Queen's Economics Department Working Paper No. 1342

Which explanations for gender differences in competition are consistent with a simple theoretical model?

Christopher Cotton
Queen's University

Cheng Li
Mississippi State University

Frank McIntyre
Rutgers University

Joseph Price
Brigham Young University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

5-2015

WHICH EXPLANATIONS FOR GENDER DIFFERENCES IN COMPETITION ARE CONSISTENT WITH A SIMPLE THEORETICAL MODEL?

CHRISTOPHER COTTON, CHENG LI, FRANK MCINTYRE, AND JOSEPH PRICE

Abstract. A number of recent studies show that males may increase their performance by more than females in response to competitive incentives. The literature suggests that such a male competitive advantage may contribute to observed gender gaps in labor force pay and achievement. Understanding which factors may be driving these gender differences is essential for designing policies that promote equality. Using a game theoretic model of contests, we consider a variety of explanations for the male competitive advantage that have been proposed in the empirical and experimental literature. Comparing the testable predictions of the model with the empirical evidence from past papers, we reject explanations involving male over-confidence, misperceptions about relative ability, and some types of preference differences. Explanations involving female under-confidence and differences in risk aversion are consistent with the significant evidence. Two explanations provide perfect matches to observed performance patterns: (i) males are better than females at handling competitive pressure, and (ii) males enjoy competition more or have greater desire to win than females.

Date: **Current version: June 2015**.
Key words and phrases. contests; gender differences; effort and productivity
JEL: D74, J16, J24, J78.

Cotton: *corresponding author*, cotton@econ.queensu.ca, Department of Economics, Queen's University, Kingston, Ontario K7L 3N6. Li: cheng.li@msstate.edu, Department of Finance and Economics, Mississippi State University, Mississippi State, MS 39762. McIntyre: frank.mcintyre@business.rutgers.edu, Rutgers Business School, Newark, New Jersey 07102. Price: joe_price@byu.edu, Department of Economics, Brigham Young University, Provo, Utah 84602. Cotton acknowledges financial support associated with his position as the Jarislowsky-Deutsch Chair in Economic and Financial Policy at Queen's University. We appreciate feedback from Christopher Parmeter, Laura Giuliano and Brennan Platt.

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 a contest with other students. They show that competition causes males to increase their performance by more than females. Gneezy and Rustichini (2004) find similar results in footraces between young children: males increase their performance in the face of competition, while females do not. Cotton, McIntyre and Price (2013) 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). Understanding which factors may be driving these gender differences is therefore essential for designing policies to promote equality.

The experimental and empirical literature suggest a variety of factors which may drive the observed male performance advantage in competition. Possibilities involve real or perceived differences in confidence, ability, or risk aversion (see Gneezy, Niederle and Rustichini 2003, Niederle and Vesterlund 2007, 2011, Gneezy and Rustichini 2004, Croson and Gneezy 2009, Günther et al. 2010, Ong and Chen 2012). None of these articles provide a game theoretic framework to assess the merits of the explanations. However, there exists an extensive theoretical literature modeling contests and tournaments that can provide insight into the causes of the observed gender differences. By comparing theoretical predictions with the empirical observations, we can learn which of the proposed explanations for the observed gender differences are consistent with the theoretical predictions.

The current article presents a simple model of competition adapted from the game theoretic literature on contests. In the model, agents simultaneously choose effort, where their performance in the contest is a function of both effort and ability, with the probability of winning the contest 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 between asymmetric players.[2] We adapt the model to consider various explanations of the male advantage, allowing for male competitors to differ from female competitors in terms of their preferences, confidence and ability; in an extension

[1]Various recent articles work to determine in which settings the male advantage exists. Cotton, McIntyre and Price (2013) 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. (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 where the male advantage does exist.

[2]See also Dixit (1987), Nitzan (1994), Stein (2002), and Brown (2011). Skaperdas and Gan (1995) incorporates differences in risk aversion into such contests. Application of contest theory to workplace achievement include O'Keeffe, Viscusi and Zeckhauser (1984), Main, O'Reilly and Wade (1993), Chan (1996), Tsoulouhas, Knoeber and Agrawal (2007) and Carpenter, Matthews and Schirm (2010).

we consider differences in risk aversion. In order to isolate the effects from each individual explanations on performance differences, we consider each possible explanation separately. For example, when considering explanations involving ability differences, the model incorporates differences in ability while holding differences in preferences and confidence fixed. When considering explanations involving differences in risk aversion, the model considers a game in which males and females differ only in their risk aversion. To further simplify the analysis, we also assume that the two groups are homogeneous. Although we model differences between males and females, we abstract from within-gender differences amongst males or amongst females. The result is a simple framework that allows for the straightforward comparison of different explanations for the male-performance advantage that exists in certain environments; it is not a general theory of gender differences in competition.

In equilibrium of our games, a competitor's effort and performance depends on his or her own type, as well the type of his or her opponent. The analysis is therefore concerned about the relative performance of four player types—males in single-gender competition of males only (MvM), males in mix-gender competition of half females and half males (MvF), females in mix-gender competition of half females and half males (FvM), and females in single-gender competition of females only (FvF)—and whether the predicted performance differences between these four groups matches the patterns observed in the data. We solve cases with two competitors and six competitors separately, showing how the number of competitors affects the predicted performance patterns in certain settings.

We compare the theoretical predictions of the model with the empirical evidence from Gneezy, Niederle and Rustichini (2003) (which involved six competitor experiments) and Cotton, McIntyre and Price (2013) (which involved head to head contests between two competitors). We chose these papers due to data availability, because Gneezy, Niederle and Rustichini (2003) is the seminar paper on the topic, and because Cotton, McIntyre and Price (2013) involved an experiment involving head-to-head contests that closely resemble the underlying theoretical model with two players. The data allow us to draw conclusions about the relative performance of the four player types: MvM, MvF, FvM, and FvF. We review the data in detail in Section 2. In that section, we describe the criterion that we use to determine how well alternative theoretical explanations match the data.

The theoretical model predicts behavior that, while well known amongst contest theorists, may go against popular intuition. For example, the model shows that competitors put in the most effort in contests in which they are evenly matched against a single 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 than his opponent, but rather because he achieves higher performance with equal 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, for example, a player is over-confident in his own ability, 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 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 over-confident about their abilities and hence their chances of actually winning the tournament" (p 1060). This statement and others found in the literature are 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. 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.

We find that other explanations for the male performance advantage are more consistent with the theory, to different degrees. When females are under-confident in their own abilities or more risk-averse, they tend to underperform compared to males. While this outcome is consistent with the most significant empirical evidence, it does not predict the exact same effect of opponent gender on performance as observed in the data. Because of this, we view the female under-confidence and higher female risk aversion models as feasible but moderately less-likely explanations of the performance patterns compared to models which perfectly predict the performance differences observed in the data.[3]

Two explanations are perfectly aligned with the empirical patterns. First, males may be better at dealing with competitive pressures. This explanation does not imply that males are inherently better at solving math questions, competing mazes, or running races.[4] Rather, it means that they remain more focused or less nervous in the face of competition, and are therefore better able to convert effort into performance. Second, males may have greater desire to win, or enjoy competition more than females.

Taken together, our results may help guide policy intended to improve gender equality in school, the workplace, and other potentially competitive environments. We find

[3]Although there is substantial empirical evidence that "women are indeed more risk averse than men" (Croson and Gneezy 2009, p448) in a variety of settings, we are the first to illustrate how risk aversion alone may result in the within competition performance differences between males and females. Dohmen and Falk (2011) empirically show that females are no less likely to sort into competition than males when controlling for differences in risk aversion. However, their analysis focuses on selection into contests, rather than performance in competition. Ong and Chen (2012) argue that risk aversion may be the dominate factor for explaining performance differences within a contest.

[4]Cotton, McIntyre and Price (2013) and Gneezy and Rustichini (2004) find gender differences even when controlling for a student's performance in a non-competitive setting.

that the observed gender differences are unlikely driven by incorrect beliefs about the relative ability of males and females. Therefore, campaigns to inform people of the fact that females are just as capable as males will be ineffective in eliminating the performance differences. To the extent that performance differences are also not driven by female under-confidence in their own abilities, working to improve the self confidence of females is unlikely to effectively reduce the observed performance differences. More effective policies may aim to give females greater exposure to competitive environments at earlier ages, with the intention of improving their ability to deal with the pressures of competition.

The contribution of the paper is threefold. First, we provide a theoretical framework in which to assess between-group performance differences in reaction to competition. Although our focus is on gender differences, the model may be applied to help explain performance differences in settings where player type is defined by some other characteristic rather than gender. Just as explanations involving overconfidence, for example, cannot explain performance differences between males and females, it will also be unable to explain average performance differences between other demographic or socioeconomic groups. The framework may also be extended to consider additional explanations for gender differences in competition that we did not consider here. Second, we use the model to provide a formal assessment of a number of popular explanations for the male performance 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 that have been put forth in the literature. These results can help guide policies intended to decrease gender inequality in graduate school or the workforce. Third, our model provides guidance for future researchers in designing new experiments that will differentiate among the alternative explanations that are consistent with the data.

2. Empirical Observations

A significant literature has emerged testing whether males and females react differently to competitive incentives and in contests. Gneezy, Niederle and Rustichini (2003) run a series of experiments on college students and show that males respond more favorably than females to competition when solving mazes. Gneezy and Rustichini (2004) produce similar results when they have primary school students run footraces. Similarly, Günther et al. (2010) identify a male advantage in maze competitions, Shurchkov (2012) identifies a male advantage in high-pressure contests, and Cotton, McIntyre and Price (2013) identify a male advantage the first time students compete in a math contest; although all three of these papers also identify contest settings in which there is no evidence of a male advantage (e.g., female oriented tasks, low-pressure situations, and repeated interactions).[5] Niederle and Vesterlund (2011) provide an excellent survey of the literature on gender differences in competition. They summarize the literature: "It therefore seems that the task and maybe the way the task is administered may matter

[5]Günther et al. (2010) find no significant difference between male and female performance in contests involving word games; Cotton, McIntyre and Price (2013) show that the male advantage disappears after an initial contest round, and does not exist when the contest is not framed as a race. Shurchkov (2012) finds that the male advantage disappears in low-time-pressure contests.

when looking for gender differences in changes in performance across incentive schemes. However, in a stereotypical-male task, gender differences in competitiveness have been confirmed."

Our theoretical analysis makes predictions about performance patterns which can be used to assess different explanations for the male advantage for settings where it does exist. In this section, we provide a detailed look at some of the empirical evidence from the literature. Our detailed review of the evidence focuses on the two experiments for which we have individual level data: Gneezy, Niederle and Rustichini (2003) and Cotton, McIntyre and Price (2013).[6] Access to the individual level data allows us to aggregate the data across the two experiments to make stronger predictions about the relative performance of the four gender combinations—males in single-gender competition (MvM), males in mix-gender competition of half females and half males (MvF), females in mix-gender competition of half females and half males (FvM), and females in single-gender competition (FvF).

The CMP data involves 253 total participants (136 males and 117 females) competing in contests to solve math problems 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.

Our model gives predictions about the rank ordering of performance for males and females in single-gender and mix-gender competitions. 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, as the theoretical predictions sometimes depend on whether there are six competitors (as in GNR) or two competitors (as in CMP). The second panel provides p-values for a set of exhaustive t-tests comparing the means of each of the groups. Table 3 presents the same analysis using MWU tests rather than t-tests. 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.

Both data sets present evidence that males on average outperform females in competitions. In GNR, the best performance comes from the MvF group, followed by MvM, then FvF, and finally FvM. In CMP, the best performance comes from the MvM group,

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

[7]Cotton, McIntyre and Price (2013) also presents data from multiple rounds of competition. They show that after the first round of competition, the gender differences disappear. Because we are interested in settings in which the gender differences do exist, 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]The paper also includes data from non-competitive treatments, which we do not include in our analysis.

Table 1. Normalized Performance Differences by Gender and Opponent Gender

	CMP	GNR
Gender Treatment		
FvM	-0.275	-0.454
	[0.115]	[0.179]
FvF	-0.156	-0.118
	[0.146]	[0.178]
MvF	0.052	0.353
	[0.120]	[0.179]
MvM	0.247	0.218
	[0.117]	[0.179]
T-Tests of Equality (p-values)		
FvM vs. FvF	0.588	0.185
FvM vs. MvF	0.065	0.001
FvM vs. MvM	0.002	0.008
FvF vs. MvF	0.272	0.063
FvF vs. MvM	0.032	0.185
MvF vs. MvM	0.246	0.592

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 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
FvM vs. FvF	0.836	0.103
FvM vs. MvF	0.090	0.0004
FvM vs. MvM	0.015	0.010
FvF vs. MvF	0.219	0.026
FvF vs. MvM	0.072	0.135
MvF vs. MvM	0.433	0.563

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.

followed by MvF, then FvF, and finally FvM. 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. This brings us to the first condition, which we compare to the predictions of the theoretical models.

C 1 (Perfect match)**.** *A model is a perfect match with the data if it predicts the exact same performance ordering observed in the CMP data when $N = 2$:*

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

AND predicts the exact same performance ordering observed in the GNR data when $N = 6$:

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

We say that a model is a perfect match with CMP (GNP), when it satisfies the perfect match criteria for $N = 2$ ($N = 6$). It is feasible that a model is a perfect match for one experiment but not the other.

Although both of our experimental data sets present evidence that males on average outperform females during contests, the results are not always significant when we consider the joint effects of both own gender and opponent gender. In GNR, 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). Other insignificant relationships in GNR include the difference between MvF and MvM performance (p-value = 0.592) and the difference between FvM and FvF (p-value = 0.185). In CMP, FvM is also 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). Other insignificant relationships in CMP include the difference between MvF and MvM performance (p-value = 0.246) and the difference between FvM and FvF (p-value = 0.588).

Because some of the relationships in C1 are not significant, we do not require that a model predicts that the insignificant relationships to be considered a feasible match.

C 2 (Feasible match)**.** *A model is a feasible match for the empirical data if it violates some part of C1, but predicts all of the significant performance orderings observed in the CMP data when $N = 2$:*

$$p^*_{FvM} < p^*_{MvM} \quad AND \quad p^*_{FvM} < p^*_{MvF} \quad AND \quad p^*_{FvF} < p^*_{MvM}$$

AND predicts all of the significant performance orderings observed in the GNR data when $N = 6$:

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

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.

3.1. **A Simple Contest.** N players, indexed $i \in \{1, 2, ..., N\}$, engage in competition for a prize. The players simultaneously and 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 each player, 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 positive 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 captures differences in both cost of effort and value of winning, which are indistinguishable in the model.

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 $p_i / \sum_{j=1}^{N} p_j$ if $p_i > 0$ and 0 when $p_i = 0$. Therefore, a player's expected utility equals

$$Eu_i = \frac{a_i e_i}{\sum_{j=1}^{N} a_j e_j} - \tau_i e_i. \tag{1}$$

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. To ensure an interior solution, we assume that

$$\frac{\tau_i}{a_i} < \frac{1}{N-1} \sum_{j=1}^{N} \frac{\tau_j}{a_j} \quad \forall i = 1, 2, ..., N. \tag{2}$$

Thus, no player faces such a large cost of effect that he chooses not to participate in equilibrium.

The first order condition of Eu_i with respect to e_i is

$$\frac{a_i(\sum_{j=1}^{N} a_j e_j - a_i e_i)}{(\sum_{j=1}^{N} a_j e_j)^2} - \tau_i = 0. \tag{3}$$

Solving the system of N equations for $\{e_1, ..., e_N\}$ gives the equilibrium solution, where each player's effort maximizes his own payoff conditional upon the equilibrium effort of the other players.[9] We find that for each player,

$$e_i^* = \frac{1}{a_i} \frac{N-1}{\sum_{j=1}^{N} (\tau_j / a_j)} - \frac{\tau_i}{a_i^2} \left(\frac{N-1}{\sum_{j=1}^{N} (\tau_j / a_j)} \right)^2. \tag{4}$$

From this expression, we can also determine equilibrium performance, $p_i^* = a_i e_i^*$.

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, or is less susceptible to stereotype threat.

[9]A check of the second order conditions assures that this achieves a maximum.

It is worth pointing out two results regarding the equilibrium solution that are helpful in the later analysis. Both are consistent with formal results 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 other demographic or socioeconomic characteristics. To keep the analysis as straightforward and intuitive as possible, we assume that all players of the same gender share the same values of τ_t and a_t.[10]

From here forward, we focus on cases where there are two players ($N = 2$) or six players ($N = 6$). In each case, we consider situations where there are all males, all females, or an even split. The $N = 6$ case corresponds to the experiment in Gneezy, Niederle and Rustichini (2003) (henceforth, GNR) where there are six competitors in each group, and the $N = 2$ competition corresponds to Cotton, McIntyre and Price (2013) (henceforth, CMP) and other experiments that consider person contests.

We formally define a player's type in terms of both one's own gender and opponent gender, as the analysis is concerned with the effect of both. Denote a player's type by $t \in \{FvF, FvM, MvF, MvM\}$, where the first letter denotes one's own gender. FvF denotes single-gender competition of females only, MvM denotes single-gender competition of males only, and FvM and MvF denote mix-gender competition of half females and half males.[11] When a variable is independent of opponent gender, we may simply use F and M to denote player type.

We can simplify (4) further, given the restrictions to $N \in \{2, 6\}$, the equal proportion of males and females in mixed gender contests, and homogeneity within gender.

When $N = 2$,

$$e^*_{MvM} = \frac{1}{4\tau_M}, \quad e^*_{MvF} = \frac{\tau_F a_M a_F}{(\tau_M a_F + \tau_F a_M)^2}, \quad e^*_{FvF} = \frac{1}{4\tau_F}, \quad e^*_{FvM} = \frac{\tau_M a_M a_F}{(\tau_M a_F + \tau_F a_M)^2}.$$

[10]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.

[11]In the 2 player competition, FvM and MvF denote mix-gender competition of one female and one male. In the 6 player competition, FvM and MvF denote mix-gender competition of 3 females and 3 males.

When $N = 6$,

$$e^*_{MvM} = \frac{5}{36\tau_M}, \quad e^*_{MvF} = \frac{5}{9} \cdot \frac{a_F(3\tau_F a_M - 2\tau_M a_F)}{(\tau_M a_F + \tau_F a_M)^2},$$

$$e^*_{FvF} = \frac{5}{36\tau_F}, \quad e^*_{FvM} = \frac{5}{9} \cdot \frac{a_M(3\tau_M a_F - 2\tau_F a_M)}{(\tau_M a_F + \tau_F a_M)^2}.$$

We walk through the derivation of the above equations for equilibrium effort in the appendix. In the following subsections, we adapt this framework to consider various explanations of gender differences in performance.

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 isolate the effect of ability differences, here we assume $\tau = \tau_M = \tau_F$.

In a male/female competition, the players differ in terms of a but not τ. We can simplify the expressions for equilibrium effort. When $N = 2$,

$$e^*_{MvM} = e^*_{FvF} = \frac{1}{4\tau}, \; e^*_{MvF} = e^*_{FvM} = \frac{a_M a_F}{\tau(a_M + a_F)^2}.$$

When $N = 6$,

$$e^*_{MvM} = e^*_{FvF} = \frac{5}{36\tau}, \; e^*_{MvF} = \frac{5}{9}\frac{a_F(3a_M - 2a_F)}{\tau(a_M + a_F)^2}, \; e^*_{FvM} = \frac{5}{9}\frac{a_M(3a_F - 2a_M)}{\tau(a_M + a_F)^2}.$$

Males and females put in equal effort in single gender competitions, and in the case of $N = 2$, males and females also put in the same effort in a mixed gender competition. This goes against standard intuition. This is because the marginal return on effort depends on how competitive the contest is. If one player is much more able than his opponent, 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, as given by $p^* = a_i e_i^*$.

We assume that a_F is consistent with (2), which implies that males have higher ability, but not such high ability compared to females that they choose not to participate. In the case of $N = 2$, this is never a concern. In the case with $N = 6$, we require $(2/3)a_M < a_F$. Given this, it is straightforward to compare 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*

$$p^*_{FvM} < p^*_{FvF} < p^*_{MvF} < p^*_{MvM} \;\; \textit{when } N = 2,$$

$$p^*_{FvM} < p^*_{FvF} < p^*_{MvM} < p^*_{MvF} \;\; \textit{when } N = 6.$$

Notice that the effect of opponent gender on performance depends on the number of competitors. Males are predicted to exert more effort and perform better when the other competitor is also male in the case of head to head competition. But, males are predicted to exert less effort when all other competitors are male in a six person competition. Therefore, the predictions of the model depend on the number of players.

The reason for this is because the incentive that someone of the highest ability has to exert effort is highest when there is a competitor of similar ability, which means he cannot only rely on an ability advantage and expect to win. But, his incentive to exert effort is decreasing in the number of additional competitors (beyond the first) who are also of similar ability. This is because the additional competitors reduces the probability that he wins even if he puts in additional effort, which reduces his incentive. In the game with six players, a high ability player exerts more effort when he faces only two other high ability competitors, than he does when all five other competitors are also high ability. The result would carry over if we alternatively assumed that mixed gender groups involved two people of one gender and four people of another.

Result 1 is consistent with condition C1 as described in Section 2. The differences in performance predicted in the $N = 2$ and $N = 6$ cases are perfectly aligned with the differences in performance patters observed in the CMP and GNR data.

That is, a model in which males have an ability advantage perfectly match the performance patterns observed in CMP and GNR. An ability advantage may refer to either differences in one's absolute ability to compete a task, or differences in how well one deals with the pressures of competition (e.g., how able one is to convert effort to performance when competing). Although the model does not distinguish between these factors, we know that task-ability alone cannot explain the gender differences. The literature shows, for example, that if you take males and females who are equally good at running or solving math problems outside of competition, and have them do the same task in a competition, then the male students will typically outperform the female students of similar underlying ability (Gneezy and Rustichini 2004, Cotton, McIntyre and Price 2013). This suggests that the ability measure in our analysis must at least partially capture the ability to handle pressures of competition.

3.3. **Under- or Overconfidence.** The second set of explanations for the competitive performance gap that we consider involve female under-confidence 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 *under-confident*. If $\hat{a}_i > a_i$, then i is said to be *over-confident*. 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.[12] 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 have the same level of confidence. This means that in equilibrium,

$$e_i^* = \frac{a}{\tau}\frac{N-1}{(N-1)\hat{a}_i + a} - \frac{a^2}{\tau}\left(\frac{N-1}{(N-1)\hat{a}_i + a}\right)^2 \quad \text{and} \quad p_i^* = ae_i^*,$$

where performance equals *actual* ability times effort.

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

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 $\hat{a}_i < a$, then e_i^* is strictly increasing in $\hat{a}_i$; and if $\hat{a}_i > a$, then e_i^* is strictly decreasing in $\hat{a}_i$.

If males are over-confident in their ability, then $\hat{a}_M > a$, and the equilibrium expression for e_M^* and e_F^* simplify further.

When $N = 2$,

$$e_M^* = \frac{a\hat{a}_M}{\tau(5\hat{a}_M + a)^2} \text{ and } e_F^* = \frac{1}{4\tau}.$$

When $N = 6$,

$$e_M^* = \frac{5(5a\hat{a}_M - 4a^2)}{\tau(5\hat{a}_M + a)^2} \text{ and } e_F^* = \frac{5}{36\tau}.$$

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

Result 2. *If males are over-confident (i.e., $\hat{a}_F = a < \hat{a}_M$ and $\tau_M = \tau_F$), then when either $N = 2$ or $N = 6$*

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

If females are under-confident in their ability, then $\hat{a}_F < a$, and simplified expressions for e_t^* may be calculated in similar fashion. Deriving and comparing p_t^* for these cases is similarly straightforward.

Result 3. *If females are under-confident (i.e., $\hat{a}_F < a = \hat{a}_M$ and $\tau_M = \tau_F$), then when either $N = 2$ or $N = 6$,*

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

The above results are consistent with the idea that players put in more effort in competitions they believe to be evenly matched. If players believe that a competition is lopsided, they will put in less effort regardless of whether they are the player with an advantage. In equilibrium, over-confident males decrease their performance just as under-confident females decrease theirs.

Of the two confidence stories, only the female under-confidence model is feasible, satisfying C2. However, it does not predict the exact same performance ordering, as it predicts that both male and female performance is independent of opponent gender, failing C1. Because of this, we view of the female under-confidence model as a possible, but less consistent, reason for observed gender differences in the competition.

Male over-confidence, on the other hand, is completely inconsistent with the observed performance patterns in both GNR and CMP. The theory illustrates how over-confident males exert too little effort and end up performing worse than females, which is the opposite pattern than what we observe in the data. It represents a remarkably poor match for the empirical evidence.

3.4. **Perceived Differences in Ability.** A related explanation for performance differences between males and females is that there is a common belief that males tend to be 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. It is also not the same as differences in self-confidence studied above, 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, beliefs affect effort and performance only when competing against an opponent of another gender. Therefore, when $N = 2$,

$$e^*_{MvM} = e^*_{FvF} = \frac{1}{4\tau}.$$

and when $N = 6$,

$$e^*_{MvM} = e^*_{FvF} = \frac{5}{36\tau}$$

Consider the possibility that all players underestimate female ability. In the game with $N = 2$, both males and females put in the same effort when facing an opponent of the opposite gender, with

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

This effort is lower than it would be if the players did not underestimate female ability. Because both genders have the same true underlying ability, the performance of males and females is the same.[13]

In the game with $N = 6$, males put in more effort than females, as they compete against other males within their group for the prize. Females within the group exert less effort, as they (incorrectly) expect that their performance will be worse than males, even if they put in the same effort. In this case,

$$e^*_{MvF} = \frac{5}{9}\frac{a(3a - 2\hat{a}_F)}{\tau(\hat{a}_F + a)^2} \quad \text{and} \quad e^*_{FvM} = \frac{5}{9}\frac{\hat{a}_F(3\hat{a}_F - 2a)}{\tau(\hat{a}_F + a)^2}.$$

We assume that $\hat{a}_F$ is consistent with (2), which implies that players believe that males have higher ability, but not such high ability compared to females that they choose not to participate. In the case of $N = 2$, this is never a concern. In the case with $N = 6$, we require $(2/3)a < \hat{a}_F$. Given this, is straightforward to compare equilibrium performance for each competitor-opponent combination.

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

$$\begin{aligned} p^*_{FvM} = p^*_{MvF} < p^*_{FvF} = p^*_{MvM} &\quad \text{when } N = 2, \\ p^*_{FvM} < p^*_{FvF} = p^*_{MvM} < p^*_{MvF} &\quad \text{when } N = 6. \end{aligned}$$

This result assumes that both males and females underestimate female ability or over estimate male ability. In the appendix, we consider additional possibilities, including cases where only one group has incorrect beliefs.

[13] Interestingly, both males and females are better off in this environment compared to one in which players have accurate beliefs.

Result 4, as well as the similar explanations involving incorrect beliefs that we consider in the appendix, are inconsistent with both C1 and C2. This allows us to rule out explanations in which performance differences are caused by *perceived* differences in ability.

3.5. **Differences in Enjoyment or Costs.** 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 intrinsic value of winning (i.e., $v_M > v_F$).[14] Both of these cases affect the preference parameter $\tau_i \equiv c_i/v_i$ in the same way, implying that $\tau_M < \tau_F$. Here, we consider the possibility that performance differences are driven by differences in τ. 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

$$e_i^* = \frac{N-1}{\sum_{j=1}^N \tau_j} - \tau_i \left(\frac{N-1}{\sum_{j=1}^N \tau_j} \right)^2 .$$

Equilibrium effort simplifies as it did in the previous sections. When $N = 2$,

$$e_{MvM}^* = \frac{1}{4\tau_M}, \quad e_{FvF}^* = \frac{1}{4\tau_F}, \quad e_{MvF}^* = \frac{\tau_F}{(\tau_M + \tau_F)^2}, \quad e_{FvM}^* = \frac{\tau_M}{(\tau_M + \tau_F)^2}.$$

and when $N = 6$,

$$e_{MvM}^* = \frac{5}{36\tau_M}, \quad e_{FvF}^* = \frac{5}{36\tau_F}, \quad e_{MvF}^* = \frac{5}{9}\frac{3\tau_F - 2\tau_M}{(\tau_M + \tau_F)^2}, \quad e_{FvM}^* = \frac{5}{9}\frac{3\tau_M - 2\tau_F}{(\tau_M + \tau_F)^2}.$$

We assume that τ_M is consistent with (2), which implies that males have lower cost of effort, but not such low cost of effort compared to females that they choose not to participate. In the case of $N = 2$, this is never a concern. In the case with $N = 6$, we require $(2/3)\tau_F < \tau_M$. Comparing performance under this effort gives the following result.

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*

$$p_{FvM}^* < p_{FvF}^* < p_{MvF}^* < p_{MvM}^* \quad \text{when } N = 2,$$
$$p_{FvM}^* < p_{FvF}^* < p_{MvM}^* < p_{MvF}^* \quad \text{when } N = 6.$$

Result 5 satisfies C1, providing an excellent match with the empirical evidence. Therefore, the observed gender differences in competitive performance may be driven by males having a higher intrinsic value of winning a contest, or getting more enjoyment out of competition (and therefore finding effort less costly) compared to females.[15]

[14] In both experiments, males and females received the same prizes for winning, therefore differences in value of winning should be interpreted as differences in the inherent, non-monetary value of winning.

[15] 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

3.6. **Players Dislike Losing to Females.** We extend the model of competition to consider the possibility that males dislike losing to females. Suppose that $a_M = a_F = a$ and $\tau_M = \tau_F = \tau$. New to the analysis is the assumption that when a males loses to a female, he experiences a loss of utility equal to $h > 0$.

In single gender competition, h does not play a role, and $e^*_{MvM} = e^*_{FvF}$ equals $1/(4\tau)$ when $N = 2$, and $5/(36\tau)$ when $N = 6$.

In the mixed gender contests, males exert additional effort to decrease the probability of losing to a female. Females, on the other hand, do not have any incentive to exert additional effort, as they have no aversion to losing to a male. They respond by putting in less effort, anticipating that winning is more difficult given the additional effort exerted by the males. In the mixed gender competition, when $N = 2$,

$$e^*_{MvF} = \frac{(1+h)^2}{(2+h)^2\tau} > \frac{1}{4\tau} \text{ and } e^*_{FvM} = \frac{1+h}{(2+h)^2\tau} < \frac{1}{4\tau},$$

and when $N = 6$,

$$e^*_{MvF} = \frac{(1+3h)(5+9h)}{9\tau(2+3h)^2} > \frac{5}{36\tau} \text{ and } e^*_{FvM} = \frac{5+9h}{9\tau(2+3h)^2} < \frac{5}{36\tau}.$$

A similar analysis can be used to determine performance differences when all players, including females, dislike losing to females.

Result 6. *If males dislike losing to a female, then when either $N = 2$ or $N = 6$,*

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

If all players dislike losing to a female, then when either $N = 2$ or $N = 6$,

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

When players dislike losing to females, the model predicts performance differences that are inconsistent with the empirical evidence, violating C1 and C2.

3.7. **Differences in Risk Aversion.** A final explanation we consider involves differences in risk aversion between males and females. 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 $a = a_M = a_F$, and $\tau = \tau_M = \tau_F$. Competitors have the same initial resource $Y > 0$, and as before the value of the prize is normalized to 1. Let y denote his payout at the end of the competition, where $y = Y + 1 - \tau e_i$ if he wins, and $y = Y - \tau e_i$ 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 $0 < \rho_M < \rho_F$.

ordering in Result 5 continues to hold if males enjoy competition more than females, and experience extra enjoyment when competing against other males. 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).

We formally derive the result for this situation in the appendix.[16] 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 player's 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 moderately more risk averse than males (i.e., $\rho_M < \rho_F$ and the values are not too different), then when either $N = 2$ or $N = 6$,*

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

This result predicts the exact performance pattern observed in the GNR, and all of the significant performance orderings observed in the CMP. It does not, however, predict the same performance effect of opponent gender on male performance that we observe in the CMP data. Therefore, differences in risk aversion satisfies C2, but fails C1 in one of the data sets. We therefore see differences risk aversion as a feasible explanation for the observed performance patters.

3.8. **Additional Explanations.** In the appendix, we rule out a number of additional explanations involving incorrect beliefs about ability and preference based explanations. These include situations in which only males or only females underestimate females or overestimate males, and situations where one's preference parameter τ depends on opponent gender.

4. Summary and Conclusion

The literature suggests a number of explanations for the male advantage during competition. We present a game theoretic framework to model these explanations, and compare the model predictions with 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. Explanations that involve female under-confidence and differences in risk aversion are consistent with the most significant evidence, but fail to predict the exact same performance orderings observed in the data. Because of this, we view the female under-confidence story and the differences in risk aversion story as feasible but not perfectly aligned explanations of the performance patterns.

Table 3 summarizes how well each of the models is aligned with the empirical observations from GNR and CMP.

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

[16] 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). The appendix also provides a numberical analysis to show that the results carry over for larger differences in ρ.

TABLE 3. Alignment of theory and evidence

	Perfect Alignment to CMP	Perfect Alignment to GNR	Feasible Match
Males are higher ability	**yes**	**yes**	**yes**
Males are over-confident	no	no	no
Females are under-confident	no	no	**yes**
Players underestimate females or overestimate males	no	no	no
Males get greater enjoyment or have lower costs	**yes**	**yes**	**yes**
Players dislike losing to females	no	no	no
Females are more risk-averse	no	**yes**	**yes**

the performance differences in competitive settings are unlikely due to misperceptions about the abilities of oneself or others. Therefore, policies aimed to eliminate stereotypes, for example, may be unlikely to effectively reduce the observed performance differences. More effective policies may aim to give females greater exposure to competitive environments at earlier ages, with the intention of improving their ability to deal with the pressures of competition.

Additional experimental work may be able to narrow down the set of possibilities further. Better data, perhaps collected from a larger scale experiment, may allow for more precise estimates regarding the effects of opponent gender on performance. Empirically showing that males perform better when competing against males in a two player contest, for example, would rule out the female under-confidence and the risk aversion explanation.

Our theoretical analysis has focused on a setting with no heterogeneity within gender groups, i.e. all males are similar and all females are similar. This assumption assures that the only asymmetries that arise in the model arise at the group level, which keeps the analysis as straightforward possible, and allows us to focus on developing the fundamental intuition. With that said, such assumptions may be relaxed in future work in order to further examine within gender performance differences. Allowing ability heterogeneity within gender groups would allow an analysis to consider ability-contingent explanations of the observed performance patterns. For example, one could consider situations in which only high ability females are under confident or low ability males are over-confident. Our data is not precise enough for estimating within group performance distributions, and a significantly larger experiment would be needed. We therefore reserve such considerations for future research.[17]

[17]Existing evidence does suggest that high and low ability females may respond differently to competitive pressures. Anbarci, Arin and Lee (2014) for example show that high ability females may respond in similar ways as males to changing incentive. A look at the dynamic performance patterns in Cotton, McIntyre and Price (2013) reveals that the first period male competitive advantage may be driven by

Finally, our analysis has focused exclusively on performance within a contest. The literature on gender differences in competitive settings is also concerned with differences in willingness to select into competitive environments (e.g. Niederle and Vesterlund 2007, Dohmen and Falk 2011). If we incorporated a selection stage at the beginning of our games, those who anticipated lower competitive performance would be less willing to enter the contest, at least in the models in which players have correct beliefs about their performance. This suggests that differences in competitive ability, enjoyment of competition or desire to win, or risk aversion could all be responsible for females shying away from competition. However, such an analysis assumes that players will enter the contest whenever it is rational to do so, and therefore provides no insight into why females may compete too little, or males may compete too much. Future research may consider these issues further.

References

Anbarci, Aejat, K. Peren Arin, and Jungmin Lee. 2014. "Gender differences in response to contingent rewards: Evidence from a natural experiment of junior tennis." *Journal of Behavioral and Experimental Economics*, 53: 131–137.

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

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

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

Brown, Jennifer. 2011. "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars." *Journal of Political Economy*, 119: 982–1013.

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

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

Cotton, Christopher, Frank McIntyre, and Joseph Price. 2013. "Gender Differences in Repeated Competition: Evidence from School Math Contests." *Journal of Economic Behavior and Organization*, 86: 52–66.

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

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

Dohmen, Thomas, and Armin Falk. 2011. "Performance pay and multidimensional sorting: Productivity, preferences, and gender." *American Economic Review*, 101: 556–590.

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

the underperformance of the highest-ability females combined with the overperformance of lower ability males, relative to their performance in later contests. Allowing heterogeneity in ability within gender groups would enable an analysis to consider such performance patterns in more details, but would lead to a less straightforward analysis and less intuitive results.

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

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

Günther, Christina, Neslihan Arslan Ekinci, Christiane Schwieren, and Martin Strobel. 2010. "Women can't jump? An experiment on competitive attitudes and stereotype threat." *Journal of Economic Behavior and Organization*, 75(3): 395–401.

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

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." *AREC: Annual Review of Economics*, 3.

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

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

Ong, David, and Zhouqiong Chen. 2012. "Tiger women: An all-pay auction experiment on gender signaling of desire to win." Peking University, Working Paper.

Shurchkov, Olga. 2012. "Under pressure: Gender differences in output quality and quantity under competition and time constraints." *Journal of the European Economic Association*, 10: 1189–1213.

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

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

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

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

5. Appendix

5.1. **Derivation of equilibrium effort in the initial model.** Here we walk through the analysis to determine the equilibrium effort functions presented in Section 3.1.

For each player i:

$$Eu_i = \frac{a_i e_i}{\sum_{j=1}^{N} a_j e_j} - \tau_i e_i$$

In equilibrium, each player chooses e_i to maximize Eu_i, while correctly anticipating the e_j of the other players. The first order conditions of player i's optimization problem are:

$$\frac{\partial Eu_i}{\partial e_i} = \frac{a_i(\sum_{j=1}^{N} a_j e_j - a_i e_i)}{(\sum_{j=1}^{N} a_j e_j)^2} - \tau_i = 0$$

Which we can simplify:

$$a_i e_i = \sum_{j=1}^{N} a_j e_j - \frac{\tau_i}{a_i}(\sum_{j=1}^{N} a_j e_j)^2 \tag{5}$$

Summing $a_i e_i$ over all players gives

$$\sum_{j=1}^{N} a_i e_i = N \sum_{j=1}^{N} a_j e_j - \sum_{j=1}^{N} \frac{\tau_j}{a_j}(\sum_{j=1}^{N} a_j e_j)^2$$

Solving for $\sum_{j=1}^{N} a_j e_j$, we get

$$\sum_{j=1}^{N} a_j e_j = \frac{N-1}{\sum_{j=1}^{N} \frac{\tau_j}{a_j}} \tag{6}$$

Substituting (6) into (5) and then dividing both sides by a_i gives the equilibrium condition from the body of the paper

$$e_i = \frac{1}{a_i} \frac{N-1}{\sum_{j=1}^{N} \frac{\tau_j}{a_j}} - \frac{\tau_i}{a_i^2} \left(\frac{N-1}{\sum_{j=1}^{N} \frac{\tau_j}{a_j}} \right)^2 \tag{7}$$

In single gender contests, all N players have the same τ and α, in which case $\sum_{j=1}^{N} \frac{\tau_j}{a_j} = N\tau/a$ and (7) simplifies to

$$e_i = \frac{1}{\tau} \frac{N-1}{N^2}$$

Plugging in $N = 2$ and $N = 6$ gives the expressions for e^*_{MvM} and e^*_{FvF} in the paper.

In a mixed gender contest (with an even number of players), there are an equal number of males and females. Therefore, $\sum_{j=1}^{N} \frac{\tau_j}{a_j} = (N/2)\tau_M/a_M + (N/2)\tau_F/a_F = (N/2)(\tau_M a_F + \tau_F a_M)/(a_M a_F)$ and (7) simplifies to

$$e_i = \frac{1}{a_i} \frac{2(N-1)(a_M a_F)}{N(\tau_M a_F + \tau_F a_M)} - \frac{\tau_i}{a_i^2} \left(\frac{2(N-1)(a_M a_F)}{N(\tau_M a_F + \tau_F a_M)} \right)^2.$$

where $i \in \{M, F\}$ depending on whether player i is male or female. For a male in the contest,

$$e_{MvF} = \frac{2a_F(N-1)(Na_M t_F - (N-2)a_F t_M)}{N^2(t_M a_F + t_F a_M)^2}$$

with an analogous result females. Plugging in $N = 2$ and $N = 6$ gives the expressions for e^*_{MvF} and e^*_{FvM} in the paper.

5.2. **Differences in Risk Aversion.** When player i puts forth effort e_i and his opponents put forth total effort EO, player i earns expected utility

$$Eu_i = -\frac{ae_i}{ae_i + aEO}\exp[-\rho_i(Y + 1 - taue_i)] - \frac{aEO}{ae_i + aEO}\exp[-\rho_i(Y - \tau e_i)].$$

First order conditions with respect to e_i simplify to

$$EO + \tau e_i^2 \rho_i + \tau e_i EO \rho_i + \exp[\rho_i] EO((e_i + EO)\tau\rho_i - 1) = 0. \tag{8}$$

Checking second order conditions and verifying that $e_i > 0$ provides higher Eu_i than $e_i = 0$ assures that we are solving for a global maximum.

In the rest of the section, we walk through the derivation of Result 7 for the case where $N = 2$. Proving the result for the case where $N = 6$ follows the same method, although the expression are messier given the additional players. A Mathematica file working through both the $N = 2$ and $N = 6$ cases is available from the authors upon request.

To find the symmetric equilibrium for the case when the two competitors are the same type, we solve the first order conditions given $e_i = e_{-i}(= EO)$, and ρ_i. This gives equilibrium effort in MvM and FvF contests.

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

Equilibrium e^* from a single sex contest 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, $e^*_{MvM} > e^*_{FvF}$.

We are interested in how e^*_{MvF} and e^*_{FvM} compare to e^*_{MvM} and e^*_{FvF} when there are only marginal differences between ρ_M and ρ_F. To do this, we determine $\partial e^*_{FvM}/\partial \rho_F$ and $\partial e^*_{MvF}/\partial \rho_F$ evaluated at $\rho_F = \rho_M$. Equation (8) my be rewritten for $i \in \{M, F\}$ in a mixed gender contest:

$$e^*_{MvF} + \tau e^{*2}_{FvM}\rho_F + \tau e^*_{FvM} e^*_{MvF}\rho_F + \exp[\rho_F] e^*_{MvF}((e^*_{FvM} + e^*_{MvF})\tau\rho_F - 1) = 0, \text{ and} \tag{9}$$

$$e^*_{FvM} + \tau e^{*2}_{MvF}\rho_M + \tau e^*_{MvF} e^*_{FvM}\rho_M + \exp[\rho_M] e^*_{FvM}((e^*_{MvF} + e^*_{FvM})\tau\rho_M - 1) = 0. \tag{10}$$

In equilibrium, the expressions for e^*_{FvM} and e^*_{MvF} depend on ρ_F and ρ_M. Taking the derivative of (9) and (10) with respect to ρ_F gives

$$\begin{aligned} &\tfrac{\partial e^*_{FvM}}{\partial \rho_F}(1 - exp[\rho_M]) + 2\tau\rho_M e^*_{MvF}\tfrac{\partial e^*_{MvF}}{\partial \rho_F} + 2\tau\rho_M exp[\rho_M] e^* FvM \tfrac{\partial e^*_{FvM}}{\partial \rho_F} + \\ &\tau\rho_M(1 + exp[\rho_M])(e^*_{FvM}\tfrac{\partial e^*_{MvF}}{\partial \rho_F} + e^*_{MvF}\tfrac{\partial e^*_{FvM}}{\partial \rho_F}) = 0 \end{aligned} \tag{11}$$

$$\begin{aligned} &\tfrac{\partial e^*_{MvF}}{\partial \rho_F} + \tau\rho_F e^*_{FvM}\tfrac{\partial e^*_{MvF}}{\partial \rho_F} + exp[\rho_F](\rho_F e^*_{MvF} + \rho_F e^*_{FvM} - 1)\tfrac{\partial e^*_{MvF}}{\partial \rho_F} + \\ &\rho_F e^*_{MvF}\tfrac{\partial e^*_{FvM}}{\partial \rho_F} + 2\tau\rho_F e^*_{FvM}\tfrac{\partial e^*_{FvM}}{\partial \rho_F} + \tau\rho_F exp[\rho_F] e^*_{MvF}(\tfrac{\partial e^*_{MvF}}{\partial \rho_F} + \tfrac{\partial e^*_{FvM}}{\partial \rho_F}) = 0 \end{aligned} \tag{12}$$

We are interested in the signs on $\partial e^*_{MvF}/\partial\rho_F$ and $\partial e^*_{FvM}/\partial\rho_F$ at the point where $\rho_M = \rho_F$. We substitute into (11) and (12) the values $\rho_F = \rho_M = \rho$ and

$$e^*_{FvM} = e^*_{MvF} = e^*_{MvM} = e^*_{FvF} = \frac{\exp[\rho]-1}{2\tau\rho(1+\exp[\rho])},$$

and we solve the two resulting equation for $\frac{\partial e^*_{MvF}}{\partial\rho_F}$ and $\frac{\partial e^*_{MvF}}{\partial\rho_F}$. Simplifying gives

$$\left.\frac{\partial e^*_{FvM}}{\partial\rho_F}\right|_{\rho_F,\rho_M\to\rho} = \frac{(3+exp[\rho])(1+2\rho exp[\rho]-exp[2\rho])}{8\rho^2(1+exp[\rho])^2}, \tag{13}$$

$$\left.\frac{\partial e^*_{MvF}}{\partial\rho_F}\right|_{\rho_F,\rho_M\to\rho} = \frac{(1-exp[\rho])(1+2\rho exp[\rho]-exp[2\rho])}{8\rho^2(1+exp[\rho])^2}. \tag{14}$$

One can show that $(1+2\rho exp[\rho]-exp[2\rho]) < 0$ for all $\rho > 0$. To do this, first notice that the expression is strictly decreasing in ρv. Plugging in to the expression $\rho v = 0$ gives the maximum value it approaches as $\rho v \to 0$. This value is 0; thus for all $\rho v > 0$, the expression is negative. Furthermore, note that $(1-exp[\rho]) < 0$ for all $\rho > 0$, and that all other components of (13) and (14) are positive. This means that

$$\left.\frac{\partial e^*_{FvM}}{\partial\rho_F}\right|_{\rho_F,\rho_M\to\rho} < 0 \quad\text{and}\quad \left.\frac{\partial e^*_{MvF}}{\partial\rho_F}\right|_{\rho_F,\rho_M\to\rho} > 0.$$

If females are marginally more risk averse than males, then $e^*_{FvM} < e^*_{MvM} < e^*_{MvF}$. It remains to be shown that $e^*_{FvM} < e^*_{FvF}$. To do this, one can show that

$$\left.\frac{\partial e^*_{FvM}}{\partial\rho_F}\right|_{\rho_F,\rho_M\to\rho} < \left.\frac{\partial e^*_{FvF}}{\partial\rho_F}\right|_{\rho_F,\rho_M\to\rho},$$

which implies that a marginal change in ρ_F decreases e^*_{FvM} by more than it decreases e^*_{FvF}. Therefore,

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

The above analysis establishes that the performance ordering holds for small enough differences between ρ_F and ρ_M for the case where $N = 2$. The analysis of the case where $N = 6$ follows the same methodology, and is available in a Mathematica file from the authors.

Below, we present numerical examples showing that at least for the case where $N = 2$, the performance ordering holds more generally than for arbitrarily small differences in risk aversion. For these examples, we assume that $v = c = 1$ and we calculate equilibrium effort for both players in a contest for different ρ_i and ρ_{-i} combinations. Table 4 reports the equilibrium effort of player i for 25 (ρ_i, ρ_{-i}) pairs.

TABLE 4. Equilibrium Performance of Player i in Risk Aversion Game

	$\rho_{-i} = 0.25$	$\rho_{-i} = 0.5$	$\rho_{-i} = 1$	$\rho_{-i} = 2$	$\rho_{-i} = 4$
$\rho_i = 0.25$	0.2487	0.2490	0.2496	0.2432	0.1882
$\rho_i = 0.5$	0.2443	0.2449	0.2471	0.2457	0.1918
$\rho_i = 1$	0.2250	0.2259	0.2311	0.2440	0.1984
$\rho_i = 2$	0.1549	0.1530	0.1536	0.1904	0.2080
$\rho_i = 4$	0.0559	0.0524	0.0460	0.0367	0.1205

Reports the numerical performance measure for player i in the risk aversion game. The calculations assume $v = c = a = 1$ for both players.

Consider the following examples, with performance measures taken from Table 4.

- If $\rho_M = 0.25$ and $\rho_F = 1$, then
$$p^*_{FvM} = 0.2250 < p^*_{FvF} = 0.2310 < p^*_{MvM} = 0.2487 < p^*_{MvF} = 0.2496.$$
- If $\rho_M = 0.5$ and $\rho_F = 2$, then
$$p^*_{FvM} = 0.1530 < p^*_{FvF} = 0.1904 < p^*_{MvM} = 0.2449 < p^*_{MvF} = 0.2457.$$
- If $\rho_M = 1$ and $\rho_F = 2$, then
$$p^*_{FvM} = 0.1536 < p^*_{FvF} = 0.1904 < p^*_{MvM} = 0.2311 < p^*_{MvF} = 0.2449.$$

Each of these examples satisfies C2, showing how even substantial differences in risk aversion between males and females produce predictions that are consistent with the significant empirical evidence.

5.3. **Considering additional explanations.** Consider the possibility that only males underestimate female ability. In this case, $e^*_{FvM} = e^*_{MvM} = e^*_{FvF} = (N-1)/(\tau N^2)$, and e^*_{MvF} is the same as in the case where both males and females underestimate female ability. Similarly, when only females underestimate their ability, $e^*_{MvF} = e^*_{MvM} = e^*_{FvF} = (N-1)/(\tau N^2)$ and e^*_{MvF} is the same as in the case where both males and females 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 8. *If only males underestimate female ability or overestimate male ability, then*
$$p^*_{MvF} < p^*_{FvM} = p^*_{FvF} = p^*_{MvM} \ \ when N = 2,$$
$$p^*_{FvM} = p^*_{FvF} = p^*_{MvM} < p^*_{MvF} \ \ when N = 6.$$
If only females underestimate female ability or overestimate male ability, then in when either $N = 2$ or $N = 6$,
$$p^*_{FvM} < p^*_{FvF} = p^*_{MvF} = p^*_{MvM}.$$

Additionally, we can consider differences in preferences, while allowing them to depend 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. In this section, we consider an alternative model of that allows for such possibilities. To rule out these explanations, it is sufficient to show that the model fails to predict the significant evidence in the

CMP data. So we focus on the two player competitions in this section. As before, allow males and females to differ only in their preference parameter τ. Thus,

$$e^*_{FvF} = \frac{1}{4\tau_{FvF}} \qquad e^*_{MvM} = \frac{1}{4\tau_{MvM}}$$

$$e^*_{FvM} = \frac{\tau_{MvF}}{(\tau_{MvF}+\tau_{FvM})^2} \quad e^*_{MvF} = \frac{\tau_{FvM}}{(\tau_{MvF}+\tau_{FvM})^2}.$$

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

Result 9. *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

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

These models fail to predict the significant evidence in the CMP data, and therefore fails to satisfy C1 and C2. We can rule out these explanations of the empirical evidence.

Next consider 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 $\tau_L = \tau_{FvM} = \tau_{MvM}$, and

$$e^*_{FvF} = \frac{1}{4\tau_H} \qquad e^*_{MvM} = \frac{1}{4\tau_L}$$

$$e^*_{FvM} = \frac{\tau_H}{(\tau_L+\tau_H)^2} \quad e^*_{MvF} = \frac{\tau_L}{(\tau_L+\tau_H)^2}.$$

From these values we can derive the following results.

Result 10. *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 less enjoyment from competition against females (i.e., $\tau_{MvM} < \tau_{FvF} = \tau_{FvM} < \tau_{MvF}$), then

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

Each of these possibilities fails to predict the significant evidence in the CMP data, and therefore fails to satisfy C1 and C2. They therefore provide a poor explanation for the empirical evidence.