For any names which are not included in the SSA database, we use an additional database created by Geoff Peters which calculates gender ratios by first name, using the internet to analyze patterns of name-usage for over 100,000 first names.[16] This second database is also used as an additional check on the SSA-based assignments in cases where the gender ratio

[14] This is made clear in the treatment email describing compensation, where we indicate that returning the questionnaire completes their application. In the rare cases where a subject sent us a second email of interest in the position (i.e. after receiving our email describing the salary regime) without filling out the questionnaire, we emailed them a reminder that their application is not complete until we receive the questionnaire.

[15] We use the SSA database to calculate, for each name, a weighted gender probability. The database reports figures on the most common 1,000 names for men and the most common 1,000 names for women born in any given year. We take a given name, proceed to use the number of men and women born each year with that name, and then create a gender ratio for that name in each given year. We then look across multiple years to create a weighted average of this gender ratio. We first restrict our range to birth years 1944-1993, so as to focus on the most likely birth years of our subjects. Then, since information in the CVs suggests the majority of our subjects are between 22 and 30 years old, we assign lower weights to birth years earlier than 1979 and later than 1987. The weights diminish as birth-years move further from those of the 22-30 year-old bracket. We do this for each name that appears in our sample. If the resulting first name-based weighted probability of being female is larger than 50%, we define the subject as a woman. Otherwise, we define the subject as a man. Note that there are very few names for which gender cannot be inferred accurately. For example, in our sample of job-seekers only 0.8% of the subjects have names with gender ratios less than 2:1, and only 4.7% have a ratio less than 10:1.

[16] For names included in both databases, the Geoff Peters database (available at http://www.gpeters.com/names/baby-names.php) generates gender ratios very close to the ratios generated by the SSA database ($r=0.89$, $p<0.0001$, Pearson).

derived from the SSA database is too low to confidently assign one gender or the other. Finally, for all names where neither database yields a large enough gender ratio to make a confident assignment, we perform internet searches for gender identifiers of the actual subjects themselves, e.g., by finding the subjects on social networking websites.

The remaining individual characteristics of interest were gathered from the resumes sent to us by the subjects. These include level of education, job experience (i.e., whether job-seeker has already worked as administrative assistant), and age.

3. Do Competitive Workplaces Create a Gender Gap in Applications?

3.1 Descriptive overview

In total, we collected data from 8,969 individuals who responded to one of the two job advertisements. The gender from most individuals (83.5%) was assigned using the SSA database. The Geoff Peters database was used to assign gender to 8.9%, and the internet search allowed us to assign gender to 7.6%.[17] Table 2 provides detailed information on the number of observations in each cell of our two natural field experiments and application probabilities for selected categories. Table 3 provides information about the sixteen cities where the advertisements were placed, including the number of observations and application probabilities in each location. 4,239 individuals responded to the female ad, 4,730 responded to the male ad (2,540 in the first and 2,190 in the second natural field experiment). Of these interested job-seekers, 3,525 individuals ultimately applied for one of our jobs – 1,566 for the female version and 1,959 for the male version (1,136 in the first experiment, 823 in the second). This represents 39.3% of the job seekers in our sample (30%–48% of the subjects in each city).

As expected, the more general administrative assistant ad attracts a substantially higher proportion of women than men, consistent with the notion that it is largely a "female" job. For this ad, 80.3% of the job-seekers are women.[18] By contrast, for the version which adds a

[17] The application patterns reported in this paper are not subject to the inclusion of the gender identifiers from the Geoff Peters database or the gender identifiers from the internet search. We excluded from the analysis individuals for whom we could not identify gender. There were also 241 individuals who responded to both job advertisements in the first natural field experiment. For these individuals, we only consider their application decision with respect to the job advertisement to which they responded first so that each job-seeker constitutes exactly one observation. Excluding these individuals from the analysis has no impact on our findings. Further, we had to exclude 48 individuals due to technical problems during the e-mail exchange (e.g. we could not reply to some job-seekers as their email accounts had been closed).

[18] This is very close to the percentage of office support jobs actually occupied by women – 79% in 2001, according to the Bureau of Labor Statistics. The close parallel between expressed interest and realized outcomes suggests a

"masculine" flavor, the gender distribution of interested individuals is roughly equal: 53.5% of the first experiment's job-seekers who expressed interest in the male ad are women. (The difference in gender composition of interested job-seekers across the two ads is significant at the $p<.01$ level using a Fisher's exact test).

The median age of our subject pool is 26 years. In terms of education and job experience, we observe that 44.9% of the applicants have a bachelor's degree and 57.4% have had at least some college. 70.5% of applicants has had at least some work experience as an administrative assistant. The characteristics of our sample differ from the average administrative assistant employed in the US, who tends to be older and less educated. However, since we do not sample from people currently working as administrative assistants, but instead from job-seekers interested in such a position, this difference appears natural.

The length of the application questionnaire does not significantly affect application probabilities: 40.6% of the job-seekers apply if the questionnaire is short and 39% if the questionnaire is long (Fisher's exact test, $p=0.188$). Yet, the questionnaire length does appear to significantly affect application costs. Applicants respond to the long questionnaire with roughly three times the writing output as that for the short questionnaire (long questionnaire applicants use 338 words on average, while short questionnaire applicants use an average of 117 words , T-test, $p<0.0001$).

3.2 **Application patterns across the compensation treatments**

Figure 2 illustrates the mean overall application probabilities for women and men for each of the six compensation treatments in the first natural field experiment. We first note that, conditional on expressing initial interest in the job, men are in general more likely to apply: the overall application probability for men is 0.484, significantly higher than for women, who apply with a 0.362 probability (Fisher exact test, $p<0.0001$). Second, note that the difference between the application probabilities of men and women is most pronounced in the high-competition treatment T4, where men are 55.5% more likely to apply than women (0.485 for men versus

strong relationship between gender distributions in the actual workforce and gender distributions among applicants. This underscores the importance of the effects that different payment-regimes can have on the composition of job applicants, since the composition of applicants may strongly influence the realized composition of workers in the labor force.

0.312 for women, Fisher exact test, $p<0.0001$). In the low-competition treatment T3, men are 36.4% more likely to apply than women.

In the team competition treatment, T5, men's application probability is 31.9% higher, in the uncertainty treatment, T6, it is 32.5% higher. In the two baseline treatments T1 (team-fixed) and T2 (individual-fixed), the probability that males apply is 23.1% and 26.5% higher than females' probabilities. Appendix B, Figure A, illustrates the application patterns for each city separately. The data patterns lead to our first result:

Result 1: *Competitive workplaces can significantly increase the gender gap in application probabilities: For the position which advertises pay that depends heavily on relative performance, women's propensity to apply substantially drops relative to that of men.*

For further evidence of this result, we consider a Logit regression model in which we regress the application decision on all possible interactions between treatments × gender. Table 4 reports the results. The regression models 1-3 control for location fixed effects and the coefficients reported are marginal effects. Model 1 shows the treatment effects for men, model 2 for women, and model 3 shows the gender difference men minus women, i.e. the treatment × male interactions. The reference groups for the regressions are T1 and T2. The models 4-6 are identical to models 1-3 but control in addition for education and treatments × education interactions. We pool the two baseline treatments, as the application probabilities and patterns in T1 and T2 are very similar, suggesting these two compensation schemes were perceived as virtually the same.[19] In Appendix Table A, we provide regression analyses showing that our findings are similar if we do not pool data from treatments T1 and T2 or use T6 as the comparison group. Appendix Table B, includes a robustness check where we control for subjects' education, job experience, their interactions with the treatments, and other potential covariates, none of which affect the application patterns.

Table 4 indicates whether the gender gap in application probabilities in treatments T3 through T6 significantly differs from the gender gap in the two fixed-wage baseline treatments under several different specifications using the full sample from both jobs. We find in models 3 and 6 that the coefficient on the interaction T4 × male (variable T4) is positive, and significant at

[19] Recall that the compensation structure is in fact identical across these two treatments. The only difference between T1 and T2 was in the framing of the job as individual or teamwork-based.

the p=0.064 (model 3) and p=0.046 (model 6). That is, as we hold everything but the payment regime constant, the amount by which men are more likely to apply than women substantially *increases* as we move from the baseline to compensation that depends heavily on individual relative performance. Such contracts make women significantly less likely than men to apply. While the average marginal effect of being male on propensity to apply in the fixed-wage treatments T1&T2 is 0.096-0.106, the effect of being male under T4 increases by 0.07-0.08. That is, T4 widens the gender gap in propensity to apply by 66% (model 3) – 83% (model 6).

While the estimated magnitude of the effect is substantial, it is instructive to consider for a moment the precision of this estimate in the context of the literature on gender and competition. Studies on choices in the lab over compensation for solving mazes or math problems typically have much smaller samples, often less than 100 subjects. Yet they generally find results on preferred pay scheme that are more precise than our estimate of the influence of relative performance pay on application choices of job-seekers. Selecting into a job differs in many ways from choosing the preferred pay regime for a task performed in a lab, and it is possible the myriad other factors affecting job choice introduces greater variance in the errors. However, there are two other important sources of lower precision in this initial estimate, both of which shed new light on the underpinnings of the gender-competition relationship. While we discuss both later in the paper, they deserve brief mention here.

First, Result 1 is based on the pooled data. In comparison to previous studies of competition, which generally use a single undifferentiated task, our data is generated using a job split into two slightly different versions which differ in their gender-connotations. The overall effect we detect in Result 1 is across both versions of the job. As we discuss further below, the effect of T4 on the gender gap in job interest is sensitive to these adjustments in the position. Pooling both samples in the same regression thus lowers the precision (and magnitude) of estimated effects. Second, our sample is more diverse than standard lab samples in several ways that appear to be important – most prominently in terms of age. If the gender gap is driven primarily by younger subjects (e.g. those in their late teens and early twenties, the standard age-range in lab experiments), then the broader spectrum of ages in our sample will weaken precision and lower the magnitude of estimates compared to what is typically reported in the literature. We examine age-dependency further in section 7.2, and discuss the role of task-dependency in detail in section 5.

4. Are Men Attracted to Competitive Workplaces?

Existing laboratory studies suggests gender differences in choices over competition are driven partly by men preferring competitive to non-competitive settings. We might therefore expect men in the field to be more attracted by competitive workplaces than non-competitive ones. However, we find that:

Result 2: *The observed gap is not driven by men being attracted to the competitive work environment and women shunning it, but rather by a significantly stronger aversion to competitive workplaces among women compared to men.*

Figure 2 suggests that women and men are *both* more interested in non-competitive workplaces than competitive ones, as application probabilities drop for both genders when moving from fixed-wage compensation (T1 & T2) to compensation depending on relative performance (T3 & T4). The estimated coefficients for the treatment dummies T3–T6 in Table 4, models 1 and 2 show the changes in men's and women's application probabilities in each of the four treatments T3–T6, as compared to the baseline treatments. Several insights follow.

First, women are significantly less likely to apply for jobs in which compensation is not fixed. This is shown in Table 4, model 2, where all coefficients for the four treatment dummies are significantly negative and range in magnitude from -0.068 (T6) to -0.127 (T5). In particular, the probability that women apply in T4 is 0.117 lower as compared to treatments T1 and T2, which represents a 27.3% drop in application probability in moving from fixed-wage to high-competition.

If we focus on men in treatments T1 and T2 as the reference group for the regressions (models 1 and 4), we see that men's application probabilities also decrease in treatments T3-T6 as compared to the baseline T1&T2, yet the reductions are smaller than for women. In particular, for T4, the estimated average marginal effect suggests an average drop of only 8.8% in application probability from the baseline (-0.047, p=0.130).

We thus observe on the one hand that contracts characterized by relative performance-based pay tend to have a dissuasive effect on *both* genders, while on the other hand that the effect is substantially stronger for women.[20] The gender gap induced by competition-based contracts is

[20] As the effect for men is not significant at conventional levels, its negative sign should be interpreted with caution. However, the effect on men is clearly not positive, in contrast to robust findings on compensation choices in the lab.

thus driven by the fact that women are *more* deterred from applying by competitive compensation regimes than are men. This finding is important. It suggests a more nuanced mechanism may drive responses in the field to workplace compensation regimes than that suggested by the result that has emerged from the literature in lab experiments, in which men are observed to be significantly *attracted* to competition-based pay while women shy away from it. Note that this may derive not only from differences in the choice environment between the field setting and previous experimental settings, but also from the broader distribution of ages among our field participants compared to the standard subject pools in the literature – a hypothesis we further explore in section 8.2.[21]

5. Determinants of the Gender Gap

In this section, we qualify our main result. This analysis helps us understand the determinants of the deviations in application probabilities as well as the factors that attenuate and exacerbate them. A first result emerges:

Result 3: *Several factors related to the workplace affect the gender gap, including: i) the degree to which compensation depends on relative performance, ii) whether the job is team based, and iii) minor changes in the job task.*

We observe a substantial attenuation of the effect of competitive workplaces on the gender gap in applications as we lessen the intensity of the competition incentive from 50% (treatment T4) of the base wage to 22% (treatment T3). In Table 4, model 3 (and 6), we see that the estimated coefficient on the interaction T3 × male drops to approximately one-third of the magnitude of the coefficient on T4 × male; although it remains positive, it becomes insignificant (average marginal effect of 0.022, p=0.555). That is, ceteris paribus, changing the contract from a fixed wage to compensation *mildly* dependent on individual relative performance has no significant impact on the gender gap in application probabilities.

We also find that adjusting the relative performance contract to make it team-based reduces the gender gap considerably. Table 4, model 3 (and 6) shows that the coefficient on the

[21] In section 8.2, where we examine non-experimental data on the role of age on the gender gap, we find evidence that *younger* males may in fact be attracted by the relative performance pay scheme (rather than simply being less dissuaded). If we restrict the sample to males under the age of 26, the sign of estimates is consistent with previous findings (based typically on young adults) that men tend to be competition loving (see Table 10). The magnitudes are, however, considerably smaller than observed in previous studies and the estimates are not statistically significant.

interaction T5 × male is small in magnitude and insignificant (p=0.937). Recall that under T5, the bonus raises the wage from $12 to $18, just as for T4, but that in T5 workers are evaluated on the joint performance of their team. Thus, the gender gap in application probabilities remains constant when moving from a fixed wage (T1 & T2) to a wage heavily dependent on *team* relative performance (T5).

New evidence that is beginning to emerge from lab studies suggests the different responses by men and women to competition may depend partly on the tasks performed (Grosse and Riener, 2012; Shurchkov, 2012). To examine whether and how changes in aspects of the job that have nothing to do with the contract affect the gender gap in job-interest caused by relative performance pay, we investigate gender differences in application probabilities for each of the two versions of the position separately. We start with the sports job advertisement. Figure 3 replicates Figure 2 but restricts the data summary to the sports job advertisement. As before, we observe that among individuals expressing initial interest in the job, men are more likely to ultimately apply than women. This is true across all six treatments, the mean application probability being 0.534 for men and 0.372 for women, a significant difference at the $p < .01$ level using a Fisher exact test. Yet, note the striking difference of probabilities in T4, compared to the other five treatments. The gender gap is fairly similar across T1, T2, T3, T5, and T6 (0.109–0.172), where the probability that men apply ranges from 32.8% to 50.8% higher than that for women. It is considerably more pronounced in T4 (0.281; application probability women = 0.298, men = 0.579), where the probability that men apply is 94% higher than it is for women. Table 5 is identical to Table 4, with the exception that it is restricted to responses for the sports job advertisement. The treatment dummies show that women are significantly less likely to apply in T3-T6 as compared to the baseline (model 2), and the male dummy in model 3 indicates that the application probability is 0.124 higher for men in the baseline. As in Table 3 models 3 and 6, all treatment × gender interactions are insignificant besides T4 × male. Note that now, however, the coefficient on T4 × male approximately *doubles* in magnitude and substantially increases in significance. The difference-in-difference in application probabilities in moving from the fixed hourly wages to the high-stakes individual competition treatment is now 0.153 in model 3 (0.140 in model 6), and significant at the $p<.01$ level in model 3 ($p=0.022$ in model 6). The difference-in-difference in application probabilities between T3 × gender (low bonus) and T4 × gender (high bonus) is significant at the $p<.05$ level and the difference between T4 × gender and T5 ×

gender is significant at the p<.02 level. Importantly, this result shows that when we restrict the sample to the sports job-task and hold everything but the payment regime constant, the gender gap in application probabilities observed in the baseline T1 and T2 increases by *approximately 125%* when compensation depends heavily on individual relative performance.

Figure 4 is identical to Figure 3, but now shows only the application probabilities for the general job advertisement. We observe again that, conditional on expressing initial interest in the job, men are generally more likely to apply, but that the difference in overall application rates is much smaller than for the male ad (application probability of men= 0.414, application probability of women= 0.358; Fisher exact test, p=0.003). We also see that the gender gap in application probabilities is quite similar across *all* six treatments. The smallest gender difference is in T4 (0.021) and the largest in T2 (0.072). In none of the six treatments are women significantly less likely to apply than men (p>0.107).

Table 6 is identical to Tables 4 and 5, with the exception that it is restricted to responses to the general job advertisement. The results are similar to Tables 4 and 5, with one crucial exception. As in Tables 4 and 5, the coefficients on the treatment dummies show that women are less likely to apply in treatments T3-T6 as compared to the baseline (models 2 and 5). However, turning to the treatment × gender interactions (models 3 and 6), we now see that *none* of the coefficients are significant (p>0.525). That is, none of the treatments significantly alter the gender gap in propensity to apply among job-seekers interested in the general ad position. The T4 × male interaction is now small and statistically insignificant at conventional levels (p>0.555): the gender gap in application probabilities caused by the high-stakes competition incentive vanishes when the job task does not have masculine connotations.

A closer look at the data reveals that the disappearance of the gender gap in T4 as we move from the sports ad to the general ad is driven mostly by a large drop in the fraction of men who apply. Among men, the application probability remains constant when moving from the baseline to T4 under the sports ad (estimated change=0.006, p=0.879, model 1), but it significantly drops (by 0.139; p=0.005, model 1) when moving from the baseline to T4 under the general advertisement. Among women, the application probability decreases by 0.147 when moving from the baseline to treatment T4 under the male ad, but drops by only 0.104 when moving from the baseline to treatment T4 under the female ad (models 1). That is, while men are unaffected

and women severely dissuaded by treatment T4 under the sports ad, both genders are dissuaded by treatment T4 from applying under the general advertisement.

Thus, we find that when switching the framing of the job from sports- to general job duties, men switch from being indifferent to high-stakes individual competition to being deterred by it, while women remain similarly deterred by it, causing the gender gap to disappear (see also Table 7). One potential explanation for this finding is that subjects have expectations about the gender of their competitor (who they may believe is likely to be male in the sports job, female in the general job) and different preferences for competing against members of the opposite gender than competing against members of the same gender. For example, men may enjoy competing against men, but prefer not to compete against women, whereas women may dislike competing against men more than they dislike competing against women. Alternatively, women may believe they have better chances to win the bonus when they are competing against women, and are thus less dissuaded by high-stakes competition in the general job, if they believe they will compete against a woman. These possibilities are addressed in greater detail in section 6.

Turning our attention to treatment T6, we find that when compensation depends on uncertain elements, but without competition, the effect on application probabilities is also negative. In Table 4, models 1 and 4 for men, and models 2 and 5 for women show that the estimated coefficients on the dummy for T6 are clearly negative – e.g. -0.056 for men (model 1) and -0.068 for women (model 2). While the difference leads to a mild increase in the gender gap, the coefficient on the gender-treatment interaction T6 × male is insignificant ($p>0.726$).

Note that the coefficients for T6 and the T6 × male interaction in Table 4 are similar to the coefficients for T3 and the T3 × male interaction. Recall that the bonus in both treatments is the same, and that the only difference between T3 and T6 is that the uncertainty in pay for T3 stems from competition. This similarity in estimates suggests uncertainty may play an important role in job-seekers' responses to relative performance-based pay. Since T6 was structured to mirror the bonus of T3, comparing the findings on uncertainty with the findings from T4 is less insightful. However, we return to this issue in section 6, where direct comparison of the effects of high-stakes competition (T4) and uncertainty is made possible by the treatments in the second experiment.

6. Drivers of the Effects of Competitive Workplaces

Thus far, we have shown that exogenously introducing relative performance based rewards in the workplace can have a substantial impact on the gender composition of workers interested in a position. We have also shown that the impact is sensitive to non-remunerative aspects of job characteristics (in looking at the general version of the job compared to the version which involves sports), as well as the structure and stakes of the competition incentives (placing relative performance in a team context or by lowering the stakes of success and failure). We are aware of no other study which has so conclusively shown the effects of these types of employment contracts on worker sorting in actual labor markets. We now turn to a closer examination of possible explanations for *why* competitive workplaces can have this effect.

6.1. Beliefs About Ability

One possibility is that the gender gap in applications caused by competition-based pay results from beliefs about differences in ability. Suppose that it is believed that men are generally better at administrative assistant duties when communication in the position revolves around sports (the version of the job in which T4 induces a large gender gap). This might cause competition-based pay to create a gender gap in job interest if the average woman believes there is a high chance her coworker will be male and thus likely to be better at the job, and therefore revises her subjective expected wage downward such that it falls below the wage of the fixed wage environment.

The experimental design of our second natural field experiment allows us to explicitly test this hypothesis by randomly assigning the gender of the coworker. If there is indeed an expected gender difference in ability for the job-tasks in the sports-related version of the job, then by revealing that the coworker is female (male), we alter the expected ability of one's competitor. If women believe that men are better at the sports-related job and thus have a higher chance to win the tournament, then we should observe that women are more likely to enter into competition against female than male coworkers.[22] Yet, in the second experiment, we do not find that the gender of the co-worker affects women's application probabilities under T4 (see Figure 5). Table

[22] There are two implicit assumptions in this argument. First, it is assumed that job-seekers do not doubt the information the employer provides; e.g., that female applicants believe that the employer is committing to hire a female coworker in treatment 2T4b. Second, it is assumed that the distribution of abilities is similar in terms of variance and skewness. For example, if there only existed a small group of highly able women, most women should still shy away from competing against a female coworker in the sports-related version of the job.

8, model 2 reports estimated effects of coworker gender on women's probability of applying. We see that neither knowing her coworker is male nor knowing her coworker is female has a significant effect on her propensity to apply. More importantly, the difference between the two estimates is not significant (p=0.599), and moreover is in the opposite direction than we would expect if women avoid the job under T4 – application rates are slightly higher when their coworker-competitor is male than when female.

Alternatively, suppose that only men (but not women) believe they are more skilled than the average woman at administrative assistant duties in a position that deals with sports, but do not believe males are any better at these duties in the general version of the position. Then they would expect a higher wage under T4 in the sports-related than the general version of the job, as long as they believe there is a positive probability their coworker could be a woman. This would cause T4 to attract them in the sports-related, but not the general, version of the job. However, this interpretation is also not consistent with our results from the second natural field experiment. In Table 8, column 1, it is clear that for men, the gender of their coworker-competitor does not have a significant effect on their response to T4. These findings are summarized in our fourth main experimental result:

Result 4: *The gender of the coworker-competitor in the high-stakes relative performance pay scheme has no measurable effect on the gender gap.*

If the gender gap induced by T4 in the sports-related version of the job were driven by beliefs (among men, women, or both) about differences in ability between men and women, we would expect that revealing the gender of the coworker would have an effect on the gender gap. However, we see no evidence for this in the second experiment.

A closer look at the data patterns from the first natural field experiment yields additional evidence helpful in assessing the hypothesis that beliefs about ability are driving the results. First, if behavior is being driven by the fact that women's expected wages under T4 are lower in the sports-related version of the position (where the gender gap is created) than in the more general version (where the gap disappears), we would expect T4 to have a dissuasive effect on female applications only in the sports-related version of the job. However, we see a *general aversion* to competitive workplaces among women, regardless of which position it is: Their application probabilities drop by 10.3 percentage points from T1/2 to T4 ($p<0.001$) in the general

job ad (Table 6, model 2). This drop is similar to the corresponding drop in the sports job ad (14.7 percentage points, Table 5, model 2). Moreover, we would expect that holding T4 constant and switching from the sports-related version to the other version of the job would result in a drop in application probabilities by women. However, Table 7, models 1 and 2 show that women's selection out of T4 is not significantly different across the two versions of the job ($p>0.433$). The fact that women do not shy away from the competitive environment more in the sports version than in the general version of the job suggests there is no common belief that women are likely to be less skilled at the job and therefore expect a lower wage. For these reasons, we do not find beliefs about gender differences in ability to perform the job a compelling explanation.

Perhaps this should not be a surprise. The sports version of the job was not designed to confer an actual advantage to men over women, but rather to manipulate gender associations with the position. If it is true that there is a stereotype that women are generally better than men at communication and tasks based on verbal skills, narrowing the subject matter around a position with predominantly communication-oriented job functions to the area of sports seems unlikely to have strong effects on beliefs over how good each gender will be at the job.

6.2 Expectations of Gender Discrimination

Let us next consider the hypothesis that workers are responding to beliefs about sex-based discrimination in performance evaluation. For example, workers may believe that the competition-based pay regime creates an opportunity for superiors to selectively discriminate. This would affect subjective expected wages under relative pay performance much in the same way as beliefs over relative ability: a worker who believes the evaluation of her performance will be biased downward while that of her coworker will not be subject to this bias, will believe she is less likely to earn the bonus and revise her expected wage downward accordingly. Thus, the gender gap in job-interest caused by T4 might be due to a belief among women that their performance will not be fairly judged by their bosses.

Our second natural field experiment allows us to test this question in several ways. First, we manipulate the gender of the worker's superior, who evaluates relative worker performance, to test for any evidence that women (men) believe they will be discriminated against (in favor of) by superiors of a specific gender. In Model 2 of Table 8, we see the estimated effects of supervisor gender on a female job-seeker's propensity to apply under T4 are not significant,

while Model 1 of Table 8 shows supervisor gender also has little effect on applications by men. The application probabilities are also illustrated in Figure 5. These findings inform our next result:

Result 5: *The gender of the supervisor in the high-stakes relative performance pay scheme has no significant effect on the gender gap.*

This suggests that women are not responding to a concern of being judged unfairly by a male (or female) boss and thereby ending up with a lower wage and that men are not responding to expectations of favoritism by a male (or female) boss and thereby expecting a higher wage.

It is possible job-seekers believe discrimination would be applied by both male and female superiors in the same way. If so, the lack of effects of supervisor gender might simply imply that job-seekers are reacting to expectations of discrimination that are independent of the gender of their superior. However, if women believe assessments of their performance will be biased downwards compared to those of men (regardless of the gender of their boss), their expected wage should increase when their coworker switches from male to female. As already mentioned, in the regression results reported in Table 8, we see no measurable increase in a woman's probability to apply under T4 when her coworker switches from being male (2T4a) to female (2T4b). By the same logic, expectations of pro-male discrimination should increase men's expected wage when their competitor is a woman, but there is no discernible increase in men's willingness to apply when their coworker switches from male to female.

As a final piece of evidence, we note once again that in the first natural field experiment women are *generally* deterred by T4. Since women are deterred by T4 at similar rates across the two versions of the job, it does not appear that the gap induced by T4 in the sports-version of the job is caused by women fearing that their performance will be unfairly judged by their superiors in comparison to that of men. Nor does it appear that it is caused by men believing they will be the beneficiaries of sex-based discrimination.

6.3. Task-Dependency: Job Task vs. Gender Associations

Recent findings suggest that men in laboratory experiments may not be more competitive than women across all tasks, but that it depends on the type of task being performed. Shurchkov (2012) replicates the standard sharp gender difference first found in Niederle and Vesterlund (2007) with a simple computational task, and then shows that when switching to a very different

task based on verbal skills, men are no longer more willing to compete than women. She posits gender-task stereotypes/associations as an explanation for this new finding. Grosse and Riener (2012) find similar results in a lab experiment that uses a verbal task. The extent to which this pattern predicts behavior in the field has significant implications for the effects of different performance incentives on gender composition of workers. One interpretation, for example, would be that compensation based on relative performance may induce a gender gap in job-interest for jobs requiring computational skills, but not for positions based on verbal skills.

One challenge in interpreting these nascent patterns on task dependency in the lab is that the experimental designs manipulate two factors simultaneously. In addition to changing the potential gender connotations of the task, these designs also fundamentally alter the nature of the task and the skills used in performing it. It may indeed be subjective gender associations that individuals attach to the task which matters. However, on the other hand, it is possible that men and women have different preferences over the tasks themselves, and combinations between task and pay regime, and this may have little to do with gender stereotypes. Whether it is the task itself, or gender connotations surrounding the task, is important to uncover whether a fundamental change in the work performed is necessary to affect the gender difference in competitiveness.

Turning to the behavior of individuals on labor markets, this suggests several key questions. The first is whether the tasks performed in a job affect the extent to which competition-based pay can create a gender difference in job-interest. The second is, if so, whether it is necessary to switch between tasks that require fundamentally different skillsets to turn the gender-competition effect on or off. A third is whether it is job functions themselves, or gender connotations linked to those functions, which play the pivotal role.

A key feature of our experimental design is that the job we created was modified slightly in one case so as to create two different versions within the same employment position – one without any overt gender connotations, and another with connotations that tend to be more masculine. This enables us to test several things at once. *First*, it allows us to examine whether changes in the job besides the pay-regime can affect the gender difference induced by competition-based pay. *Second*, we can test whether the effect of competition incentives on gender composition is sensitive only to switches across fundamentally different jobs, or if instead minor adjustments within the same position are sufficient to turn the effect on or off. *Third*, we

are able to examine more closely whether it is the general job function itself that matters or gender connotations surrounding the work.

As discussed in Result 3, we find a very large difference in the effect of T4 across the two versions of the job. We conclude on the one hand that aspects of the job which have nothing to do with the contract can be critical in causing relative performance based pay to create a gender gap in job interest. On the other hand, it does not appear that it is the functional duties themselves (proofreading, correspondence, writing reports, sending emails, answering phones, etc.) which are pivotal. A mild adjustment within the same position such that it is surrounded by the topic of sports was enough to trigger the effect of competition incentives in causing a gender imbalance, suggesting that *gender connotations* linked with a position are important.

6.4. Tastes for Uncertainty and Competition

There is evidence that men and women respond differently to risk and uncertainty (e.g. Borghans et al, 2009; for a review see Croson and Gneezy, 2009), which may explain why women sometimes shy away from competitions. In their influential study, Niederle and Vesterlund (2007) argue for the existence of systematically different preferences between men and women over competition per se, independent from preferences over risks and other factors. However, as the authors point out, their controls are imperfect and the residual gender gap may still be driven by a gender difference in factors such as risk aversion, rather than by an independent taste for competition.

Our second natural field experiment sheds light on whether job-seekers in the field are as sensitive to uncertainty as to competition per se. In Table 8, we see that application probabilities of men (model 1), women (model 2), and the gender-difference (model 3), are not significantly different under large wage uncertainty ($6/hour) than they are under high-stakes competition. (The coefficient in model 3, for example is just .027, p=0.71.) This leads to our next result:

Result 6: *Large wage uncertainty has almost the same effect on the gender gap in application rates as the high-stakes relative performance scheme.*

One way to interpret this finding is that it suggests preferences over uncertainty can potentially explain choices over competition across men and women; another possibility is that uncertainty and competition have independent effects of similar magnitude. We are not aware of any study investigating the importance of gender differences in uncertainty on the gender gap in

competitiveness in a natural field setting involving large stakes. In this light, our finding here suggests this as a potentially fruitful area for further research.

Another possibility for the observed gender gaps in competitiveness is that women like cooperative workplaces with egalitarian rewards more than men (Kuhn and Villeval, 2011). However, in our first experiment, where the only difference between T1 and T2 is that in T2 work is done in teams, women's application probabilities are almost identical (T1: 42.6%, T2: 43.1%), and there is also almost no difference for men (T1: 52.5%, T2: 54.5%). We see no strong evidence that women prefer to work in teams more than men in our job setting. This makes us hesitant with regard to the hypothesis that female preferences for cooperative work environments explain the pattern that we observe.

We conclude that gender differences in preferences over uncertainty, and possibly also over competition per se, are the most plausible explanations underlying the effects we find of competition incentives in the workplace on gender composition applicants.

7. Heterogeneous Treatment Effects

This section analyzes the relationship between the compensation regime's effect on propensity to apply and three additional variables: *(i)* the economic environment, *(ii)* the age of job-seekers, and *(iii)* the qualifications of job-seekers.

7.1 Local market conditions, relative-performance pay, and the gender gap

By placing our employment advertisements in multiple cities, we are able to explore the link between gender-dependent responses to different contract-types and local labor market conditions. To obtain precise estimates of prevailing local wages in each of the sixteen different labor markets from which our subjects are sampled, we collected data on the median hourly wages from comparable administrative assistant jobs posted on the same online job board, for positions which offered fixed-wage compensation. Median wages range from $10 -$14 per hour.[23] For gender differences in wages across cities we use Bureau of Labor Statistics 2009 data.

[23] While lower than the most recent Bureau of Labor Statistics data on *mean* compensation for administrative office jobs (ranging in the same cities from $14.99-$20.03 in May 2009), the two data sets for wages of jobs comparable to the one we advertised are highly correlated ($r=0.914$, $p<0.0001$).

When women have fewer (or less attractive) outside options to work in a job with fixed wages, i.e. when market wages are comparatively low and there are larger gender differences in wages, they may be less reluctant to enter competitive workplaces – even if they dislike competing. Figure 6 compares median wage levels to the size of the effect of T4 on the gender gap in application probabilities in each city (estimated from the specifications in Table 4, model 6, run separately for each city). The pattern suggests a positive relationship between the cities' hourly market wage and the T4 × male interactions. In particular, the three cities with the highest wage levels in our sample (San Francisco, Boston, and Washington DC) are also the cities with highly positive estimates of the T4 × male interactions. In these cities median wages are at least $13, and they are the closest to the mean wage that we offer of $15. Overall there is a positive correlation between market wages and T4 × male interactions ($r = 0.618$, $p=0.011$, Pearson).

To more closely examine the extent to which the application patterns are subject to market wages, we divide the sample into one group consisting of the three cities where wages are most comparable to our offered wage, and another group where local wages are considerably lower. Appendix Table C reports results from the regression using the first group. Among these cities the gender gap caused by T4 across both job types is more than three times as large as the gap caused by T4 across all sixteen cities taken as a whole (the estimated marginal effect rises from 0.080 to 0.266, significant at $p<0.01$, in model 1). When we restrict our focus to the sports-related version of the job (model 2 of Table C), we observe that the estimate for the T4 × male interaction is roughly twice as large as it is when all sixteen cities are combined (a marginal effect of 0.315, compared to 0.153). In appendix Table D, which reports results from the second group of cities, we see that the estimated effect of T4 on the gender gap across both job types is no longer significant ($p=0.438$, model 1). It is still significant for the sports advertisement sample, though its magnitude is roughly one third of that for the group of cities with higher local wages (0.119, $p=0.066$, model 2).

Further evidence for the role of market wages is presented in Appendix Table 9. This table presents three models where we control for the triple interactions treatment × gender × market wage (models 1 and 2) and treatment × gender × gender difference in market wage (model 3). Model 1 uses a binary wage variable indicating whether the market wage is lower than $13. Model 2 uses a continuous wage variable and model 3 a continuous variable for gender

differences in market wages. Models 1 and 2 corroborate the previous analysis in showing that the T4 × male × market wage interaction is significant at p=0.022 in model 1 and p=0.088 in model 2. That is, we observe that as local wages for similar jobs rise towards our offered wage, the extent to which women are disproportionately deterred from applying to the highly competitive job increases. Model 3 shows that the T4 × male × gender difference in market wage interaction is negative and significant at p=0.075, which suggests that the gender gap in applications to competitive workplaces is smaller in cities where the gender gap in wages is larger. Finally, Appendix Table E further investigates the role of local labor market conditions by controlling for different city-covariates and their interactions with treatment and gender. We observe that the T4 × male findings (Result 3) are robust in all five regression models.

7.2 The role of age on the gender-gap in responses to relative-performance pay

Two additional sources of variation in our study come from the individual characteristics of the job-seekers. First, individuals who responded to our job advertisements are of varying ages (mean = 27.9 years, approximately two-thirds are 22-30 years, min. = 18, max. = 68).[24] Such a broad range of age-diversity is unusual in studies of gender differences in competitiveness. With the exception of a few studies which analyze the behavior of children (e.g. Sutter and Ruetzler, 2010, Dreber et al, 2011), much of the research in this area has focused on the behavior of university students. The only studies that we are aware of that include substantial age variation among adults are Charness and Villeval (2009), and Flory et al. (2012), both of which use experiments in the lab. While the former does not examine the relationship of the gender gap with age, the latter finds evidence that gender differences in willingness to compete close as adults grow older.

If it is true that age has a negative effect on gender differences in competitiveness then one might expect that the gender gap in applications for competitive workplaces is less pronounced as the age of job-seekers increases. Figure 7 shows application probabilities (linearly fitted values) under T4 by job-seeker age. Propensity to apply significantly decreases among both genders as age goes up. In fact, younger men's application probabilities (up to approximately 25 years of age) are higher in T4 than in T1/T2 – an observation that is consistent with previous

[24] We derived age from information in CVs such as graduation year. We could infer the age only from 4620 of our 6779 subjects in the first experiment.

conclusions from the literature which suggest that men are competition-loving. However, older men are clearly more attracted to a workplace which offers a fixed wage. Confidence intervals for men and women start intersecting at approximately 32 years, indicating that the gender gap in competitive workplaces is more pronounced in younger job-seekers.

Table 10 cuts the sample of job-seekers in two roughly equally sized age categories (25 and younger, and over 25). Models 1-3 present the estimates for the younger sample for women (model 1), men (model 2), and the gender difference (model 3). Models 4-6 present the corresponding estimates for the older sample. The coefficient on T4 for men switches signs when moving from younger to older job-seekers (coefficient 0.037 in model 1, -0.092 in model 2) whereas it is always negative for women (coefficient -0.08 in model 2, -0.119 in model 4). The T4 × male interactions are more than four times larger in the younger job-seeker pool (coefficient=0.117, model 3) than in the older job-seeker pool (coefficient=0.027, model 6).

7.3 Competition-incentives, job qualifications and adverse selection

Individuals in the sample also vary with respect to measures of human capital and qualifications for the job (44.9% of the applicants have a bachelor's degree, 70.5% have work experience as an administrative assistant) and this variance allows us to test whether competition incentives cause adverse selection. Niederle and Vesterlund (2007), for example, find that even high-ability women (those likely to win a tournament) select out of competitive environments in the lab, while low-ability men select into them. This raises a serious concern with regard to attracting the best talent and efficiently allocating human capital. If competition-incentives deter talented individuals or alter applicant pools such that it lowers abilities, this could lead to substantial productivity losses and important costs to society. We examine evidence on effects of this type in a labor market by analyzing job-qualifications across the different pools of applicants.

We first consider the net effect on qualifications of applicants as a whole. Comparing education and job experience across those who choose to apply under each treatment, we find no statistically significant differences. Neither mean job experience (Kruskal-Wallis equality of populations rank test, p=0.654) nor mean education (p=0.263) are statistically different across the six groups of applicants in T1-T6. That is, we do not find evidence that any of the treatments lowers indicators of worker-ability in the applicant pool.

To test whether the compensation schemes affect the qualifications of male and female applicants differently, we regress education and job experience of those who chose to apply on treatments, gender, and the treatment × gender interactions. The results are reported in Appendix Tables F (general ad) and G (sports ad). With the exception of marginal significance for T3 and T6 for job experience in Table F, model 3, none of the 21 other coefficients on the interactions are significant, suggesting that the effect of each treatment on the qualifications of applicants does not differ by gender. In particular, competition incentives do not affect job-qualifications of female applicants any differently than males.

Finally, to examine whether the high-stakes competition regime deters highly-qualified women, we turn our attention to the coefficient for T4 in models 2 and 4 of Tables F and G. Recall that T4 reduces the percentage of women in general that apply. The coefficients in models 2 and 4 on T4 in both tables show that women who apply under T4 do not significantly differ in education and job-experience than those who apply under T1 & T2. This suggests that women of low and high experience both opt out of applying in T4, doing so at similar rates. Together, these results suggest that highly-qualified women are in fact deterred by T4. While deterring any talented individuals is generally undesirable, deterring talented women can be of particular concern. While job qualifications are important in their own right, they may be an imperfect proxy for ability and performance. To the extent that this is true, inferences with regard to ability levels of job applicants and worker performance should be interpreted with caution.

8. Discussion

A pillar of economics is that incentives matter. Sociology and management literatures have long studied the role of both financial and non-financial incentives within firms and organizations. The seminal contract theory work of Hart and Holmstrom (1987) and the proliferation of personnel economics (Lazear 1995) helped to usher such incentive questions into the mainstream of economics.

Our research takes this broader literature into a different direction to explore how the compensation scheme influences worker sorting and is related to other field experimental studies considering gender differences (Bagues and Esteve-Volart, 2010; Booth and Leigh, 2010; Barankay, 2011; Buser et al, 2012; Delfgaauw et al, forthcoming). Buser et al (2012), for example, show the extent to which laboratory competition measures are related to individual education choices, which can have important later impacts on labor market outcomes. Delgaauw

et al (forthcoming) conduct a field experiment and find that sales competitions affect sales growth differently depending on the gender composition of the employees.

Our study is also related to the literature relying on surveys to investigate gender differences in psychological traits and their relationship to gender gaps in labor markets (for a recent review see Bertrand, 2011). Fortin (2008), for example, uses high school longitudinal surveys and finds that gender differences in traits such as self-esteem and importance of money/work have a modest effect in explaining the gender wage gap. Manning and Swaffield (2008) use data from a set of psychological attributes of men and women before labor market entry and find that they can roughly explain half of the gender wage gap. Babcock et al (2006) find evidence for pronounced gender differences in the willingness to negotiate which may explain parts of the gender wage gap.

An earlier generation of empirical studies exploited firms' personnel data to measure the productivity effects of compensation schemes on individual workers (see List and Rasul 2010, for a review). An econometric challenge facing these studies is that observed incentive contracts are endogenously determined (Prendergast 1999, Chiappori and Salanie 2003). This means that identifying causal effects of incentives on behavior is confounded by the presence of unobservables, such as managerial practices or workplace amenities that determine both which compensation schemes are chosen, and worker productivity. In earlier research relying on cross-sectional variation this concern has been addressed using instrumental variables (Groves et al., 1994). However, this concern applies even if such effects are identified purely from within firm observations (Jones and Kato, 1995; Ichniowski et al., 1997; Paarsch and Shearer, 1999; 2000; Lazear, 2000).

Field experiments introduce exogenously timed variation in incentive structures that are orthogonal to other management practices or workplace amenities. This provides us with the opportunity to identify the causal impact of compensation regimes on the behavior of individual job seekers. In this way, our study extends a rich literature on gender-based behavioral differences in several meaningful ways. First, by taking the experiment to the field we are provided with a glimpse of whether, and to what extent, gender-dependent responses to different pay regimes influence the first stage of the hiring decision. Second, by randomly varying the compensation scheme faced by an individual who has already expressed interest in a job, we isolate the impact of the salary regime on the gender balance of applicants, which is likely to

influence gender composition of employees**.** In doing so, we find that women are generally more sensitive to divergences from fixed-wage compensation. In particular, we find that competitive workplaces can in fact significantly decrease the propensity of women to apply for a job compared to that of men. These findings are robust to variations in the gender of competitors and supervisors rendering it unlikely that the gender gap in applications for competitive workplaces is the result of gender-dependent expectations to win the competition.

Complementary treatments provide important qualifications for this result. For instance, while the dissuasive effect of competition-based compensation on women can be strong and significant when stakes of the competition are on the order of 50% of base-wage, the effect substantially diminishes when the stakes of relative performance drop to 22% of the base-wage. In addition, including a team element to the competitive environment, such that compensation depends on the relative performance of a group of workers to which one belongs can considerably attenuate the gender gap. Moreover, we find that the gender gap in high-stakes competition jobs is roughly the same as that caused by high wage uncertainty. This suggests that preferences over uncertainty can be just as important as preferences over competition in affecting job-entry choices and gender imbalances.

Whether competition disproportionately discourages women from applying to a job also appears to heavily depend on minor changes in the job-task and possibly gender norms surrounding the task. Among job-seekers interested in the general ad, which redefines the job-task so as to remove associations with activities that may be construed as "male," neither of the competition treatments changes the gender balance among applicants. This result suggests that previous findings regarding gender differences in choices over competition are based at least in part on the task performed. It is consistent with recent lab evidence that the task used (math versus verbal) can have significant effects on gender differences in competitiveness (Shurchkov, 2012; Grosse and Riener, 2012). We show that, in the field, even a marginal adjustment of the job-task can be sufficient to induce or eliminate large gender differentials in job applications caused by competition-incentives. The change does not have to be as drastic as switching a job from math computations to tasks relying on verbal skills, but can be a simpler matter of framing. In field settings, when economic agents are faced with choices regarding job-selection, the nature of the work matters. Therefore, descriptive framing of the jobs and their potential gender associations may play a key role in the ultimate sorting of workers.

In addition, we find that age appears to play a crucial role in gender-dependent responses to highly competitive workplaces. The gender gap induced by high-stakes competition incentives is only significant for the pool of younger job-seekers. The findings suggest that younger men are more likely to apply for a job in a competitive workplace than for the same job in a workplace which offers a fixed wage. This is consistent with previous findings in the literature which show that (young) men are competition-loving. However, older men are clearly more attracted to a workplace which offers a fixed wage. This suggests that our second experimental result (that the gender gap we observe in the field is not driven by a general attraction by men to competitive workplaces) may be driven in part by the broader age distribution of agents in the field than has been the norm for subjects studied in the lab.

We furthermore find that the effect is strongest in labor markets where good outside options are available. This suggests that broader market conditions are important for the strength and presence of the gender difference, and that women are responding appropriately to 'prices,' in the sense that they are trading off lower fixed wages with more competitive job environments.

In closing, some words of caution are in order. First, it is important to recognize that we have chosen a specific type of job in our natural field experiment. We chose an office support position both because it is the most common occupation in the US and gender differences in wages among office support and low skilled occupations nearly equal that of high-skilled occupations. Furthermore, Maxwell (2006) finds that 90% of low-skilled jobs have opportunities for promotional advancement; competitions for promotions and more lucrative jobs as well as compensation schemes based on performance[25] (and their potential impacts on gender imbalances) are not limited to high-profile careers. Our findings thus have a direct bearing on a large part of the labor market. However, if office support positions attract types that respond differently to competitive environments, the implications of our results for individuals in high-profile careers become less clear. In examining observable characteristics, we see that the typical administrative assistant did not attend university and is significantly older than the typical subject in a lab experiment. This study therefore represents a different demographic than the university undergraduate volunteers for experiments underlying previous findings. The

[25] In the job boards where we posted our job advertisement, we observe that roughly 10-15% of the advertised office support positions refer to bonus compensation schemes. The actual fraction of office support positions offering such compensation schemes may be substantially higher as many job advertisements (like ours) do not specify compensation details in the initial advertisement itself.

robustness of behavior patterns across these different subsets of the population (once we control for age) and very different choice environments suggests our results are likely to extend to other occupational settings. Whether the job type interacts with the effect of competition incentives on the gender gap is an important issue that remains unresolved.

Second, we examine job applications and not final acceptance of the job. In this manner, we are observing a critical first step in the hiring process to learn more about gender-dependent responses to contract types. In as much as learning about the ultimate distribution of workers, we are more limited as additional aspects of the labor market may also play a role in final worker sorting.[26] These represent fruitful avenues for future research.

[26] However, recall that the gender distribution of individuals interested in the general job closely mirrors the gender distribution in the US workforce for this type of job, underscoring the importance of application patterns for realized labor force composition.

REFERENCES

Akerlof, George A. and Kranton, Rachel E. 2000. "Economics and Identity." *Quarterly Journal of Economics*, 115(3): 715-753.

Altonji, Joseph G., and Rebecca Blank. 1999. "Race and Gender in the Labor Market." In *Handbook of Labor Economics, Vol. 3C*, eds. O. Ashenfelter, and D. Card, 3144–3259. Amsterdam: Elsevier Science.

Azmat, Ghazala, Maia Guell, and Alan Manning. 2006. "Gender Gaps in Unemployment Rates in OECD Countries." *Journal of Labor Economics*, 24(1), 1-37.

Babcock, Linda, Michele Gelfand, Deborah Small, and Heidi Stayn. 2006. "Gender differences in the propensity to initiate negotiations." In *Social Psychology and Economics*, eds. David De Cremer, Marcel Zeelenberg, and J. Keith Murnighan. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bagues, Manuel and Berta Esteve-Volart, 2010. "Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment". *Review of Economic Studies*. 77, 1300-1328.

Balafoutas, Loukas and Matthias Sutter. 2012. "Affirmative action policies promote women and do not harm efficiency in the lab" *Science*, 335: 579-582.

Barankay, Iwan. 2011. "Gender differences in productivity responses to performance rankings: Evidence from a randomized workplace experiment". Working paper.

Barber, Bret M. and Terrance Odean. 2001. "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics*, 116(1): 261-292.

Bertrand, Marianne. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics, Vol. 4B*, eds. Orley Ashenfelter, and David Card, 1543-1590. Amsterdam: Elsevier North Holland.

Bertrand, Marianne, and Kevin F. Hallock. 2001. "The Gender Gap in Top Corporate Jobs." *Industrial and Labor Relations Review*, 55(1): 3–21.

Blau, Francine, and Lawrence M. Kahn. 2000. "Gender Differences in Pay." *Journal of Economic Perspectives*, 14(4): 75–99.

Booth, Alison, and Andrew Leigh. 2010. "Do employers discriminate by gender? A field experiment in female-dominated occupations?" *Economics Letters*, 107(2):236-238.

Borghans, Lex; Golsteyn, Bart H.H.; Heckman, James J.; Meijers, Hubb. *2009.* "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association, 7(2-3): 649-658.*

Bureau of Labor Statistics *Employment and Earnings*, January 2010, Vol. 57, No. 1. Available online: http://www.bls.gov/opub/ee/empearn201001.pdf

Buser, Thomas; Niederle, Muriel; Oosterbeek, Hessel. 2012. "Gender, Competitiveness and Career Choices. Working paper, Stanford University.

Card, David; Mas, Alexandre; Moretti, Enrico and Saez, Emmanuel. 2010. "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *NBER Working paper #16396*.

Cason, Timothy N.; Masters, William A. and Sheremeta, Roman M. 2010. "Entry into Winner-Take-All and Proportional-Prize Contests: An Experimental Study," *Journal of Public Economics*, 94(9-10): 604–611.

Charness, Gary, and Marie-Claire Villeval, 2009. "Cooperation and Competition in Intergenerational Experiments in the Field and the Laboratory." *American Economic Review*, 99(3): 965-978.

Chiappori, Pierre-Andre and Bernard Salanié. 2003. "Testing Contract Theory: A Survey of Some Recent Work." In *Advances in Economics and Econometrics: Theory and Application, Eighth World Congress (Econometric Society Monographs)*, eds. M. Dewatripont, L. Hansen, and P. Turnovsky, 115–149. Cambridge: Cambridge University Press.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature*, 47(2): 448-474.

Dargnies, Marie-Pierre. Forthcoming. "Men Too Sometimes Shy Away from Competition: The Case of Team Competition" *Management Science, Forthcoming.*

Delfgaauw, Josse; Dur, Robert; Sol, Joeri and Willem Verbeke. Forthcoming. "Tournament Incentives in the Field: Gender Differences in the Workplace." *Journal of Labor Economics.*

Dohmen, Thomas and Falk, Armin. 2011. "Performance Pay and Multidimensional Sorting - Productivity, Preferences and Gender." *American Economic Review*, 101(2): 556-590.

Dreber, Anna, von Essen, Emma and Eva Ranehill. 2011. "Outrunning the gender gap – boys and girls compete equally." *Experimental Economics*, 14(4): 567-582.

Fortin, Nicole. 2008. "The Gender Wage Gap Among Young Adults in the United States: The Importance of Money vs. People." *Journal of Human Resources*, 43(4): 886-920.

Gabriel, Paul E. and Susanne Schmidtz. 2007. "Gender Differences in Occupational Distributions Among Workers." *Monthly Labor Review,* 130(6): 19-24.

Galor, Oded, and David N. Weil. 1996. "The Gender Gap, Fertility, and Growth." *American Economic Review*, 86(3): 374-387.

Gneezy, Uri, Kenneth L. Leonard and John A. List. 2009. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *Econometrica*, 77(3): 909-931.

Gneezy, Uri, Muriel Niederle and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, 118(3): 1049–1074.

Goldin, Claudia. 1990. *Understanding the Gender Gap: An Economic History of American Women*, New York: Oxford University Press.

Goldin, Claudia, and Cecilia Rouse. 2000. "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians." *American Economic Review*, 90(4): 715–741.

Grosse, Niles D., and Gerhard Riener. 2012. "Explaining the Gender Differences in Competitiveness: Gender-Task Stereotypes." *Jena Economic Research Paper No 2010-017*, Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

Groves, Theodore, Yongmiao Hong, John McMillan and Barry Naughton. 1994. "Autonomy and Incentives in Chinese State Enterprises." *Quarterly Journal of Economics,* 109(1):183–209.

Datta Gupta, Nabanita, Anders Poulsen and Marie-Claire Villeval. 2011. "Gender Matching and Competitiveness: Experimental Evidence." *Economic Inquiry*, 51(1): 816-835.

Flory, Jeffrey, Gneezy, Uri, Leonard, Kenneth and John List. 2012. "Sex, Competitiveness, and Investment in Offspring: On the Origin of Preferences." *Working paper.*

Harrison, Glenn and John A. List. 2004. "Field Experiments." *Journal of Economic Literature*, 42(4): 1009–1055.

Hart, Oliver, and Bengt Holmstrom. 1987. "The Theory of Contracts." In *Advances in Economic Theory: Fifth World Congress*, ed. Truman F. Bewley. Cambridge: Cambridge University Press.

Hotchkiss, Julie L., and Melinda M. Pitts. 2007. "The Role of Labor Market Intermittency in Explaining Gender Wage Differentials." *The American Economic Review*, 97(2): 417-421.

Ichniowski, Casey, Kathryn Shaw and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *American Economic Review*, 87(3): 291-313.

Jones, Derek and Takeo Kato. 1995. "The Productivity Effects of Employee Stock-Ownership Plans and Bonuses." *American Economic Review*, 85(3):391–414.

Kuhn Peter, Marie-Claire Villeval. 2011. Do Women Prefer a Co-operative Work Environment ? IZA Discussion Paper 5999, Bonn, and GATE WP 11-27

Lazear, Edward. 1995. *Personnel Economics*, Cambridge, MA: MIT Press.

Lazear, Edward. 2000. "Performance Pay and Productivity." *American Economic Review*, 90(5): 1346-1361.

Lemieux, Thomas, MacLeod W. Bentley, and Daniel Parent. 2009. "Performance Pay and Wage Inequality." *Quarterly Journal of Economics*, 124(1): 1-49.

List, John A. and Imran Rasul. 2010. "Field Experiments in Labor Economics." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 16062, June 2010.

Manning, Alan, and Joanna Swaffield. 2008. "The gender gap in early-career wage growth." *Economic Journal*, 118 (530): 983-1024.

Maxwell, Nan L. 2006. *The Working Life: The Labor Market for Workers in Low-Skilled Jobs*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

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

Niederle, Muriel, and Lise Vesterlund. 2011. "Gender and Competition." *Annual Review of Economics*, 3: 601-630.

Paarsch, Harry J., and Bruce Shearer. 1999. "The Response of Worker Effort to Piece Rate: Evidence from the British Columbia Tree-Planting Industry." *Journal of Human Resources*, 34(4): 643-667.

Phipps, Shelley, Peter Burton, and Lynn Lethbridge. 2001. "In and Out of the Labor Market: Long-Term Income Consequences of Child-Related Interruptions to Women's Paid Work." *Canadian Journal of Economics*, 34(2): 411-29.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37(1): 7–63.

Reuben, Ernesto, Sapienza, Paola and Luigi Zingales. 2010. "The glass ceiling in experimental markets." *Working paper*.

Salop, Joanne and Steven Salop. 1976. "Self-Selection and Turnover in the Labor Market." *Quarterly Journal of Economics*, 90(4): 619–627.

Shurchkov, Olga. 2012. "Under Pressure: Gender Differences in Output Quality and Quantity Under Competion and Time Constraints." *Journal of the European Economic Association*, 10(5): 1189-1213.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*, 87(3): 355–374.

Spencer, Steven J., Claude M. Steele, and Diane M. Quinn. 1999. "Stereotype Threat and Women's Math Performance." *Journal of Experimental Social Psychology*, 35, 4-28.

Sutter, Matthias and Ruetzler, Daniela. 2010. "Gender Differences in Competition Emerge Early in Life," IZA D.P. 5015, 2010.

Wilkinson, Richard and Kate Pickett. 2009. "The spirit level: why more equal societies almost always do better." Penguin Books.