Causes of Gender Differences in Competition: Theory and Evidence

Christopher Cotton
Frank McIntyre
Joseph Price*

September 28, 2010

Abstract

We use a game theoretic model of contests to assess different explanations for the male performance advantage in competition. Comparing the testable predictions of the model with the empirical evidence, we reject explanations involving male overconfidence, misperceptions about relative ability, and some preference differences. Explanations involving female underconfidence, stereotype threat, and adverse female reaction to competition are consistent with only some of the evidence, and an explanation involving lower male risk aversion is consistent with most of the evidence. Two explanations are consistent with all of the evidence: (i) male ability to perform may increase in the face of competition, possibly due to changes in testosterone or adrenaline; or (ii) males may care more about winning or get greater enjoyment from competition than females.

(*JEL* D74, J16, J24, J78)

Keywords: contests; gender differences; effort and productivity

*Cotton: cotton@business.miami.edu; Department of Economics, University of Miami, Coral Gables, Florida 33146. McIntyre: mcintyre@byu.edu; Department of Economics, Brigham Young University, Provo, Utah 84602. Price: joe_price@byu.edu; Department of Economics, Brigham Young University, Provo, Utah 84602.

1 Introduction

A number of recent articles present evidence that males tend to respond more favorably than females when faced with competition. Gneezy, Niederle and Rustichini (2003) conduct experiments in which college students solve mazes, either on their own or in competition with other students. They show that competition causes males to increase their performance by more than females. Gneezy and Rustichini (2004) find similar results when they study footraces between young children: males increase their performance in the face of competition, while females do not. Cotton, McIntyre and Price (2010) conduct multiple-round math competitions and find evidence that males outperform females of the *same* ability during the initial round of competition.[1] This male advantage may help explain achievement differences between males and females that have been documented in competitive academic and workplace settings (Blau and Kahn (2000)).

In their influential article, Gneezy, Niederle and Rustichini (2003) discuss various explanations for the observed male advantage. Some explanations involve real differences between males and females. For example, males may be inherently better competitors, or males may enjoy competing or care about winning more than females. Some explanations are behavioral in nature. For example, females may be under-confident about their chances of winning, or males may be over-confident. Similarly, people may have inaccurate beliefs about the relative abilities of males and females, which could affect their effort during competition. Such explanations, as well as the possibility that the performance gap is due to differences in risk aversion, have been discussed in other papers as well, including Niederle and Vesterlund (2007), Gneezy and Rustichini (2004), Croson and Gneezy (2009), and Günther et al. (in press, 2010). None of these articles provide a theoretical framework to assess the merits of these explanations. However, there exists an extensive theoretical literature modeling contests and tournaments that can be applied to provide insight into the causes of gender differences in response to competition.

The current article provides a theoretical assessment of the possible explanations for the male advantage during competition. We present a simple game theoretic model of a two-person contest in which both players simultaneously choose effort. Performance in the contest is a function of both effort and ability, with the probability of winning the contest

[1] Various recent articles work to determine in which settings the male advantage exists. Cotton, McIntyre and Price (2010) show that the male advantage only lasts for one round in such a setting, and depends crucially on the framing of the competition as a race. Günther et al. (in press, 2010) find that the male advantage is task dependent, as they show that it exists during maze competitions, but not during competitions involving word games, pattern matching, or memory tasks. Gneezy, Leonard and List (2009) present evidence that gender differences in the face of competition depend on participant background and social norms. Our analysis is only applicable to settings in which the male advantage exists.

increasing in one's own performance and decreasing in opponent performance. Players may differ in their ability or preference parameters, as well as in their (potentially inaccurate) beliefs about ability. The model, based on Tullock (1980) and Baik (1994), is the standard framework in the theoretical literature for modeling contests with asymmetric players.[2] Next, we adapt the model to account for various explanations of the male advantage, then compare the equilibrium predictions with the empirical evidence. Both the empirical and theoretical analyses define the competitors' type by their own gender and the gender of their opponent. Therefore, instead of just comparing the performance of males and females, the analysis is concerned about the relative performance of four player types: males when competing against males (MvM), males when against females (MvF), females when against males (FvM), and females when against females (FvF).

The empirical analysis considers datasets from both Gneezy, Niederle and Rustichini (2003) and Cotton, McIntyre and Price (2010). Together, the data allow us to draw conclusions about relative performance. We show that males tend to outperform females, regardless of opponent gender. The condition is statistically significant, and we use the relationship to reject any model that does not predict that the performance of each of the male types (MvM and MvF) is greater than the performance of each of the female types (FvM and FvF).

The theoretical model predicts behavior that is not always in line with initial intuition. For example, the model shows that competitors put in the most effort in contests in which they are evenly matched against an opponent. Starting from an evenly-matched contest, increasing one player's ability results in a less competitive contest, and in equilibrium both players respond by putting forth less effort. The high-ability player puts forth less effort because he can expend less effort and still perform better than before. The low-ability player puts forth less effort because her marginal expected return from effort is decreasing in opponent ability. This means that in a lopsided contest, both players put forth less effort than in a contest between two same-ability players.

A high-ability competitor is more likely to win a contest against a low-ability competitor, not because he puts in more effort, but rather because he achieves higher performance with the same effort. This is an important distinction when considering explanations in which players have misperceptions about their own ability or the ability of their opponent. If a player is overconfident in his own ability, for example, then he underestimates the competitiveness of the contest, and puts in less effort than if he had accurate beliefs about his ability. If his ability advantage was real, his lower effort would not fully offset the advantage

[2]See also Dixit (1987), Nitzan (1994), Stein (2002), and Brown (2010). Skaperdas and Gan (1995) incorporates differences in risk aversion into such contests. Application of contest theory to workplace achievement include O'Keefe, Viscusi and Zeckhauser (1984), Main, O'Reilly and Wade (1993), Chan (1996), Tsoulouhas, Knoeber and Agrawal (2007) and Carpenter, Matthews and Schirm (2010).

of higher ability, and he would still experience an increase in performance. However, because he overestimates his ability, his lower effort results in lower equilibrium performance.

The case of overconfidence illustrates the importance of formally considering the theoretical model. Gneezy, Niederle and Rustichini (2003) hypothesize that "It might be that men are solving 'too many' mazes, because they ... are overconfident about their abilities and hence their chances of actually winning the tournament" (p 1060). This statement, and others found in the literature, is inconsistent with a game theoretic model of contests. If a male overestimates his ability, he underestimates the competitiveness of the contest which causes him to put in less effort and perform worse than opponents who have correct beliefs about ability. The theoretical analysis shows that overconfidence has the opposite effect on performance than what has been assumed in the literature, and by comparing the model to the empirical requirement, we are able to reject the male-overconfidence explanation of gender differences. For similar reasons we can also reject a model in which players have incorrect beliefs about male or female ability. Additionally, we rule out other explanations for the male advantage including explanations in which players dislike losing to females, or have preferences for competing against an opponent of the same gender. This leads us to reject explanations for the male advantage involving male-overconfidence or general misperceptions about ability, as well as a number of explanations involving preference differences.

A closer look at the empirical evidence allows us to rule out a model of female underconfidence. First, such a model predicts that male performance is independent of opponent gender, which is inconsistent with the observations in the data. Second, it assumes that the performance differences are the result of females underperforming in competition. This is in contrast to evidence in the literature that the gender differences are due to males increasing their performance when faced with competition rather than from females decreasing their performance (e.g., Gneezy, Niederle and Rustichini 2003, Cotton, McIntyre and Price 2010). These same data also allow us to rule out some explanations involving females having a lower ability to deal with the pressures of competition, including those involving stereotype threat.[3]

Two explanations are consistent with all of the empirical evidence we consider. First, males may increase their ability when faced with competition. This explanation does not imply that males are better at solving math questions, competing mazes, or running races.[4]

[3]Stereotype threat refers to the idea that people suffer from additional anxiety when completing a task that their "type" (e.g., gender, race) is stereotypically not good at performing. This increases the likelihood they choke under the pressure of competition. See Gneezy et al. (2003, p 1059) and Günther et al. (in press, 2010).

[4]Cotton, McIntyre and Price (2010) and Gneezy and Rustichini (2004) find gender differences even when controlling for a student's performance in a non-competitive setting.

Rather, it means that competition improves male ability to convert effort into performance. This could be the result of males benefiting more from increases in testosterone or adrenaline.[5] Second, males may care more about winning or enjoy competition more than females. This explanation is also consistent with evidence that males choose to compete more often than females (e.g., Niederle and Vesterlund 2007). A third explanation–that performance differences are driven by lower male risk aversion–is consistent with most of the empirical evidence.

The contribution of the paper is threefold. First, we provide a theoretical framework in which to assess gender differences in reaction to competition. In addition to the explanations considered here, the model may be adapted to consider alternative explanations, or to assess the causes of performance differences in settings where player type is defined by some other characteristic rather than gender (e.g., race, age, income). Second, the model provides guidance for future researchers in designing new experiments that will differentiate among the alternative explanations that are consistent with the data. Third, we provide a formal assessment of a number of popular explanations for the male advantage in competition. Because of the strategic nature of competition, one's choice of effort and performance during competition is not always as intuitive as people may assume. A careful consideration of the competitive interaction allows us to rule out a number of explanations for the male advantage in competition. At the same time, we provide evidence consistent with other explanations.

2 Evidence

This section presents experimental evidence which we then compare to the testable predictions from the theoretical models in Section 3. The theoretical analysis develops predictions involving the ordering of performance from the four gender combinations: males competing against males (MvM), males against females (MvF), females against males (FvM), and females against females (FvF). This means that the empirical evidence should provide results about the actual effect of one's own and opponent gender on performance.

Our data come from Gneezy, Niederle and Rustichini (2003) (henceforth, GNR) and Cotton, McIntyre and Price (2010) (henceforth, CMP).[6] Both data sets present significant evidence that males on average tend to outperform females during competition. However, the

[5]Competition increases testosterone in males, but not for females, compared to the period leading up to competition (Kivlighan, Granger and Booth 2005). Both testosterone and adrenaline improve physical and mental performance (e.g., Clark et al. 1989, Chmura et al. 1998, Brisswalter et al. 2002, Salvador et al. 2003).

[6]The GNR data can be found in the May 30, 2001 working paper version of the published paper, and the CMP data is available from the authors.

predictions are weaker when we consider the effects of both own gender and opponent gender. Specifically, each data set is limited by the number of observations, and some predictions regarding gender effects are not statistically significant. By combining the two sources into an aggregate data set, we have enough observations to draw statistically significant conclusions about the effect of gender on performance.

The CMP data involves 253 total participants (136 males and 117 females) competing in contests to solve math problms as quickly and accurately as possible.[7] Participants were randomly matched with a single opponent, and competed in a one-on-one contest. The GNR data involves 120 total participants (60 males and 60 females) competing in contests to solve mazes as quickly as possible.[8] Participants were randomly selected into groups of six people, which were either mixed gender (three males and three females) or single gender. The person to solve the most mazes in each group won a prize. To simplify the analysis and allow us to combine the data sources, we categorize someone in a mixed-group as competing against an opponent of the other gender.[9]

Analysis

For each data set, we form a standardized measure of performance by subtracting the mean and dividing by the standard deviation. After this normalization a nonparametric Mann-Whitney U (MWU) test cannot reject that the GNR and CMP data come from the same distribution (p=0.957).[10] The similar distributions in the two data sets make us more comfortable combining the data for an aggregate analysis. As we'll see below, the point estimates are also similar across the two.

Our model gives predictions about the rank ordering of performance for males and females based on whether they are competing against their own or opposite gender. Thus in the first panel of Table 1 we list the average performance for MvM, MvF, FvF, and FvM. We do this for the GNR and the CMP data separately and then for the aggregated data that combines the two datasets. The second panel provides p-values for a set of exhaustive t-tests comparing

[7]Cotton, McIntyre and Price (2010) also presents data from multiple rounds of competition. They show that after the first round of competition, the gender differences disappear. Because of this, we only use data from the first round of competition. Because the data come from a field experiment done in classrooms, the number of participants was not predetermined by the researchers but by the class size. In classes with odd numbers of students, one pair of competitors had three students instead of two.

[8]Gneezy, Niederle and Rustichini (2003) also presents data from non-competitive treatments, which we do not include in our data.

[9]Allowing for multiple opponents in the theoretical model greatly complicates the analysis without changing the results. See for example, Stein (2002).

[10]The distributions are both roughly symmetric with skewness close to zero (CMP = -0.13 and GNR = 0.17) and kurtosis slightly less than the normal (CMP = 2.86 and GNR = 2.62). Thus the two datasets have similar higher moments.

the means of each of the groups. Table 2 presents the same analysis using MWU tests rather than t-tests.[11] In GNR, the best performance comes from the MvF group, followed by MvM, then FvF, and finally FvM. However, while FvM is significantly different from both male groups (p-values of 0.001 and 0.008), FvF is not significantly different from MvM (p-value = 0.185), though it is different from MvF at the ten percent level (p-value = 0.063). In CMP, the best performance comes from the MvM group, followed by MvF, then FvF, and finally FvM. Once again, FvM is significantly different from both male groups (p-values of 0.065 and 0.002). CMP can reject what GNR could not—FvF is different from MvM (p-value = 0.032)—but the data are uninformative about FvF vs. MvF (p-value = 0.272).

The performance ordering of the four gender groups is almost identical in the two data sets, with the exception being that in GNR males perform better when competing against females, and in CMP males perform better when competing against males. In neither data set is the interaction effect of opponent gender significant. As we wish to aggregate the data, we test each of the coeffiicents to see if it is statistically different across the two datasets. We run four individual t-tests and a joint test of all four restrictions and never reject the hypothesis that both datasets are converging to the same estimate.[12] Given these similarities, we focus most of our attention below on the aggregated data, which, because of the larger sample size, have more precise standard errors.

Neither of the individual data sources provides enough evidence on its own that, regardless of opponent gender, males tend to outperform females. From GNR, we are not confident that FvF is significantly different than MvM (p-value = 0.185) and from CMP, we are not confident that FvM performs worse than MvF (p-value = 0.272). Using both datasets allows us reject equality across genders. Furthermore, when we aggregate the data, we see that males significantly outperform females, regardless of opponent gender. That is, we can reject that either of the female groups (FvF or FvM) performs as well as either of the male groups (MvM or MvF). This is true either as individual tests (the p-values on the relevant comparisons range from 0.0001 for FvM vs. MvM, to 0.057 for FvF vs. MvF), or as a joint test of all four restrictions at once (p=0.0002 in the aggregate data, in 0.007 GNR and 0.015 in CMP). This brings us to the first condition, which we rely on in the theoretical analysis.

C 1 (required) *Males outperform females, regardless of opponent gender.*

[11] The Mann Whitney U test has the advantage of not relying on an explicit distributional assumption, asymptotic or otherwise. Unfortunately it is not truly a test of means but rather of distributions, as it can also reject because of differences in variance or higher order moments. Thus we base our discussion on the t-tests in Table 1 but provide the additional MWU tests for reference. As one can see, the MWU tests provide even stronger rejections than those in Table 1.

[12] p-values for CMP vs. GNR coefficients are 0.869 (FvF), 0.355 (FvM), 0.163 (MvF), and 0.890 (MvM). The joint test of all four restrictions cannot reject with a p-value of 0.583

The theoretical analysis in Section 3 rejects any model that is inconsistent with C1. In Section 4 we introduce a number of additional conditions from our analysis and the literature. The additional conditions are not statistically significant, and we therefore reject a model outright only if it violates C1.

3 Theoretical Model

We apply theoretical models of contests to assess the various explanations for male-female performance differences during competitions. The underlying model is based on Baik (1994)'s adaption of Tullock (1980)'s rent seeking model to allow for player asymmetries. Later, we incorporate risk aversion in a way similar to Skaperdas and Gan (1995).

3.1 A Simple Contest

Two players, $i = 1, 2$, engage in competition for a prize. The players independently choose a level of costly effort, $e_i \geq 0$, which affects their probability of winning the contest. Let $W_i \in \{0, 1\}$ indicate whether player i wins the prize. Payoffs equal $u_i^* = W_i v_i - c_i e_i$ for $i = 1, 2$, where $v_i > 0$ is i's benefit from winning the contest, and $c_i > 0$ is his cost of effort. Since behavior is unchanged by affine transformations, we may rewrite the utility function $u_i \equiv u_i^*/v_i = W_i - \tau_i e_i$, where $\tau_i \equiv c_i/v_i$. Thus, differences in τ_i account for differences in both cost of effort and value of winning.

Parameter $a_i > 0$ denotes player i's *ability*, and $p_i \equiv a_i e_i$ denotes a player's *performance*. That is, ability represents how effective one is at turning effort into results. Player costs, valuations, and ability parameters are common knowledge. The probability that player i wins the contest depends on the performance of *both* players, and is equal to $\frac{p_i}{p_1+p_2}$ if $p_i > 0$ and 0 when $p_i = 0$. Therefore, a player's expected utility equals

$$Eu_i = \frac{a_i e_i}{a_1 e_1 + a_2 e_2} - \tau_i e_i. \tag{1}$$

We solve for the Nash equilibrium of this simultaneous move game. In equilibrium, neither player can have an incentive to deviate from their effort choice given the effort choice of the other player. Solving $\frac{\partial Eu_1}{\partial e_1} = 0$ and $\frac{\partial Eu_2}{\partial e_2} = 0$ for e_1 and e_2 gives the equilibrium values of e_1 and e_2. That is,

$$e_i^* = \frac{a_i a_{-i} \tau_{-i}}{(a_i \tau_{-i} + a_{-i} \tau_i)^2}, \tag{2}$$

where subscript $-i$ denotes the other player. A check of the second order conditions assures that this achieves a maximum. From this expression, we can determine equilibrium performance,

$$p_i^* = \frac{a_i^2 a_{-i} \tau_{-i}}{(a_i \tau_{-i} + a_{-i} \tau_i)^2}. \tag{3}$$

Players may differ in terms of their valuation and cost parameters, which enter our model through τ_i, and in terms of their ability, a_i. A player has a lower τ_i than his opponent, for example, if he puts a higher value on winning the contest (a higher v_i), or if he is more eager to compete (a lower c_i). A difference in a represents actual differences in players' ability to convert effort into performance. This is the case, for example, if one of the players is better at converting effort into results under the pressures of competition, benefits more from jumps in adrenaline or testosterone that result from competition, or is less susceptible to stereotype threat.

It is worth pointing out two results regarding the equilibrium solution that are helpful in the later analysis. Both are formally derived in Baik (1994). First, starting out from an even contest in which $a_i = a_{-i}$ and $\tau_i = \tau_{-i}$, suppose that we decrease one player's preference parameter τ (meaning we increase his enjoyment of competition or his value of winning). Doing so causes the player with the higher τ to expend more effort and achieve higher performance, and for his competitor to expend less effort and achieve lower performance. Second, starting from an even contest, suppose that we increase one player's ability a. The player with the ability advantage can decrease his effort and still perform better than he was performing. In equilibrium, he chooses to decrease effort but not by enough to fully offset the effect of the ability increase on performance. The other player also cuts her effort, but for her the cut in effort is not accompanied by an increase in ability and her performance suffers. That is, increasing one player's ability decreases both players' effort, decreases the performance of the disadvantaged player, and increases the performance of the advantaged player.

Incorporating Gender Differences

Our analysis uses this contest model to assess different explanations for why males tend to outperform females during competition. The remainder of the analysis assumes that players are either "female" or "male," although the categorization could just as easily be based on race, age, or other observable characteristics. We formally define player type in terms of both one's own gender and opponent gender, as the analysis is concerned with the effect of both. Let $t \in \{FvF, FvM, MvF, MvM\}$ denote a player's type, where the first letter denotes one's own gender, and the final letter denotes opponent gender. When a variable is independent of opponent gender, we may simply use F and M to denote player type. To keep the analysis as straightforward and intuitive as possible, we assume that all players of the same type share the same values of τ_t and a_t.[13]

3.2 Ability Differences

The first possibility we consider is that males are higher-ability competitors than females. We incorporate this into the model by assuming that $a_M > a_F$. To focus the analysis on the effect of ability differences, here we assume $\tau = \tau_M = \tau_F$.

Most empirical analyses identify a male advantage during competition, even when they account for performance in non-competitive settings. The literature shows, for example, that if you take males and females who are equally good at math and have them take a math quiz in a competition, then the male students will typically outperform the female students (Cotton, McIntyre and Price 2010). This means that the male advantage during competition cannot be explained simply by males being better at the task at hand than females. Ability differences, in our context, must refer to differences in the ability to compete. This may be due to males receiving a greater benefit from changes in testosterone or adrenaline, or males otherwise having an advantage at converting effort into performance *during* competition.[14]

In a male/female competition, the players differ in terms of a but not τ. In this case, equilibrium effort simplifies to $e_i^* = \frac{a_i a_{-i}}{(a_i + a_{-i})^2 \tau}$. This expression goes against standard intuition, as $e_i^* = e_{-i}^*$, which means that both players put in the same amount of effort independent of which one has higher ability. This is because the marginal return on effort depends on how competitive the contest is. If one player is much more able than his competitor, this decreases the return from effort for both players. When competing against a low ability opponent, a high ability player can put in less effort and still win the competition most of the time. Similarly, the low-ability player also puts in less effort since the expected returns from effort are lower when competing against a high-ability player than when competing against another low-ability player. This does not imply that the two players perform equally well in equilibrium. When both players put in the same amount of effort, the high-ability player is better able to turn that effort into performance with $p_i^* \equiv a_i e_i^*$.

[13]One could alternatively assume that same type players may differ in terms of a and τ, but that the distributions are such that one type tends to have an advantage over another type. For example, males may tend to have higher ability or enjoy competition more than females. In such a setting, our results would continue to hold for average performance, although they would not hold for every single competition.

[14]The testosterone and adrenaline stories are supported by evidence from the medical and psychological literatures showing that male testosterone increases in the face of competition, and that testosterone and adrenaline can improve psyical and mental performance. See for example Kivlighan, Granger and Booth (2005), Chmura et al. (1998) and Salvador et al. (2003).

In a male/male or female/female competition, the players share the same a and τ, in which case equilibrium effort simplifies to $e_i^* = \frac{1}{4\tau}$ and performance simplifies to $p_i^* = \frac{a_i}{4\tau}$.

It is straightforward to calculate equilibrium performance for each competitor-opponent combination.

Result 1 *If males are higher-ability competitors than females (i.e., $a_M > a_F$ and $\tau_M = \tau_F$), then*

$$p^*_{FvM} < p^*_{FvF} < p^*_{MvF} < p^*_{MvM}.$$

For this result to be consistent with the empirical evidence as described in Section 2, it must predict that each of the male performance variables exceed each of the female performance variables (i.e., C1). The theoretical result satisfies this condition, and is thus consistent with the significant finding in the empirical data.

3.3 Under- or Overconfidence

Another explanation of the competitive performance gap involve female underconfidence and male overconfidence. To assess these explanations, we consider a version of the model in which players may have incorrect beliefs about their *own* ability. (The next section considers misperceptions about an entire type's ability.) Here, a player acts as if he has ability $\hat{a}_i$. If $\hat{a}_i < a_i$, then i is said to be *underconfident*. If $\hat{a}_i > a_i$, then i is said to be *overconfident*. To keep the analysis as straightforward as possible, players are naive in that they do not recognize that beliefs about themselves and others may be incorrect.[15] This means that in equilibrium,

$$e_i^* = \frac{\hat{a}_i a_{-i} \tau_{-i}}{(\hat{a}_i \tau_{-i} + a_{-i}\tau_i)^2}, \text{ and} \tag{4}$$

$$p_i^* = \frac{\hat{a}_i a_i a_{-i} \tau_{-i}}{(\hat{a}_i \tau_{-i} + a_{-i}\tau_i)^2}. \tag{5}$$

Notice that performance still equals *actual* ability times effort.

If a player's misperceptions about his ability causes him to believe that the competition is less-evenly matched, this causes him to decrease his effort and performance. That is, if $\hat{a}_i \tau_{-i} < a_{-i}\tau_i$, then e_i^* is strictly increasing in $\hat{a}_i$; and if $\hat{a}_i \tau_{-i} > a_{-i}\tau_i$, then e_i^* is strictly decreasing in $\hat{a}_i$.

We consider the impact of differences in self confidence in the absence of any actual gender differences in a and τ. That is, $a = a_M = a_F$ and $\tau = \tau_M = \tau_F$. All players of the same type

[15] Future work may consider the impact of sophistication about the possibility of incorrect beliefs.

have the same level of confidence. If females are underconfident in their ability, then $\hat{a}_F < a$, and the equilibrium expression for e_F^* and e_M^* simplify to $e_F^* = \frac{\hat{a}_F a}{(\hat{a}_F + a)^2 \tau}$ and $e_M^* = \frac{1}{4\tau}$. If males are overconfident in their ability, then $\hat{a}_M > a$, and simplified expressions for e_t^* may be calculated in similar fashion. Deriving p_t^* for these cases is similarly straightforward.

Result 2 *If females are underconfident (i.e., $\hat{a}_F < a = \hat{a}_M$), then*

$$p_{FvM}^* = p_{FvF}^* < p_{MvF}^* = p_{MvM}^*.$$

If males are overconfident (i.e., $\hat{a}_F = a < \hat{a}_M$), then

$$p_{MvM}^* = p_{MvF}^* < p_{FvF}^* = p_{FvM}^*.$$

This result is consistent with the idea that players put in more effort in competitions they believe to be evenly matched. If a players thinks that a competition is lopsided, they put in less effort regardless of whether they are the player with an advantage. In equilibrium, overconfident males decrease their performance just as underconfident females decrease theirs. Of the two confidence stories, only female underconfidence is consistent with C1. Male overconfidence results in males performing worse than females in a male/female competition, in contrast to the empirical evidence.

Although female underconfidence satisfies C1, we will show in Section 4 that it violates a number of other conditions. A variation of the underconfidence story, in which females are more underconfident when competing against a male, continues to satisfy C1 and is slightly consistent with more of the evidence presented in the later section.

Result 3 *If females are under-confident and relatively more so when competing against males (i.e., $\hat{a}_{FvM} < \hat{a}_{FvF} < a$), then*

$$p_{FvM}^* < p_{FvF}^* < p_{MvF}^* = p_{MvM}^*.$$

We could present a similar result for the case when male overconfidence depends on opponent gender. It should be clear at this point, however, that male overconfidence results in males underperforming females, which is inconsistent with the empirical evidence.

3.4 Perceived Differences in Ability

A related explanation for performance differences between males and females is that players believe that males generally have higher ability than females. This is not the same as the real ability differences studied earlier, since we do not require actual differences in a_F and a_M. Nor is this the same as differences in self-confidence from the previous section, as players do not believe they are more or less capable than others of the same type.

Again, assume $a = a_M = a_F$, and $\tau = \tau_M = \tau_F$. We consider two cases. In the first case, players underestimate female ability. A player who does this acts as if all females have lower ability than they actually do, acting as if female ability is $\hat{a}_F < a$. In the second case, players overestimate male ability, in which case $\hat{a}_M > a$. As before, to keep the analysis as straightforward as possible, players are naive and do not recognize that beliefs may be incorrect.

In both cases,

$$e^*_{MvM} = e^*_{FvF} = \frac{1}{4\tau}, \text{ and } p^*_{MvM} = p^*_{FvF} = \frac{a}{4\tau}.$$

We begin with the possibility that players underestimate female ability. As is the case with actual ability differences, perceived differences cause both types of players to put in less effort when they compete against a member of the other group, compared with when they compete against a member of their own group. When a male competes against a female and both underestimate female ability,

$$e^*_{MvF} = e^*_{FvM} = \frac{a\hat{a}_F}{(a + \hat{a}_F)^2\tau}.$$

Another possibility is that only males underestimate female ability. In this case, only the males underestimate the competitiveness of the contest, and

$$e^*_{MvF} = \frac{a\hat{a}_F}{(a + \hat{a}_F)^2\tau}, \text{ and } e^*_{FvM} = \frac{1}{4\tau}.$$

When only females underestimate their ability, the results are symmetric to the case when only males underestimate female ability. One can repeat this analysis assuming that players overestimate male ability rather than underestimate female ability. In each of these cases, a player who has misperceptions about ability will also have misperceptions about the competitiveness of mixed-gender contests, and will put in less effort compared to players with correct beliefs. From this, we derive the following result.

Result 4 *If both males and females underestimate female ability or overestimate male ability, then*

$$p^*_{FvM} = p^*_{MvF} < p^*_{FvF} = p^*_{MvM}.$$

If only males underestimate female ability or overestimate male ability, then

$$p^*_{MvF} < p^*_{FvM} = p^*_{FvF} = p^*_{MvM}.$$

If only females underestimate female ability or overestimate male ability, then

$$p^*_{FvM} < p^*_{FvF} = p^*_{MvF} = p^*_{MvM}.$$

All three predictions in Result 4 violate C1. This allows us to rule out explanations in which performance differences are caused by *perceived* differences in ability.

3.5 Preference Differences

Here, we consider the possibility that performance differences are driven by differences in preferences. For example, males may enjoy competition more than females, which is incorporated into the model by assuming that males have a lower cost of effort (i.e., $c_M < c_F$). Similarly, males may care more about winning the contest than females, which is incorporated into the model by assuming that males have a higher value of winning (i.e., $v_M > v_F$). Both of these cases affect the preference parameter $\tau_i \equiv c_i/v_i$ in the same way, implying that $\tau_M < \tau_F$. To focus the analysis on the effect of preference differences, we assume $a = a_M = a_F$.

In a male/female competition, the players differ in terms of τ but not a. In this case, equilibrium effort simplifies to $e^*_i = \frac{\tau_{-i}}{(\tau_i+\tau_{-i})^2}$. From this we can see that males will put in more effort than females, $e^*_M > e^*_F$, when they care more about winning or get more enjoyment from competing. In male/male and female/female competitions, equilibrium effort simplifies as it did in the previous section with $e^*_i = \frac{1}{4\tau_i}$. It is straightforward to calculate performance in each of these cases.

Result 5 *If males care more about winning or if they get more enjoyment from competing than females (i.e., $a_M = a_F$ and $\tau_M < \tau_F$), then*

$$p^*_{FvM} < p^*_{FvF} < p^*_{MvF} < p^*_{MvM}.$$

This result satisfies C1.

Other explanations involve preferences depending on opponent gender. For example, players may receive greater benefit from winning against a female opponent. Or, they may enjoy competition against one gender more than against another. The model can allow for such possibilities. As before, allow males and females to differ only in their preference parameter τ. Thus,

$$\begin{array}{ll} e^*_{FvF} = \frac{1}{4\tau_{FvF}} & e^*_{MvM} = \frac{1}{4\tau_{MvM}} \\ e^*_{FvM} = \frac{\tau_{MvF}}{(\tau_{MvF}+\tau_{FvM})^2} & e^*_{MvF} = \frac{\tau_{FvM}}{(\tau_{MvF}+\tau_{FvM})^2}. \end{array}$$

There are many cases that the model can consider in which gender differences and preferences depend on own gender and opponent gender. Here, we present results for three situations with intuitive appeal. The appendix considers other situations.

Result 6 *If all players care more about winning or get more enjoyment from competing against females (i.e., $\tau_{MvF} = \tau_{FvF} < \tau_{MvM} = \tau_{FvM}$), then*

$$p^*_{FvM} < p^*_{MvM} < p^*_{MvF} < p^*_{FvF}.$$

If only males care more about winning or get more enjoyment from competing against females (i.e., $\tau_{MvF} < \tau_{MvM} = \tau_{FvM} = \tau_{FvF}$), then

$$p^*_{FvM} < p^*_{FvF} = p^*_{MvM} < p^*_{MvF}.$$

If all players care more about winning or get more enjoyment from competing against an opponent of the same gender (i.e., $\tau_{MvM} = \tau_{FvF} < \tau_{MvF} = \tau_{FvM}$), then

$$p^*_{FvM} = p^*_{MvF} < p^*_{FvF} = p^*_{MvM}.$$

Each of these situations, as well as the alternative situations presented in the appendix, violate C1. This allows us to rule out explanations based on the idea that players want to avoid "losing to a girl," or in which players care more about winning against a same-gendered opponent. However, it is impossible for the analysis to exhaust all possible preference cases, and we recognize that other preferences may produce results that satisfy the required empirical conditions. These other preferences will likely have a similar flavor as the male-enjoyment story in Result 5. For example, the performance ordering in Result 5 continues to hold if males enjoy competition more than females, and experience extra enjoyment when competing against other males.[16]

[16]This variation of the male-enjoyment story is consistent with the behavioral science literature showing that males have a preference for interacting with other males in competitive settings (e.g., Boyatzis, Mallis and Leon 1999).

3.6 Differences in Risk Aversion

A final explanation we consider involves differences in risk aversion between males and females. Here, we show that when females are more risk averse in the face of competition, they may put in less effort than males. Allowing for this possibility, we adapt the contest model used above to allow for non-linear utility, similar to the model of Skaperdas and Gan (1995).

To focus the analysis on differences in risk aversion, we assume that the competitors are otherwise symmetric, with $a = a_M = a_F$, $v = v_M = v_F$, and $c = c_M = c_F$. Competitors have the same initial resource $Y > 0$. Let y denote his payout at the end of the competition, where $y = Y + v - ce_i$ if he wins, and $y = Y - ce_i$ if he loses. Player utility over y exhibits constant absolute risk aversion, with $U_i(y) = -\exp(-\rho_i y)$. We use exp to denote the exponential function to reduce confusion with the effort variable. The variable ρ_i denotes one's level of risk aversion. Assuming that females are more risk averse than males means $0 < \rho_M < \rho_F$.

We formally derive the result for this situation in the appendix.[17] The player who puts forth the greatest effort experiences the greatest performance since ability is the same for all players. In a competition between two same-gender players, equilibrium effort (and performance) is strictly decreasing in risk aversion. Since $\rho_M < \rho_F$, this implies that $p^*_{FvF} < p^*_{MvM}$. In a contest between two different type opponents with close-enough risk parameters, a players effort is strictly increasing in his opponent's risk aversion. It follows that $p^*_{MvF} > p^*_{MvM}$ and $p^*_{FvF} > p^*_{FvM}$. Together, this analysis implies the following result.

Result 7 *If females are relatively risk averse (i.e., $\rho_M < \rho_F$ and the values are not too different), then*

$$p^*_{FvM} < p^*_{FvF} < p^*_{MvM} < p^*_{MvF}.$$

This condition satisfies C1.

4 Additional Evidence

By comparing the equilibrium predictions from the theoretical analysis with the significant empirical results, C1 and C2, we are able to reject many of the explanations for the male advantage which have been put forth in the literature. We are unable to reject four possible explanations: (i) males are higher-ability competitors than females, (ii) females tend to be underconfident of their own abilities, (ii) males enjoy competition or care about winning more

[17]The analysis assumes that ρ_M and ρ_F are sufficiently close, an assumption that assures an interior pure-strategy equilibrium exists (see Skaperdas and Gan 1995).

than females, and (iv) males are less risk averse. To further assess these explanations for the male advantage, we consider additional evidence from our own analysis and the literature.

Both the CMP and aggregate data in Section 2 and Table 1 predict that males tend to perform better when competing against other males, and females tend to perform better when competing against other females. However, the ordering of MvM and MvF performance, and of FvF and FvM performance is not significant (p-value = 0.510 in the male case, and p-value = 0.242 in the female case). In addition to considering the effects of opponent gender separately for males and females, we also test whether opponent gender has a significant affect on performance in general, for males and females together. On average, across all players, performance is higher when competing against an opponent of the same gender (p-value = 0.107). As long as a model predicts that at least one of the genders performs better when competing against an opponent of the same gender, the model is consistent with this evidence. Any model that predicts that both males and females perform better when competing against opponents of the other gender, or that predicts that performance is independent of opponent gender is inconsistent with the evidence. This implies a second condition, which we refer to as "strongly preferred" given that its evidence is nearly significant.

C 2 (strongly preferred) *Either males, or females, or both perform better when competing against an opponent of the same gender.*

C2 is satisfied by the "ability," "enjoyment," and "risk-aversion" models. The "female underconfidence" model, which predicts that performance is independent of opponent gender, does not satisfy this condition. However, a variant of the female underconfidence story in which females are more underconfident when competing against males (e.g., Result 3) is consistent with C2.

A third condition is a stricter version of C2. It requires that opponent gender has the same affect on both male and female performance as we observe in the CMP and aggregate data. The evidence for this relationship, however, is not statistically significant. We therefore use the condition to highlight models that are most consistent with the empirical evidence, but we do not reject models that do not satisfy this requirement alone.

C 3 (preferred) *Males perform better against a male opponent. Females perform better against a female opponent.*

This condition is satisfied by the ability and enjoyment models. The risk-aversion model and both female underconfidence models satisfy the condition for females, but not for males. It should be noted, however, that the empirical evidence shows opponent gender has a less-significant effect on male performance (p-value = 0.510) than on female performance (p-value

= 0.242). Therefore, the risk-aversion model and the second underconfidence model violate C3 on the less-significant dimension.

A fourth condition places the blame for performance differences on high male performance rather than low female performance. The requirement is based on data from Cotton, McIntyre and Price (2010) who analyze how performance and gender differences change over multiple, sequential rounds of competition. They show that the gender gap disappears with exposure to competition, and present evidence that the decrease in the gender gap is due to a decrease in relative male performance rather than an increase in female performance. Similarly, Gneezy, Niederle and Rustichini (2003) and Gneezy and Rustichini (2004) show that competition causes males to increase their performance. This suggests a third condition on which to assess the models.

C 4 (preferred) *Higher male performance results from male overperformance, not female underperformance. The cause of male overperformance decreases with exposure to competition.*

A preferred explanation for the gender differences in competition should be consistent with this evidence. Some interpretations of the higher-male-ability story are not consistent with this evidence. C4 allows us to rule out ability-difference stories in which female underperform, including explanations involving stereotype threat, or in which competition increases female nerves or distraction. Other interpretations of the ability measure are consistent with C4. For example, male ability may be increased by a gender-specific boost in testosterone or adrenaline at the beginning of competition, and this boost may decrease with exposure to competition. The higher-male-enjoyment explanation is consistent with C3, as male enjoyment from competition may start off high then dissipate over multiple rounds of competition. The risk-aversion model is also consistent with C4 if males become more risk-averse over time. The female-underconfidence explanation, on the other hand, violates this preferred condition as it relies on female underperformance rather than male overperformance.

In another study, Günther et al. (in press, 2010) find that the male advantage is task dependent, as they show that it exists during maze competitions, but not during competitions involving word games, pattern matching, or memory tasks. The paper argues that performance differences exist in male-oriented tasks, but not in gender-neutral or female-oriented tasks.[18] In fact, Günther et al. (in press, 2010) presents some evidence that there exists a female advantage in competitions involving the female-oriented tasks. Similarly,

[18]Cotton, McIntyre and Price (2010) also finds no evidence of gender differences in competitions involving reading tasks.

Gneezy, Leonard and List (2009) present evidence that the the male advantage turns into a female advantage when competitions are held in a matrilineal societies. The higher-ability-male, higher-male-enjoyment, and risk-averse explanations for the gender differences are potentially consistent with this additional evidence. For example, males may experience a greater increase in adrenaline in competitions involving male-oriented tasks, and females may experience more adrenaline in female-oriented settings. Males may get extra enjoyment from competition only on male-oriented tasks, and females may get extra enjoyment from competition on female-oriented tasks. Also, males may only be less risk averse in competitions involving male-oriented tasks, and females may be less risk averse when they involve female-oriented tasks.

Only two explanations are consistent with all of the empirical evidence. These are the possibilities that male ability increases when faced with competition, and that males enjoy competition or care about winning more than females. Note that we are able to dismiss interpretations of the male-ability advantage that blame the performance gap on female underperformance rather than higher male performance. One consistent interpretation of the ability story involves males benefiting more from increases in adrenaline or testosterone, which have been shown in the medical literature to improve physical and mental performance capacity. A third explanation, that males are less risk averse than females, violates the less-significant dimension of C3, but is otherwise consistent with the evidence. Because of this, we do not feel comfortable dismissing the risk aversion story, and continue to view it as a feasible explanation. The female-underconfidence story, on the other hand, in inconsistent with C3 and C4.

5 Conclusion

The literature suggests a number of explanations for the male advantage during competition. We present a game theoretic framework in which to assess these explanations. We then compare the equilibrium predictions of the model with the empirical evidence about the relative performance of males and females, and the effects of opponent gender on performance. Doing so rules out a number of explanations for the male advantage, including explanations involving male overconfidence, misperceptions about male or female ability, and a number of explanations involving preference differences. Other explanations are consistent with the significant empirical evidence (i.e., C1), but inconsistent with all or part of the suggestive evidence (i.e., C2, C3, and C4). These include explanations involving female underconfidence, stereotype threat, and increases to female nervousness or distraction when faced with competition.

Two explanations are consistent with all of the empirical evidence we consider. These include:

- Male ability to perform increases in the face of competition, possibly due to increases in adrenaline or testosterone.

- Males enjoy competition or care about winning more than females.

A third explanation–that males are less risk averse than females–is consistent with almost all of the empirical evidence. The risk aversion model predicts that males perform better against females than against other males, which violates a single condition from the data for which significance is low. Future experimental work might be able to further limit the set of possibilities. It is also possible that all three of these feasible explanations contribute to the performance differences.[19]

Being able to narrow down the set of possible explanations for the male competitive advantage has implications for the policy debate. Policy makers should recognize that the performance differences during competition are likely not the result of females responding negatively to competition, but rather the result of males responding favorably to competition. They are also not likely due to misperception of relative ability, either because of self confidence differences or gender stereotypes. This suggests that efforts to expose females to competition in an effort to decrease underconfidence or improve misperceptions about ability differences may not have a significant effect on the gender gap. Further work may be done to gain better understanding of how the design of the contest, or the characteristics of the competitive workplace affect achievement differences between males and females. Some work has started to be done in this regard (e.g., Günther, Ekinci, Schwieren and Strobel in press, 2010, Cotton, McIntyre and Price 2010), but more is needed.

6 Appendix

More Cases of Preference Differences

We begin with the four cases in which opponent gender affects males and females in similar ways. Consider first the possibility that both males and females enjoy competition more against males than against females. This is equivalent to the case in which players earn a greater benefit from winning against a male. This means $\tau_H = \tau_{FvF} = \tau_{MvF}$ and $\tau_L =$

[19] Possibly some of the explanations we rejected are also present but are overshadowed by one of the preferred explanations and so do not provide a good description of the data.

$\tau_{FvM} = \tau_{MvM}$, and

$$\begin{array}{ll} e^*_{FvF} = \frac{1}{4\tau_H} & e^*_{MvM} = \frac{1}{4\tau_L} \\ e^*_{FvM} = \frac{\tau_H}{(\tau_L+\tau_H)^2} & e^*_{MvF} = \frac{\tau_L}{(\tau_L+\tau_H)^2}. \end{array}$$

From these values we can derive the following results.

Result 8 *If players value winning or enjoy competing more against male opponents (i.e., $\tau_{FvM} = \tau_{MvM} < \tau_{FvF} = \tau_{MvF}$), then*

$$p^*_{MvF} < p^*_{FvF} < p^*_{FvM} < p^*_{MvM}.$$

If players value winning or enjoy competing more against opponents of the other gender (i.e., $\tau_{FvM} = \tau_{MvF} < \tau_{FvF} = \tau_{MvM}$), then

$$p^*_{FvF} = p^*_{MvM} < p^*_{FvM} = p^*_{MvF}.$$

If males get extra enjoyment from competition against other males (i.e., $\tau_{MvM} < \tau_{MvF} = \tau_{FvM} = \tau_{FvF}$), then

$$p^*_{FvM} = p^*_{FvF} = p^*_{MvF} < p^*_{MvM}.$$

If males get less enjoyment from competition against females (i.e., $\tau_{MvM} = \tau_{FvM} = \tau_{FvF} < \tau_{MvF}$), then

$$p^*_{MvF} < p^*_{FvM} < p^*_{FvF} = p^*_{MvM}.$$

If males get more enjoyment from competition against males, and and less enjoyment from competition against females (i.e., $\tau_{MvM} < \tau_{FvF} = \tau_{FvM} < \tau_{MvF}$), then

$$p^*_{MvF} < p^*_{FvM} < p^*_{FvF} < p^*_{MvM}.$$

If males get more enjoyment from competition against males, and females get less enjoyment from competition against males (i.e., $\tau_{MvM} < \tau_{MvF} = \tau_{FvF} < \tau_{FvM}$), then

$$p^*_{FvM} < p^*_{MvF} < p^*_{FvF} < p^*_{MvM}.$$

Each of these possibilities violate C1. The interested reader may solve for other situations as well.

Differences in Risk Aversion

Here, we walk through the derivation of Result 7. When player i puts forth effort e_i, he earns expected utility

$$-\frac{ae_i}{ae_i + ae_{-i}} \exp[-\rho_i(Y + v - ce_i)] - \frac{ae_{-i}}{ae_i + ae_{-i}} \exp[-\rho_i(Y - ce_i)].$$

First order conditions with respect to e_i simplify to

$$e_{-i} + ce_i^2\rho_i + ce_i e_{-i}\rho_i + \exp[v\rho_i]e_{-i}((e_i + e_{-i})c\rho_i - 1) = 0. \tag{6}$$

Checking second order conditions assures that we are solving for a maximum. To find the symmetric equilibrium for the case when the two competitors are the same type, we solve the equation given $e_i = e_{-i}$, and ρ_i. This gives equilibrium effort in MvM and FvF contests.

$$e^*_{MvM} = \frac{\exp[v\rho_M] - 1}{2c\rho_M(1 + \exp[v\rho_M])}, \text{ and } e^*_{FvF} = \frac{\exp[v\rho_F] - 1}{2c\rho_F(1 + \exp[v\rho_F])}.$$

It is straightforward to show that e^* is strictly decreasing in ρ. Thus, the higher a gender's risk aversion, the lower it's effort and performance in a competition against a same-type opponent. Thus, $e^*_{MvM} > e^*FvF$.

Starting from an evenly matched contest, consider the effects of marginally increasing or decreasing one player's risk aversion parameter. To do this, we solve for $\partial e^*_i/\partial\rho_i$ and $\partial e^*_{-i}/\partial\rho_i$, evaluated at $\rho_i = \rho_{-i}$. In doing this, let Z_i denote the left hand side of player i's first order conditions, as given by Eq. 6. Take the derivative of both Z_i and Z_{-i} with respect to ρ_i, remembering that e^*_i and e^*_{-i} are functions of the two risk aversion parameters. Therefore, the resulting expressions for $\partial Z_i/\partial\rho_i$ and $\partial Z_{-i}/\partial\rho_i$ will include $\partial e^*_i/\partial\rho_i$ and $\partial e^*_{-i}/\partial\rho_i$. Since we are interested in the marginal effects of changing ρ_i starting from an even contest, we then set $\rho = \rho_i = \rho_{-i}$ and $e^* = e^*_i = e^*_{-i} = \frac{exp[v\rho]-1}{2(1+\exp[v\rho])c\rho}$. Solving the two resulting expressions for $\partial e^*_i/\partial\rho_i$ and $\partial e^*_{-i}/\partial\rho_i$ gives:

$$\frac{\partial e^*_i}{\partial\rho_i} = -\frac{\exp[2v\rho] - 2v\rho\exp[v\rho] - 1}{8c\rho^2(1 + exp[v\rho])^2}(\exp[v\rho_M] + 3), \text{ and}$$

$$\frac{\partial e^*_{-i}}{\partial\rho_i} = \frac{\exp[2v\rho] - 2v\rho\exp[v\rho] - 1}{8c\rho^2(1 + exp[v\rho])^2}(\exp[v\rho_M] - 1).$$

Since $\exp[2v\rho] > 2v\rho\exp[v\rho] + 1$, it follows that

$$\frac{\partial e^*_i}{\partial\rho_i} < 0, \text{ and} \frac{\partial e^*_{-i}}{\partial\rho_i} > 0.$$

If females are marginally more risk averse than males, then $e^*_{MvM} < e^*_{MvF}$ and $e^*_{FvM} < e^*_{FvF}$. Therefore,

$$p^*_{FvM} < p^*_{FvF} < p^*_{MvM} < p^*_{MvF}.$$

References

Baik, Kyung Hwan, "Effort Levels in Contests with Two Asymmetric Players," *Southern Economic Journal*, 1994, *61* (2), 367–378.

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

Boyatzis, Chris J., Michael Mallis, and Ileana Leon, "Effects of Game Type on Children's Gender-Based Peer Preferences: A Naturalistic Observational Study," *Sex Roles*, 1999, *40* (1-2), 93–105.

Brisswalter, Jeanick, Maya Collardeau, and Arceline Rene, "Effects of Acute Physical Exercise Characteristics on Cognitive Performance," *Sports Medicine*, 2002, *32* (9), 555–566.

Brown, Jennifer, "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars," April 2010. working paper.

Carpenter, Jeffrey, Peter Hans Matthews, and John Schirm, "Tournaments and Office Politics: Evidence from a Real Effort Experiment," *American Economic Review*, 2010, *100* (1), 504–517.

Chan, Willian, "External Recruitment versus Internal Promotion," *Journal of Labor Economics*, 1996, *14* (4), 555–570.

Chmura, J., H. Krysztofiak, A. W. Ziemba, K. Nazar, and H. Kaciuba-Uscilko, "Psychomotor performance during prolonged exercise above and below the blood lactate threshold," *European Journal of Applied Physiology*, 1998, *77* (1-2), 77–80.

Clark, C. R., G. M. Geffen, and L. B. Geffen, "Catecholamines and the covert orientation of attention in humans," *Neuropsychologia*, 1989, *27* (2), 131–139.

Cotton, Christopher, Frank McIntyre, and Joseph Price, "The Gender Gap Cracks Under Pressure: A Detailed Look at Male and Female Performance Differences During Competitions," July 2010. working paper.

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

Dixit, Avinash, "Strategic Behavior in Contests," *American Economic Review*, 1987, *77* (5), 891–898.

Gneezy, Uri and Aldo Rustichini, "Gender and Competition at a Young Age," *American Economic Review*, 2004, *94* (2), 377–381.

_ , Kenneth L. Leonard, and John A. List, "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society," *Econometrica*, 2009, *77* (5), 1637–1664.

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

Günther, Christina, Neslihan Arslan Ekinci, Christiane Schwieren, and Martin Strobel, "Women can't jump? An experiment on competitive attitudes and stereotype threat," *Journal of Economic Behavior and Organization*, in press, 2010.

Kivlighan, Katie T., Douglas A. Granger, and Alan Booth, "Gender differences in testosterone and cortisol response to competition," *Psychoneuroendocrinology*, 2005, *14*, 58–71.

Main, Brian G. M., Charles A. O'Reilly, and James Wade, "Top Executive Pay: Tournament or Teamwork?," *Journal of Labor Economics*, 1993, *11* (4), 606–628.

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

Nitzan, Shumuel, "Modelling rent-seeking contests," *European Journal of Political Economy*, 1994, *10*, 41–60.

O'Keefe, Mary, W. Kip Viscusi, and Richard J. Zeckhauser, "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics*, 1984, *1* (1), 27–56.

Salvador, A., F. Suay, E. Gonzalez-Bono, and M. A. Serrano, "Anticipatory cortisol, testosterone and psychological responses to judo competition in young men," *Psychoneuroendocrinology*, 2003, *28*, 364–374.

Skaperdas, Stergios and Li Gan, "Risk Aversion in Contests," *Economic Journal*, 1995, *105*, 951–962.

Stein, William E., "Asymmetric rent-seeking with more than two contestants," *Public Choice*, 2002, *113*, 325–336.

Tsoulouhas, Theofanis, Charles R. Knoeber, and Anup Agrawal, "Contests to become CEO: incentives, selection and handicaps," *Economic Theory*, 2007, *30* (2), 195–221.

Tullock, Gordon, "Efficient Rent Seeking," in James M Buchanan, Robert D Tollison, and Gordon Tullock, eds., *Towards a Theory of the Rent Seeking Society*, College Station: Texas A&M University Press, 1980.

Table 1: Normalized Performance Differences by Gender and Opponent Gender

	CMP	GNR	Aggregate
Gender Treatment			
FvM	-0.275	-0.454	-0.315
	[0.115]	[0.179]	[0.097]
FvF	-0.156	-0.118	-0.141
	[0.146]	[0.178]	[0.113]
MvF	0.052	0.353	0.146
	[0.120]	[0.179]	[0.100]
MvM	0.247	0.218	0.239
	[0.117]	[0.179]	[0.098]
T-Tests of Equality (p-values)			
FvM vs. FvF	0.588	0.185	0.242
FvM vs. MvF	0.065	0.001	0.001
FvM vs. MvM	0.002	0.008	0.0001
FvF vs. MvF	0.272	0.063	0.057
FvF vs. MvM	0.032	0.185	0.011
MvF vs. MvM	0.246	0.592	0.510

The top panel lists the mean performance of each gender and opponent gender combination. For example, FvM lists the average performance of females competing against males. Standard errors are listed in brackets. There are 253 observations in the CMP data and 120 observations in the GNR data. The aggregate data combines the two for 373 observations. The bottom panel reports the p-values from t-tests of equality between each pair of coefficients.

Table 2: Mann-Whitney U-tests of Normalized Performance Differences

	CMP	GNR	Aggregate
FvM vs. FvF	0.836	0.103	0.316
FvM vs. MvF	0.090	0.0004	0.0005
FvM vs. MvM	0.015	0.010	0.0004
FvF vs. MvF	0.219	0.026	0.020
FvF vs. MvM	0.072	0.135	0.021
MvF vs. MvM	0.433	0.563	0.764

The table reports p-values from Mann-Whitney U-tests of equality of the distributions between any two groups. The test, which is based on rank ordering, does not assume a particular distribution, but can give a rejection due to differences in higher moments than the mean. There are 253 observations in the CMP data and 120 observations in the GNR data. The aggregate data combines the two for 373 observations.