Salient Gender Identity and Power Imbalance in a Group Contest

Usame Berk Aktas

Florian Heine

June 2024

Abstract

The metaphor "glass ceiling" describes female under-representation in executive roles, often attributed to women's aversion to competition. This study investigates gender differences in competitive behaviour through a controlled experiment, examining how the salience of gender identity affects investment in contests between male and female groups. Our results show that power dynamics influence competitiveness, with advantaged groups investing more. Interestingly, gender identity salience affects males and females differently: it narrows the gender gap in female-dominated contests but widens it in male-dominated ones. These effects are not attributed to social identity or in-group cohesion, as our measures of the social identity parameter and Social Value Orientation (SVO) remain consistent across genders. Our study provides insights into gender-based competitive behaviour. These results have significant implications for workplace dynamics and organisational policies. In female-dominated fields, highlighting gender identity could reduce gender disparities in promotion and competition. Conversely, in male-dominated environments, it might exacerbate these disparities.

Keywords— Group Contest; Social Identity; Gender Identity; Asymmetric Groups; Group Size Paradox

1 Introduction

Applications for group contests range from conflict related to language, religion or culture, to political competition and collective rent-seeking.[1] In these situations, a collection of individuals compete for a prize via irreversible and costly investments. In the underlying applications and the associated model, each member of the winning group enjoys a share of the spoils, often irrespective of or only weakly related to the individual effort. Within a group, this creates a trade-off: On the one hand, group members have the incentive to expend effort to win the prize, and on the other hand, each member has an incentive to keep resources to oneself. Importantly, investment has no productive value in group contests, but only influences the odds for winning the prize. Therefore, globally, higher efforts imply more inefficiency. For example, consider money spent on behalf of trade syndicates to lobby for government subsidies or to impose regulations on competitors to increase market share. The lobbying expenses do not add value to the economy in any way, directly or indirectly, except for the successful syndicate.

Group identity has been theorised to be one of the major components in initiating and escalating conflict (Sen, 2007). The associated social identity theory describes salient group identity to cause a blurring of the boundaries between personal and group welfare, leading to behaviour that bolsters group benefit at the expense of personal individualistic self interest (Tajfel, Turner, Austin, & Worchel, 1979). In intergroup relations, individuals place themselves and others in different categories based on perceived similarities and differences, and according to these categorisations identify others as either in-group or out-group members (Akerlof & Kranton, 2005; Basu, 2005; Akerlof & Kranton, 2010).[2] As a result of this identification, people discriminate in favour of the in-group and against the out-group. This has been a key concept to understand phenomena such as racial and political conflicts, as well as discrimination and many more (Abdelal, Herrera, Johnston, & McDermott, 2006). We study the effect of the salience of social identity on group conflict and how this interacts with competing groups' relative power.

Akerlof and Kranton (2000) have been among the first to introduce the concept of social identity into economic frameworks. By adopting a utility maximisation framework that incorporates an individual's self-identification, social identity helps understand the "microfoundation for earlier (discrimination) models."[3] To date there are only a few studies which incorporate social identity theory into contest games. Most prominently, Chowdhury, Jeon, and Ramalingam (2016) engage two homogenous groups – East Asians and Caucasians – in a group contest either with or without revealing the groups' racial composition. They find that revealing racial composition increases effort and decreases free riding significantly. This means that a salient real identity can escalate conflict, which is also predicted by social identity theory (Sen, 2007).[4]

[1] Dechenaux, Kovenock, and Sheremeta (2015) provide an extensive literature review, including applications and field studies on group contests.

[2] Such identification happens very early in life for some categories (Powlishta, Serbin, Doyle, & White, 1994).

[3] See Costa-Font and Cowell (2015); Li (2020) for a thorough review on social identity in economics.

[4] Similarly, Chakravarty, Fonseca, Ghosh, and Marjit (2016) report a small significant impact of social identity among Hindu villagers, yet none among Muslim villagers, in two-player group contest games. Sutter and Strassmair (2009) and Cason, Sheremeta, and Zhang (2012) use communication to trigger social identity in a group contest game. Both studies find that within, or between-group communication creates better coordination within and between the groups and less free riding within groups. An important point is

In our study we employ gender identity as social identity categorisation for three concrete reasons. *1)* Gender identity has been recognised as an important concept through which we define ourselves in daily life (Sen, 2007). *2)* Categorisation in terms of gender avoids identification problems. Observations of membership by gender are usually made without any error (Akerlof & Kranton, 2002). *3)* In many situations, males are found to be more aggressive and competitive (Niederle, Segal, & Vesterlund, 2013; Croson & Gneezy, 2009; Gneezy, Niederle, & Rustichini, 2003), particularly when the conflict is physical and can sustain physical harm (Hay et al., 2011). Prior results from the contest game literature, by contrast, suggest that female participants invest more resources into the contest, i.e. compete more aggressively (Price & Sheremeta, 2015; Chowdhury et al., 2016; Heine & Sefton, 2018).[5] We investigate if this phenomenon of female competitiveness in this context of between group competition is triggered via group identity (as results from Cadsby, Servátka, & Song, 2013, suggest). This issue is of significant economic importance as group contests (e.g., for promotion decisions) are ubiquitous within firms, especially among top management.

Many related applications in the field, such as competition for promotion or tenure, are characterised by a (power) imbalance between social identity groups. Particularly considering gender identity, in many workplace settings, males are overrepresented both in numbers and in power (Shor, Van De Rijt, Miltsov, Kulkarni, & Skiena, 2015; Lang, 2010; Cotter, Hermsen, & Vanneman, 2000). In other areas like the service sector or social programme jobs, by contrast, women tend to be strongly overrepresented (Barone & Assirelli, 2020; Hurst, Gibbon, & Nurse, 2016). We study how the salience of social identity, in particular gender identity, influences the degree of engagement into a competition between groups and how being in an advantaged or disadvantaged position interacts with this.

We start by devising a theoretical model on social identity in a group contest in which a participant maximises individual utility as a weighted sum of own and others' utility (similar to Y. Chen & Li, 2009; Zaunbrecher & Riedl, 2016; Kolmar & Wagener, 2019). We hypothesise that when there is more weight placed on social identity, this leads to an increase in investment into the contest. We then conduct an experiment in which we vary the salience of social identity via the instructions, the positioning of a gender identity survey and the representation of participants by gendered or neutral emojis. Our results indicate that advantaged groups invest significantly more into the contest and that female groups invest less into the contest when gender identities are salient.

The remainder of this article is structured as follows. First, in Section 2 we explain the experimental design. Then, in Section 3, we discuss the theoretical framework by means of a social preferences model, from which we derive our hypotheses. We then present the results of our experiment in Section 4, before concluding in Section 5.

that this improved coordination happens even when it reduces, rather than enhances, efficiency i.e. higher over-spending in the within-group communication treatment. Drawing on social identity theory, Cason et al. (2012) interpret that "intra-group communication increases subjects' identification with their group and shifts their self-categorization from the individual to the group level, leading them to coordinate better with their group and compete more with the opponent group."

[5] For instance, in Chowdhury et al. (2016), the authors find that "the increase in conflict in a laboratory contest setting does not arise due to the behavior of a particular race, but due to the increase in efforts by females across racial groups." With the race treatment average efforts of the females increases from 11.718 to 18.265, compared to a slight increase from 11.523 to 12.407 in males.

2 Experimental Design

We model group competition using a repeated Tullock contest (Katz, Nitzan, & Rosenberg, 1990; Tullock, 1980) between two groups in partner matching. Figure 1 illustrates a round of the game for the Symmetric Control treatment, other treatment variations will be explained hereafter. Each player receives an individual endowment of $T_i = 60$ points per round, which they invest in the contest game to buy lottery tickets for their team for a price of one lottery ticket for one point. Endowment that is not invested will be added to the player's private account for that round. Players cannot accumulate funds for future rounds. The winning probability for a group is the sum of lottery tickets bought by the own group divided by the total amount of lottery tickets bought by both groups. If none of the players in both groups invest, one of the groups will win with equal chances. Once all investment decisions have been made, one ticket is drawn and the prize $z_i = 40$ goes to all players from the group with the winning ticket, the losing group receives nothing. Note that all players from the winning group receive the full prize, irrespective of their personal investment into the contest. The expected earnings of individual g from group A can be written as (equivalent for group B):

$$\pi_g\left(\sum_{i\in A} a_i, \sum_{j\in B} b_j\right) = T_i + \frac{\sum_{i\in A} a_i}{\sum_{i\in A} a_i + \sum_{j\in B} b_j} \cdot z_i - a_g \tag{1}$$

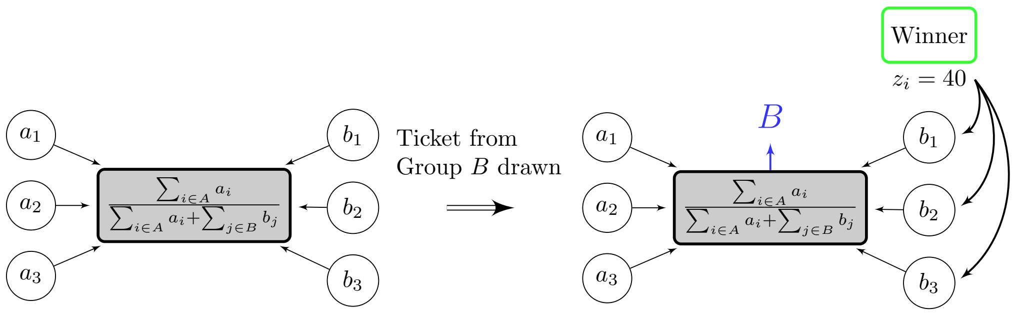


Figure 1: Groups of three ($a_{1,2,3}$ versus $b_{1,2,3}$) compete for a prize in a group contest. Suppose a winning ticket for Group B is drawn, then each player from that group receives the prize $z_i = 40$.

We explore the effect of social identity in a 3×2 experimental design. In all six treatments, participants were recruited such that each group consists of individuals from the same real social identity (as in Chowdhury et al., 2016). This means, each competing pair constitutes of one homogeneous group of participants with *female* gender identity competing against a group of participants with *male* gender identity.

Figure 2 provides an overview of our treatments. Between the treatments on the rows we vary the makeup of the groups. In the Symmetric treatments, groups of the same size – i.e. three female versus three male players – compete for the prize (as in Figure 1). In the Asymmetric Female treatments, five female participants compete against a group of three males. Equivalently, in the Asymmetric Male treatments, five male participants compete against a group of three females. As such, we vary whether groups are of equal size and power, or whether a certain social identity group is in an advantageous position and over-represented. This reflects phenomena in the field, in which certain (mostly male; sometimes female, for example in education) groups have more power or influence.

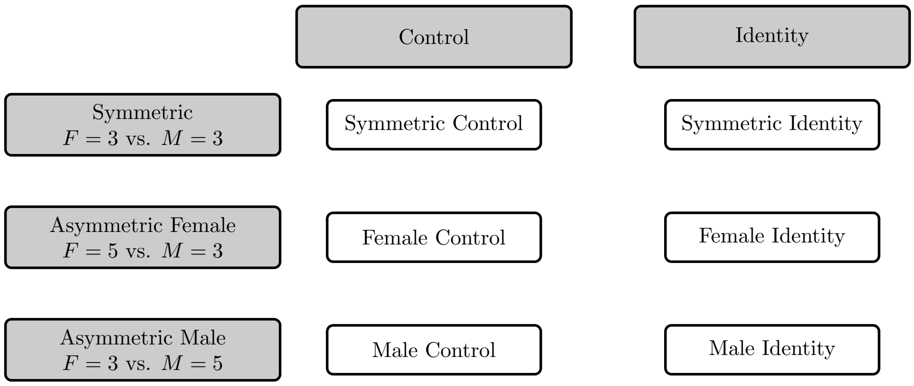


Figure 2: Overview of the Experimental Design. We apply a 3×2 design varying the salience of social identity on the one hand, and group size on the other hand.

Treatments in the two columns of Figure 2 differ in the salience of social identity. We vary salience in three ways, as visually summarised in Figure 3, where text in red font highlights the differences between the Control and Identity treatments. *First*, in the instructions for the Identity treatments we make explicit the social identity of group members, i.e. that this is a game of male versus female participants. *Second*, we vary the position of the Gender Identity Survey (Cameron, 2004) such that it comes either at the beginning of the experiment to prime gender identity, or at the end of the experiment.[6] The Gender Identity Survey serves another purpose: We are also interested in measuring the degree of identification with the own social identity, i.e. how much a participant's own identity overlaps with their gender identity, and if this have an effect on the investments. *Third*, we employ emojis that either reflect the gender identity group, or neutral ones, using emojis developed by OpenMoji (2020) (see Figure 4).[7]

As illustrated in Figure 3, the contest was repeated for 10 rounds, which allows us to observe if patterns are driven by behaviour from particular rounds only. Earnings from prior rounds cannot be saved for use in future rounds, instead participants receive a fixed endowment of 60 points in each round. At the end, one round will be selected for payment at random. This way, participants should treat each period as the payoff-relevant round, reducing potential hedging behaviour (Charness, Gneezy, & Halladay, 2016).

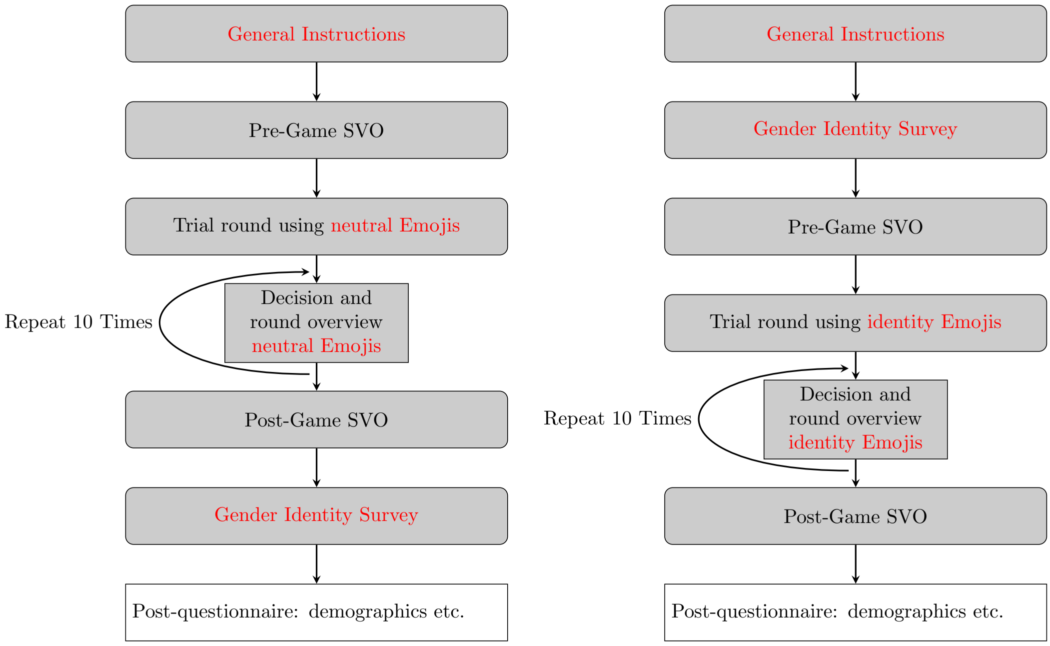
We programmed this computerised experiment in z-Tree by Fischbacher (2007) and conducted the sessions at CentERlab of Tilburg University, Netherlands, between November 2021 and September 2022.[8] A total of 350 individuals (average age 21.2 years, sd=3.47) participated in the study.[9] Participants sat in a cubicle, visually separated from each other. The experiment took about 60 minutes, including instructions and payment, average earnings were about €12.16. The following outlines the structure of the experiment in

[6] Cameron (2004) develops a three factor model representing social identity on three factors: centrality, ingroup affect, and ingroup ties. It is a twelve-item, partially reverse-scored Likert-type questionnaire including statements such as: *"I have a lot in common with other* men/women.*"* Please find a screen shot from this questionnaire in Figure 17.

[7] OpenMoji graphics are licensed under the Creative Commons Share Alike License 4.0 (CC BY-SA 4.0).

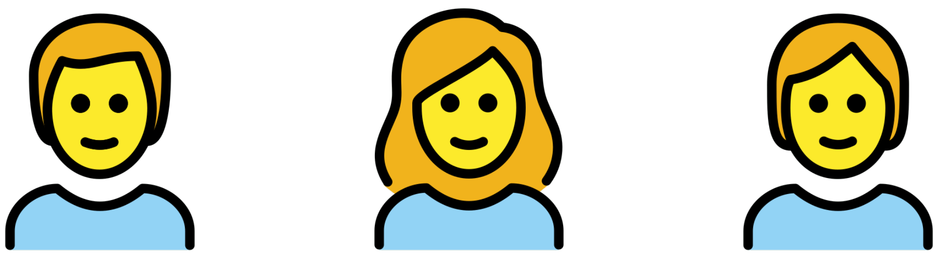
[8] Please find screen shots in Appendix A.

[9] Please find a detailed power analysis in Appendix C.



(a) Control Treatments **(b)** Identity Treatments

Figure 3: Experimental Setup. Control Treatments differ from Identity Treatments in elements highlighted with red font.



(a) Male Emoji **(b)** Female Emoji **(c)** Gender-neutral Emoji

Figure 4: Emojis used to prime gender social identity groups.

more detail (see also Figure 3). We adopted a recruitment strategy similar to the one in Chowdhury et al. (2016). Tilburg University's participant-database Sona allows us to use information on registered participants' gender identity for experiment invitations. As such, we used this information to recruit participants from each gender identity groups and applied the matching for each treatment.

2.1 The Structure of an Experimental Session

To make sure the gender identity matching was successful, participants started with a screen on demographics (gender identity, age, etc. See Figure 13). This allowed us to double check that the self-specified gender identity of participants when starting the experiment matches with what they have registered in the database. Next, participants received the instructions for the experiment, followed by a few questions about hypothetical game situations on screen to make sure participants completely understand the game. For the Identity treatments, the specific social identity of the group was made salient in the instructions and on the game screen. The emojis as in Figure 4 represented the players in the groups. While a neutral emoji, as in Sub-figure 4c represented participants in both Control treatments, players of the female and male gender identity groups were be represented by emojis as in Sub-figures 4b and 4a, respectively.

Before the contest, participants performed two social value orientation (SVO) tests, one with respect to someone from the *own* group and one with respect to someone from the *other* group. Each of the two tests consists of six different resource allocation decisions, based on Murphy, Ackermann, and Handgraaf (2011)'s slider measure.[10] The difference between giving to someone from the *own* group versus someone from the *other* group provides us with a measure for individual in-group bias. Figure 14 provides an example screen from this stage.

Then, participants played the group contest game for 10 rounds. In each round they decided how much to invest to buy lottery tickets for the contest, as illustrated in Figure 15. After each round, each player was informed about which group has won, how much their group has invested in total, how much the other group has invested in total and what the probability of winning the contest was for their group. To ensure no in-group reputation effect, we only provided information about aggregate investments at the group level instead of player level (see Figure 16).

Subsequent to the group contest, we conducted a second round of social value orientation test, again both with respect to the *own* group and the *other* group. We conducted the SVO test again to measure if the contest itself has had an effect on the level of in-group bias. In total, counting both the pre-game SVO and the post-game SVO, a player will be involved in four rounds of SVO tests as receiver and four rounds of SVO test as the dictator (each round consisting of six distinct allocation decisions, as in Figure 14). Two out of four rounds are with someone from their *own* group and the other two rounds are with someone from the *other* group. At the end of the experiment, one random resource allocation is selected from the SVO tests that the player was involved in. For this, our matching protocol avoids tacit gift exchange and ring matching situations. Furthermore, one random round from the group contest game is selected for payment. Players received the earnings from both the selected SVO payoff and the payoff from the group contest.

As last step, participants filled in a questionnaire about demographics, strategies used in the game, what the player thinks a social player should implement in this game and risk preferences.

3 Theory and Hypotheses

We begin by discussing the equilibrium predictions under standard (individualistic) preferences before presenting our social preferences model from which we derive our hypotheses. Under standard preferences, expected earnings for the group contest is the individual en-

[10] Murphy et al. (2011)'s slider measure is a parametric implementation of the SVO ring measure conceptualised by Griesinger and Livingston Jr (1973); Liebrand (1984).

dowment T_i less investment into the contest v_a plus the winning-probability weighted individual price z_i as in Equation 1. It is well-documented in the literature that for a group A, there exists a multiplicity of equilibria characterised by $\sum_{a \in A} = \frac{z_a}{4}$ (e.g., Konrad, 2009). In Appendix B we present a formal derivation of the equilibrium contribution. The calibrations of our experiment allow for a clean comparison between treatments, as group-level equilibrium predictions are equal across all treatments ($\sum_{a \in A} = 10$), as illustrated in Figure 5.

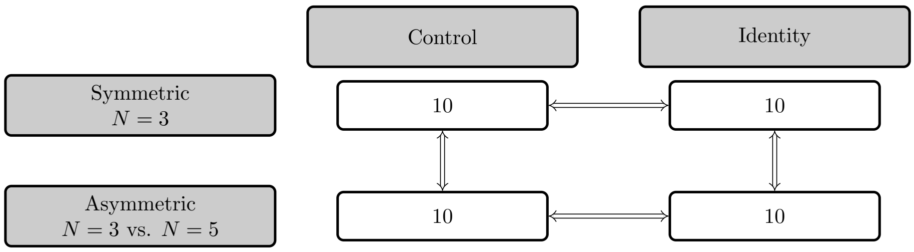


Figure 5: Group-Level Equilibrium Prediction under Individualistic Preferences.

3.1 Social Preferences Model

We use a social preferences model similar to Y. Chen and Li (2009); Kolmar and Wagener (2019), where an agent maximises a weighted sum of her own and others' payoffs. In particular, our model is strongly motivated by Zaunbrecher and Riedl (2016)'s model for the role of ingroup bias / social identity in group contests.[11] Two groups A and B compete for a fixed individual prize $z_i > 0$. As such, the prize will be divided equally among all members of the winning group, irrespective of contest investment or other factors.

Agents maximise a weighted sum of own and other group members' payoffs. Let π_g be player g's payoff, $\overline{\pi}_{A \setminus g}$ the average payoff of player g's other group members and α the social-identity parameter, i.e. the weight that g puts on her group mates' payoff. As such, parameter α reflects the strength of g's ingroup bias, where a higher α implies a stronger ingroup bias.

$$u_g(i) = (1 - \alpha) \cdot \pi_g + \alpha \cdot \overline{\pi}_{A \setminus g} \tag{2}$$

The *material* payoff for each player is as in Equation (1), analogously for both groups. Players with individualistic preferences only care about their own material payoff, so $\alpha = 0$. Incorporating social preferences into the decision problem, a player g from group

[11] Both Zaunbrecher and Riedl (2016) and Kolmar and Wagener (2019) discuss a social identity model in group contest games. Both approaches differ in some crucial aspects. While Zaunbrecher and Riedl (2016) employ a weight $\alpha \in [0.1, 1]$, in Kolmar and Wagener (2019), this is only a binary measure of whether members of a group identify with the own group. Another difference between the two seminal approaches is that while Zaunbrecher and Riedl (2016) only consider social preferences towards members of the own group, Kolmar and Wagener (2019) model a preference for parochial disutility for the competing group. Third, Zaunbrecher and Riedl (2016) consider the *average* group payoff as reference, while Kolmar and Wagener (2019) refer to the *sum* of group payoffs. Our model combines elements from both approaches.

A maximises:

$$u_g\left(\sum_{i\in A}a_i,\sum_{j\in B}b_j\right)=(1-\alpha)\left[T_i+\frac{\sum\limits_{i\in A}a_i}{\sum\limits_{i\in A}a_i+\sum\limits_{j\in B}b_j}\cdot z_i-a_g\right]+ \frac{\alpha}{N_A-1}\left[(N_A-1)\left(T_i+\frac{\sum\limits_{i\in A}a_i}{\sum\limits_{i\in A}a_i+\sum\limits_{j\in B}b_j}\cdot z_i\right)-\sum_{i\in A\backslash g}a_i\right] \tag{3}$$

In Appendix B we show that for group A the equilibrium for total investment equals to

$$\sum_{i\in A}a_i=\frac{z_i\,(1-\beta)}{(2-\alpha-\beta)^2}. \tag{4}$$

Similarly, the best response function for group B, where β is group B's social-identity parameter (group B's equivalent to what is α in group A), is

$$\sum_{j\in B}b_j=\frac{z_i\,(1-\alpha)}{(2-\alpha-\beta)^2} \tag{5}$$

In Appendix B.1 we show that group contribution depends positively on the social-identity parameter within the calibrations of the experiment, i.e. $\frac{\partial\sum_{i\in A}a_i}{\partial\alpha}\geq 0$. Intuitively, putting more weight on groupmates' earnings (meaning, increasing α) raises the attractiveness of the prize at stake as players (partially) internalise the positive within-group externalities from their contest-spendings.[12] As such, groups with a higher level of ingroup bias would be willing to chip in more resources into the contest. We expect the salience of social identity to enhance the level of identification with the own group, i.e. increasing α. We hypothesise:

Hypothesis 1. *Total investment will be greater in the Social Identity Treatments.*

The second aspect we investigate is a power imbalance between the competing groups through the relative over-representation of one group. We model this by considering groups of unequal size.[13] Prior empirical results suggest that larger groups have a higher probability of winning against smaller groups (Sheremeta, 2018; Ahn, Isaac, & Salmon, 2011; Abbink, Brandts, Herrmann, & Orzen, 2010). Explaining this empirical finding with the social identity model with N_A and N_B representing the number of players in group A and B respectively, we note that

$$N_A>N_B\ \rightarrow\sum_{i\in A}a_i>\sum_{j\in B}b_j. \tag{6}$$

[12]While having positive externalities on other members from the *own* group, contest-spending implies negative externalities on members from the *other* group and a negative *total* effect on aggregate utility considering the set of all players involved.

[13]Modelling power imbalance through varying efficiency factors per group would be an alternative. We employ unequal group size as channel for a power imbalance between groups as this is more intuitive to explain to participants and because this makes the power imbalance more implicit. This tacit modelling of power imbalance mirrors phenomena in the field like "the unseen, yet unbreachable barrier that keeps (...) women from rising to the upper rugs of the corporate ladder" (US Federal Glass Ceiling Commission, 1995).

Plugging our results from Equations 4 and 5 delivers

$$\frac{z_i\,(1-\beta)}{(2-\alpha-\beta)^2} > \frac{z_i\,(1-\alpha)}{(2-\alpha-\beta)^2},$$

which simplifies into

$$\alpha > \beta.$$

This implies that the social-identity parameter is stronger in the large group, i.e. individuals in the larger groups value other member's utility more than individuals in the small group do. We postulate that the value of the social identity parameter depends on group size. As the social identity parameter increases with group size, our theory predicts total investment to increase with group size.

Hypothesis 2. *Total investment into the contest will be higher in the large group than in the small group.*

We expect the group contest to create an in-group bias measurable through the SVO test. Therefore:

Hypothesis 3. *The in-group bias measured by the SVO will be higher after the group contest (i.e. the post-game SVO measure) compared to before (i.e. the pre-game SVO measure).*

We expect that the high salience level of the shared social identity in the Identity treatments creates a higher in-group bias already at the start of the experiment, as compared to individuals in the Control treatments.

Hypothesis 4. *The in-group bias measured by the SVO before the group contest will be higher for the Identity treatments compared to the Control treatments.*

4 Results

First we present patterns in investments into the contest both in general and over time, and look into overall treatment differences. Then we examine the effect of salience of gender identity and group asymmetries, as well as underlying interaction effects. We will show that male and female participants react very differently to the power imbalance between groups and whether or not the gender identity of groups is salient. We then explore the role at which social preferences and social value orientation can explain our results, demonstrating that particularly in disadvantaged groups, salience of gender identity affects male and female participants' in-group bias very differently. We close by an analysis of the role of the degree to which a participant identifies with his or her stated gender identity using the Gender Identity Survey by Cameron (2004).

We apply non-parametric methods for hypotheses testing: Mann-Whitney U tests (MWU) (Mann & Whitney, 1947) for independent sample tests and Wilcoxon signed-rank test (Wilcoxon, 1945) for paired tests. Furthermore, we use the Kruskal-Wallis test (KW) (Kruskal & Wallis, 1952) and Dunn's test (Dunn, 1964) with a false discovery rate (FDR) adjustment by Benjamini and Hochberg (1995) for tests involving three or more groups. We use a non-parametric test for trend developed by Cuzick (1985). Unless specified differently, we use data on paired group level (six players in Symmetric treatments, eight players in Asymmetric treatments) as independent observation and apply two-sided tests.

For regressions, we use a GLS random effect panel regression and assess the distance between the resulting parameters using a Wald test. As the case may be, the dependent variable either is individual (with error terms clustered at group level) or group effort in period t, and the independent and control variables as described below.

4.1 Results Overview

A recurring result in empirical research on contests is a robust and sizeable degree of over-investment, meaning that players invest more than the Nash equilibrium, and in turn a lot more than the social optimum. In specific, recent experimental studies on group contests find over-investment ranging from 10% to 256% (Sheremeta, 2011, 2018). Other-regarding preferences, and in specific, parochial altruism and social identity theory, have been identified as one of the main mechanisms leading to such escalation of conflict (Sen, 2007).[14]

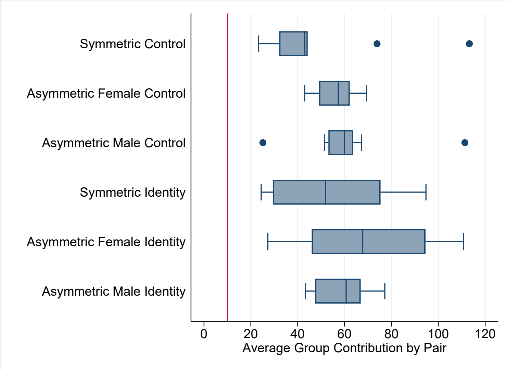
We too find a drastic and robust over-contribution with respect to both the risk-neutral individualistic equilibrium prediction (Wilcoxon test. H0: group contr. $= 10$, $N = 48$, $z = 6.031$, $p < 0.0001$) and hence also the social optimal strategy, which is even lower. Sub-Figure 6a provides an overview of group contest investment per group pair averaged over all rounds separate for each treatment.[15] The vertical red line indicates the equilibrium prediction under individualistic preferences. The figure provides a couple of first visual cues, which we will investigate more thoroughly in the upcoming subsections. *First*, asymmetric groups appear to invest more into the contest. *Second*, Investment levels do not seem to be higher in the Identity treatments. *Third*, making identity salient does however increase the noise in the data as illustrated by the increased box and whisker size in the figure.

Sub-Figure 6b illustrates investment levels separate for participants with male and female gender identity. While there again is no obvious difference in investment levels between the Control and Identity treatments, we can see some first suggestive evidence that female individuals invest more if they are in a larger group. We will zoom in on this in Subsection 4.2.

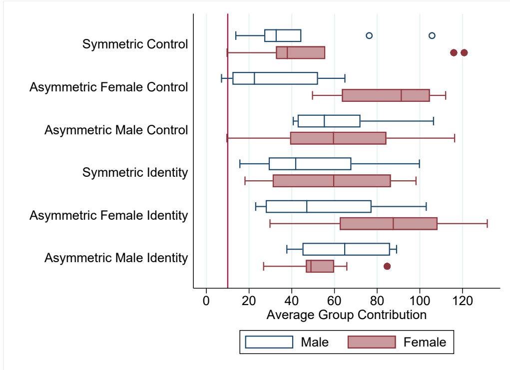
For all treatments, overall contest investment decreases over time (Cuzick Test at group level: all treatments pooled ($N = 96$) and separate ($16 \leq N \leq 18$), $z \leq -10.928$, $p < 0.0001$), as illustrated in Figure 7. The figure depicts average group investment per period for each treatment and the equilibrium prediction under individualistic preferences ($\sum_{a \in A} = 10$). The downward trend notwithstanding, contest investment exceeds the equilibrium prediction at all times. The figure also adds to the earlier observation that contest investment appears higher in the Asymmetric treatments, compared to the Symmetric treatments.

[14] Alternative explanations include non-monetary utility from winning (Delgado, Schotter, Ozbay, & Phelps, 2008; Sheremeta, 2010; Mago, Samak, & Sheremeta, 2016) and bounded rationality (Chowdhury, Sheremeta, & Turocy, 2014; Lim, Matros, & Turocy, 2014; Masiliunas, Mengel, & Reiss, 2014).

[15] Appendix D provides tables with more quantitative evidence, such as the mean and standard deviation, complementing the graphs provided in this Section.



(a) Group contest investment per group *pair* averaged over all rounds.



(b) Group contest investment per *gender identity* group averaged over all rounds.

Figure 6: Box Plots showing the medians, quartiles, octiles and outliers per category. The vertical red line indicates the equilibrium prediction under individualistic preferences.

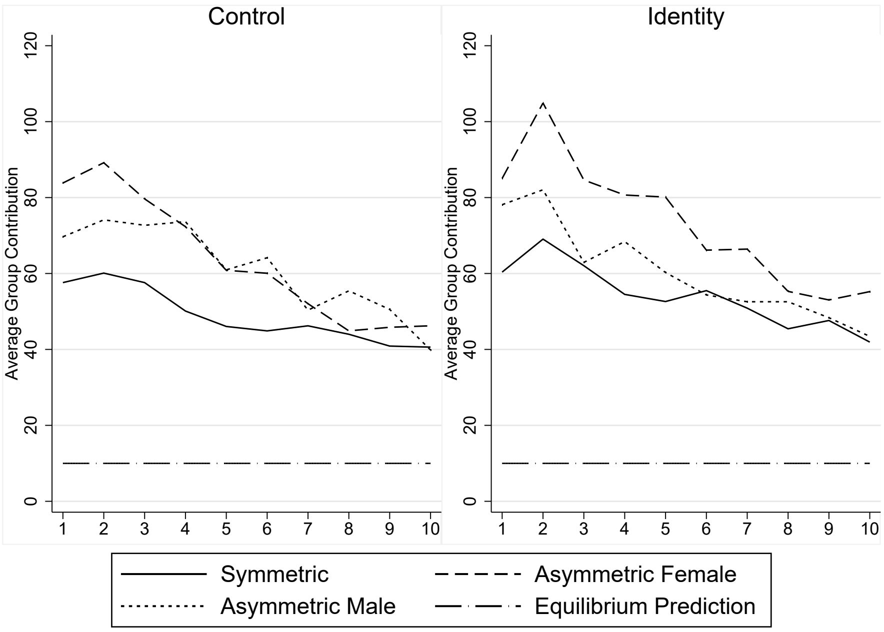


Figure 7: Average Group Contribution per Round.

4.2 Salience of Gender Identity

Using our social preferences model, we hypothesised that salience of social identity increases investment into the contest (Hypothesis 1). Analysing average contest investment aggregated at the Pair level, however, delivers no evidence that across the board, making gender identity salient would increase engagement into the contest (Wilcoxon test at Pair level. H0: group investment Control treatments = group investment Identity treatments, $N = 48$, $z = -0.804$, $p = 0.4277$). We further analyse this hypothesis using a random effects model with error terms clustered at pair-level to regress group investment in round t on treatment dummies and controls. In complement to the non-participant test, we adopt this regression approach to enhance the statistical power of our analysis. Unlike the non-parametric test, the regression method eliminates the need for clustering observations, contributing to a more robust statistical examination. The regression output in Table 1 shows that there is in fact no significant general treatment effect, echoing the observation from Sub-Figure 6a. The regression suggests that female participants invest more into the contest overall, which is in line with prior literature on group contest games (e.g., Price & Sheremeta, 2015; Chowdhury et al., 2016; Heine & Sefton, 2018).[16]

Table 1: Random effects model regressing group contest investment in round t on treatment dummies and controls.

	(1)	(2)
	Group Contribution in t	
Asymmetric Female Control	5.746	
	(9.15)	
Asymmetric Male Control	10.905	
	(11.46)	
Symmetric Identity	4.738	
	(11.74)	
Asymmetric Female Identity	17.508	
	(12.81)	
Asymmetric Male Identity	7.763	
	(9.48)	
Female	16.348***	21.832**
	(5.91)	(9.53)
Identity		10.086
		(7.90)
Female × Identity		−10.835
		(12.02)
Lagged Other Group Contribution	0.128**	0.133**
	(0.05)	(0.05)
Constant	52.754***	55.031***
	(9.12)	(5.92)
Number of observations	864	864
Number of panels	96	96
Within model R-squared	0.268	0.268
Between model R-squared	0.127	0.110
Overall R-squared	0.170	0.158

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Clustered standard errors in parentheses. Round fixed effects not reported.

Result 1 Making gender identity salient does not induce higher contest engagement overall.

We next zoom in on the observation from Sub-Figure 6b, which suggests a gender gap in contest investment between female and male groups. Figure 8 illustrates the difference

[16] Regression analyses at the individual level deliver qualitatively similar results.

between contest investment in female and male groups in a given round for each treatment. It shows that female groups invest more into the contest than their male counterparts, particularly when they are in an advantaged position, i.e. when they are in the larger group in the Asymmetric Female treatments (Wilcoxon test. H0: aggregate group investment male group = aggregate group investment female group, $N = 30,\ z = -3.215,\ p = 0.0013$). This effect appears as more pronounced if the gender identity is not salient, which suggests that female players are discouraged from competing when the male vs. female character of the game is salient. Over the course of the experiment, this gender gap decreases (Cuzick Test at group level: $N = 15,\ z \leq -7.391,\ p < 0.0001$).

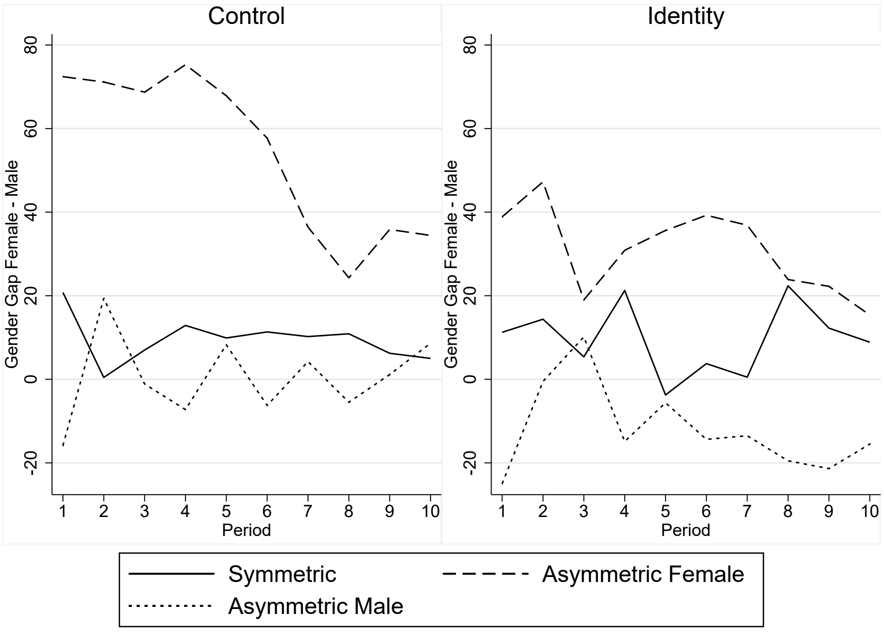


Figure 8: Gender gap in average group contribution levels between female and male groups per round.

For the Asymmetric Male treatments, by contrast, Figure 8 suggests a negative gender gap when identity is salient (Asymmetric Male Identity treatment), but not without salient gender identity (Asymmetric Male Control). This suggests that advantaged groups with male participants invest more than female groups do only if gender identity is salient. Non-parametric tests fail to confirm this relationship though (Wilcoxon test. H0: aggregate group investment male group = aggregate group investment female group, $N = 32,\ z = 0.603,\ p = 0.5641$). In contrast to the gender gap in the Asymmetric Female treatments, the gap in the Asymmetric Male treatments is persistent and does not decrease over time (Cuzick Test at group level: $N = 16,\ z \leq -0.495,\ p = 0.6207$). In the Symmetric treatments, there exists a small, yet not statistically significant difference in contest investment levels between female and male groups (Wilcoxon test. H0: aggregate group investment male group = aggregate group investment female group, $N = 34,\ z = -0.827,\ p = 0.4084$), which displays a small negative trend (Cuzick Test at group level: $N = 17,\ z \leq -2.025,\ p = 0.0428$) yet no difference between salient or non-salient treatments.

Result 2 Female participants display higher engagement into the contest when in an advantaged position, which is even more pronounced when gender identity *is not* salient. For male participants, by contrast, engagement into the contest is higher particularly when gender identity *is* salient.

4.3 Group Asymmetry

We further the analysis by looking at the effect of group competition between asymmetric groups more closely. The results so far indicate that large groups are more successful in mustering contest investments than small groups are, which is in line with evidence from the literature (Sheremeta, 2018). Our social preferences model from Subsection 3.1 suggests this may be because social identity is stronger in the advantaged group, which then causes the larger group to invest more into the contest. Our design allows to test this hypothesis. For the following analysis in this subsection, we only use data from the Asymmetric treatments, allowing for a direct comparison of behaviour when small and large groups interact. As such, we exclude data from the two Symmetric treatments for now.

Table 2 presents results of random effects regressions with error terms clustered at group-pair level using data from the Asymmetric treatments only. The regressions indicate that large groups invest about 25-34 points more and groups with female participants invest about 21-36 points more into the contest each round. In Regression (4) we interact three indicator variables for whether or not a group is large, whether it is a female group and whether gender identity is salient, which helps us identify an interesting gender-identity related dynamic. Concretely, the regression delivers weak evidence for a positive effect of salient social identity on group investment ($z = -1.87, p = 0.062$) for *male* participants, increasing their contribution by about 24 points when gender identity is salient. Female participants, by contrast, do not contribute more when gender identity is salient (Wald Test: $\beta_{\text{Identity}} + \beta_{\text{Identity}\times\text{Female}} = 0,\ \chi^2(1) = 0.6,\ p = 0.4391$). While male participants appear encouraged to compete when gender identity is salient, this effect is absent for female participants.[17]

Result 3 We find evidence that large groups invest more into the contest.

4.4 Social Preferences

In Subsection 3.1 we discussed our social preferences model which describes agents maximising a weighted sum of their own and others' payoffs. We show that our model predicts a positive relationship between the social-identity parameter α and group investment into the contest. In Appendix B.1 we analyse the sensitivity of group contribution towards α and β, which can visually be represented by a surface in the $\sum_{i\in A} a_i \times \alpha \times \beta$-space as depicted in Figure 18. We normalise the average ingroup Pre-Game SVO angle to a 0-1 scale for each group to determine a group's social-identity parameter α. Consequently, β is the competing group's social-identity parameter.

An important argument in formulating Hypothesis 1 is that making social identity salient increases the identification with the own group's shared social identity, measured

[17] Regression analyses at the individual level deliver qualitatively similar results.

Table 2: Random effects model regressing group contest investment in round t on groups size and controls.

	(1)	(2)
	Group Contribution in t	
Large	25.376***	33.750***
	(6.21)	(11.77)
Female	20.544***	36.097**
	(6.08)	(14.51)
Identity	4.158	24.487*
	(6.85)	(13.02)
Large ×		−10.349
Female		(18.87)
Large ×		−20.451
Identity		(17.12)
Identity ×		−34.822*
Female		(18.65)
Large ×		29.903
Female × Identity		(25.87)
Lagged Other Group	0.124*	0.125**
Contribution	(0.07)	(0.06)
Constant	50.263***	40.637***
	(7.69)	(10.01)
Number of observations	558	558
Number of panels	62	62
Within model R-squared	0.319	0.319
Between model R-squared	0.269	0.315
Overall R-squared	0.287	0.316

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Clustered standard errors in parentheses. Round fixed effects not reported.

by the social-identity parameter α_i. Formal tests confirm that the in-group bias is indeed about 25% larger in the Identity treatments (MWU test. H0: α_i in Control treatments = α_i in Identity treatments, $N = 350,\ z = -3.069,\ p = 0.0021$). Similarly, Hypothesis 2 may be driven by an increase in social identity within larger groups. Non-parametric tests deliver no evidence, however, that the social-identity parameter is different for players in large groups (MWU test using data from Asymmetric treatments only. H0: α_i in small groups = α_i in large groups, $N = 248,\ z = -0.264,\ p = 0.7921$). As such, our results for large groups are likely not driven by social identity, but by other motivations like the power imbalance between groups. As discussed in Subsections 4.2 and 4.3, male and female participants react very differently to the power imbalance in the contest. Similarly, female participants display a significantly higher level of ingroup bias (MWU test. H0: α_i for female players = α_i for male players, $N = 350,\ z = -2.348,\ p = 0.0189$), which can contribute at explaining the overall higher level of contest spending for female participants.

Using these parameters for α and β, we can test the relationship between the actual group investment level on the one hand, and the equilibrium predictions from the social preferences model on the other hand. As first step, pairwise tests indicate that group investment is significantly higher than what would be predicted given the imputed social-identity parameters (Wilcoxon test. H0: group contr. $= \frac{40(1-\beta)}{(2-\alpha-\beta)^2}$, $N = 96,\ z = 7.908,\ p < 0.0001$). Figure 9 depicts the relationship between the observed group investment level in our experiment as a function of average ingroup bias for a given group (α) on the x-axis and the competing group's ingroup bias (β) on the y-axis. The figure illustrates that our empirical data is close to yet mostly above the equilibrium prediction.

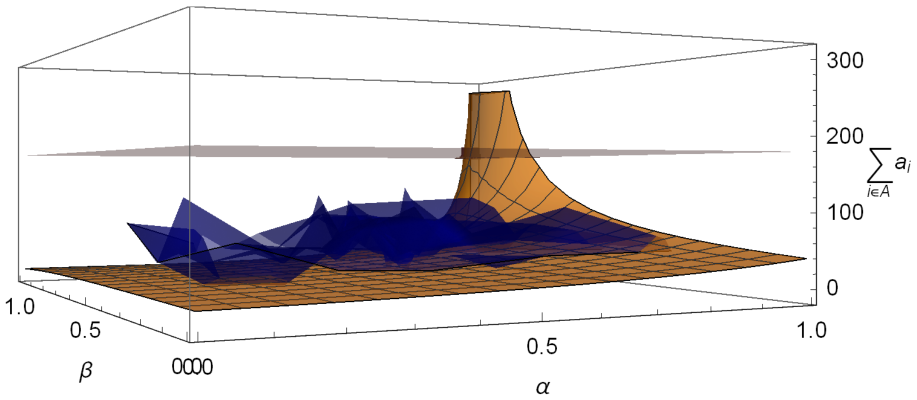


Figure 9: The *orange* surface depicts predicted group contest investment as a function of ingroup bias in the own group (α) and ingroup bias in the other group (β). The *blue* surface illustrates actual group contest investment in our experiment as a function of α and β. The *grey* surface indicates the upper bound for small groups with an endowment of $3 \cdot 60 = 180$ points.

We expand on this by regressing individual investment levels in a given round on α, β and a set of controls using a random effects model with error terms clustered at group level. The associated regression output in Table 3 shows a significant positive effect for the social-identity parameter α on group investment levels in Regressions (1) and (3). When we interact α with an indicator variable for Identity treatments, the overall alpha effect vanishes, which suggests that the social identity parameter only affects contribution in the treatments with salient gender identity. This goes in parallel with a negative baseline effect for Identity treatments, meaning that individuals with a very low alpha level actually invest less into the contest if gender identity is salient. In Regression (3) we assess whether participants' gender identity interacts with the social-identity parameter. While the Pearson's correlation coefficient between α and Female is at 0.136, we find no evidence for a different effect from α between participants with male or female gender identity on contribution. Including this interaction effect, however, renders the positive effect from female participants not significantly different from zero.[18]

4.5 Social Value Orientation

In Subsection 3 we hypothesise that the group contest increases the in-group bias (Hypothesis 3), i.e. the difference between in-group and out-group SVO. If this is true, we should be able to measure a higher in-group bias in the Post-Game SVO when compared to the in-group bias in the Pre-Game SVO.

[18]Regression analyses at the group level deliver qualitatively similar results.

Table 3: Random effects model regressing individual contest investment in round t on social preferences and controls.

	(1)	(2)	(3)
	Contribution in t		
Alpha	8.249^{***}	2.347	6.908^{**}
	(2.58)	(3.79)	(2.94)
Identity	-0.619	-4.965^{*}	-0.639
	(1.52)	(2.75)	(1.53)
Female	3.588^{**}	3.683^{**}	1.993
	(1.47)	(1.47)	(3.09)
Identity × Alpha		9.864^{**}	
		(4.93)	
Female × Alpha			3.467
			(5.51)
Beta	4.754	4.574	4.737
	(4.71)	(4.60)	(4.74)
Contribution other group members $t-1$	0.058^{***}	0.059^{***}	0.058^{***}
	(0.01)	(0.01)	(0.01)
Other Group Contribution $t-1$	0.037^{***}	0.037^{***}	0.037^{***}
	(0.01)	(0.01)	(0.01)
Constant	8.289^{***}	10.621^{***}	8.859^{***}
	(2.64)	(2.94)	(2.87)
Number of observations	3,150	3,150	3,150
Number of panels	350	350	350
Within model R-squared	0.117	0.117	0.117
Between model R-squared	0.111	0.123	0.113
Overall R-squared	0.113	0.120	0.114

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Clustered standard errors in parentheses. Round fixed effects not reported.

We conducted a Wilcoxon test to examine the systematic changes in ingroup bias through the contest. The null hypothesis (H0: Difference Pre-Game-Post-Game in-group bias = 0) was not rejected ($N = 350,\ z = -1.574,\ p = 0.1155$), indicating no significant change overall. Further analysis, considering the identity and control treatments, revealed that for control treatments, this lack of change persisted ($N = 174,\ z = 0.1683,\ p = 0.8663$). However, in the identity treatments, the group contest actually decreased ingroup bias ($N = 176,\ z = 2.2139,\ p = 0.02683$). Notably, we expected the identity treatment to increase ingroup bias, given the salience of gender identity of the players. Unfortunately, we currently lack an explanation for this unexpected result.

Result 4 The contest reduced ingroup bias in the identity treatment but had no impact on ingroup bias in the control treatment.

Our hypothesis regarding the salience of social identity creating a higher ingroup bias at the experiment's start was not confirmed on average (Wilcoxon test. H0: In-group bias in identity treatments > Control treatments, $N = 350,\ z = 1.0620,\ p = 0.1441$). This lack of confirmation is due to differential reactions between males and females to our treatments, effectively averaging each other out. However, we observe that the identity treatment has increased the Ingroup SVO (Wilcoxon test. H0: In-group SVO in identity treatments > Control treatments, $N = 350,\ z = 3.07036,\ p = 0.0010$). Interestingly, Outgroup SVO also increased with the identity treatment ($N = 350,\ z = 2.0813,\ p = 0.0187$), suggesting an order effect, potentially influenced by individuals making decisions relative to the Ingroup SVO test, which was the first they encountered.

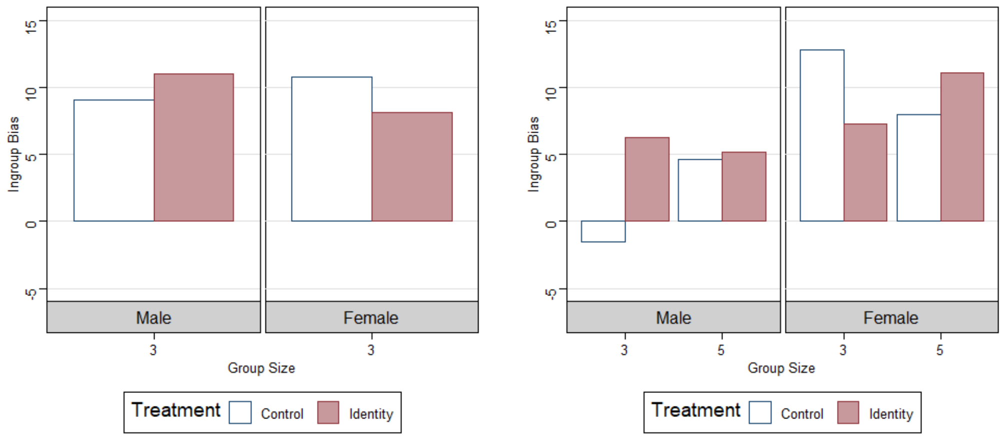


Figure 10: The figure illustrates levels of ingroup bias per treatment for males and females. The left graph represents symmetric treatments, while the right graph represents asymmetric treatments.

For symmetric treatments, revealing gender identity did not affect in-group bias for neither males nor females (Wilcoxon test. H0: In-group bias in Symmetric identity treatments > Symmetric Control treatments, Males: $N = 51,\ z = -1.0581,\ p = 0.855$, Females: $N = 51,\ z = -0.03685,\ p = 0.5147$),shown in Figure 10 left panel. In contrast, introducing asymmetry between groups had significantly different effects on male and female ingroup bias, shwon in Figure 10 right panel. In male-majority asymmetric groups, the identity treatment decreased female ingroup bias ($n = 48, p = 0.03861$),primarily due to a disproportionate increase in Outgroup SVO, while maintaining male ingroup bias at the same level ($n = 80, p = 0.5933$). In contrast, in female-majority groups, both male and female ingroup biases increased, with male bias being twice that of females (Males: $N = 45,\ z = -2.5180,\ p = 0.0059$, Females: $N = 75,\ z = 1.9172,\ p = 0.0276$). This increase was primarily due to a rise in Ingroup SVO.

This finding provides an explanation for Result 2, where female-majority groups participated more in the contest when identity was not revealed. Interestingly, despite females having a higher average ingroup bias than males, for minority female identity treatment caused a significant decrease in their ingroup bias, whereas for males, it increased. This suggests that males can leverage their collective identity more effectively when they are minority. This observation aligns with our earlier result that higher ingroup bias leads to greater effort in the group contest. Therefore, being in the minority has different effects on males and females, potentially resulting in more detrimental outcomes for females as it reduces their ingroup bias compared to males.

Observation 1 The salience of gender identity has varying effects on in-group bias for males and females. In the case of males, disclosing their gender identity either elevated or maintained the in-group bias at the same level. On the other hand, for females, the in-group bias tended to decrease unless they constituted the majority.

4.6 Gender Identity

We employ the Gender Identity Survey(GIS) by Cameron (2004) to measure the degree to which a participant identifies with his or her stated gender identity. This survey measures three key dimensions of gender identity: the significance individuals attribute to it, their affinity for being a part of that particular gender group, and the extent to which they perceive themselves as integral members of that group. While all participants self-identified with either male or female gender identity, we observe some heterogeneity in the degree to which they identify with this social dimension (average score 4.14 on a 0-6 Likert scale, sd $\approx$ 0.827). We observe no significant difference in the gender identity score between treatments (KW Test. $N = 350,\ \chi^2 = 4.290,\ p = 0.5085$) or between Control and Identity treatments (MWU test. $N = 350,\ z = -0.669,\ p = 0.5034$). Given the different timing of the Gender Identity Survey between the Control and Identity treatments (see Figure 3), the fact that the gender identity score is not affected by the treatment manipulation speaks to its stability. Note that female participants display a significantly higher level of identification with their gender identity than male participants do (MWU test. $N = 350,\ z = -3.556,\ p = 0.0004$).

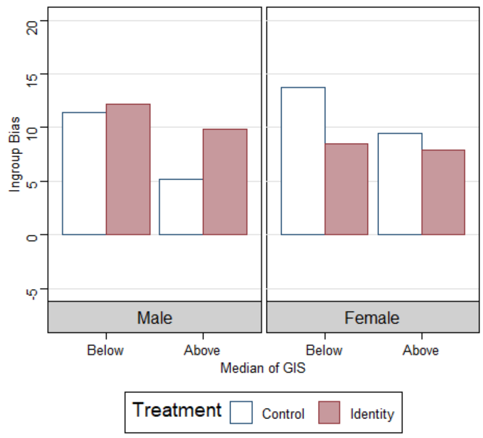


Figure 11: Levels of ingroup bias for males and females in the symmetric treatment split by the median GIS.

We explore the impact of identification with a particular gender identity on ingroup bias and its influence on earlier findings. We split males and females based on the median score from the gender identity survey. In Figure 11, the symmetrical treatment, we observe that disclosing gender identity heightens ingroup bias for males with high GIS scores. Conversely, among females, ingroup bias decreases only for those with low GIS scores. However we are not able to confirm these findings statistically due to small sample (Wilcoxon test, Males above median GIS: $N = 22,\ p = 0.3339$, Females below median GIS: $N = 18,\ p = 0.18$).

In the asymmetric male treatments, where the males are majority, GIS does not impact

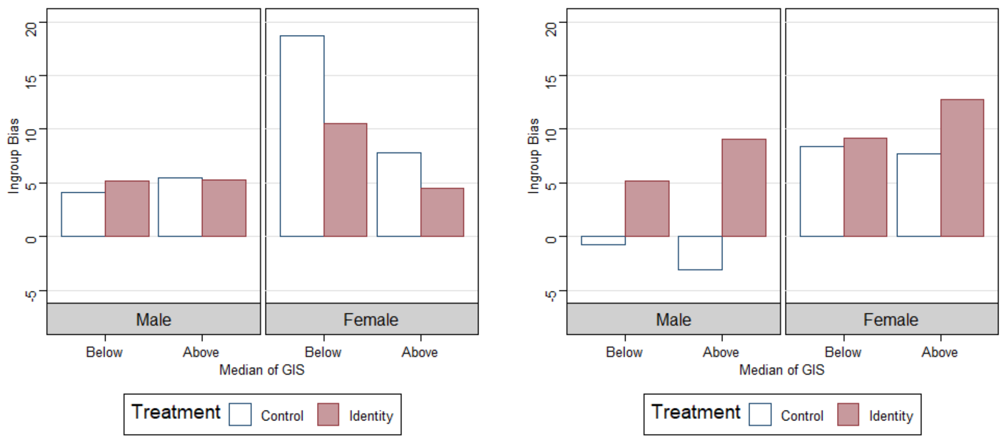


Figure 12: Levels of ingroup bias for males and females in the asymmetric treatment split by the median GIS. Left one is for Asymmetric male treatments and the right one is for the Asymmetric female treatments

male ingroup bias. However the minority female group competing against them experience a reduction in ingroup bias with the identity treatment, and this decrease is more pronounced in the low GIS group, which can be seen in left plot in Figure 12(Wilcoxon test, Females above median GIS: $N = 26,\ p = 0.1517$, Females below median GIS: $N = 22,\ p = 0.1517$). We also observe that compared to low GIS females, high GIS females have overall less ingroup bias when they are in a minority (Wilcoxon test, H0: Ingroup bias above median GIS $<$ Ingroup bias below median GIS : $N = 50,\ p = 0.0586$).

Moving to asymmetric female treatments, there's an increase in ingroup bias for both males with high and low GIS scores, but the impact is notably stronger in the high GIS group as it can be seen in the right panel of in Figure 12 (Wilcoxon test, Males above median GIS: $N = 14,\ p = 0.0165$, Males below median GIS: $N = 31,\ p = 0.0691$). Among females, the rise in ingroup bias is driven by those with high GIS scores (Wilcoxon test, Females above median GIS: $N = 20,\ p = 0.0568$). These trends persist across the three sub-aspects of gender identity. Check Appendix D for an exploration of whether identification with one's gender impacts investment decisions in the contest, where we found no evidence supporting such an impact.

Observation 2 When females are in the minority, feminine individuals tend to demonstrate lower in-group bias. Conversely, when males are in the minority, more masculine individuals are inclined to exhibit heightened in-group bias.

We check if the gender identity scores differ between various demographic characteristics. For some characteristics, such as programme of study (KW Test. $N = 350,\ \chi^2 = 16.998,\ p = 0.1994$), faculty of study (KW Test. $N = 350,\ \chi^2 = 5.749,\ p = 0.2187$) or country/region of origin (KW Test. $N = 350,\ \chi^2 = 2.773,\ p = 0.7349$), there is no difference between the categories. For other demographics, we do find an effect. Bachelor's students, for example, identify significantly stronger with their gender identity than master's student do (Wilcoxon test. H0: gender identity score Bachelor's students = gender identity score Master's students, $N = 342,\ z = -2.807,\ p = 0.005$). Similarly, the gender identity score was stronger for younger participants (Cuzick Test at individual level:

$N = 350,\ z \leq -4.445,\ p < 0.0001$).

5 Conclusion

The "glass ceiling" is a popular metaphor describing the phenomenon of female under-representation in executive positions, providing an allegory to the invisible barrier that prevents women from rising beyond a certain hierarchy level (US Federal Glass Ceiling Commission, 1995). Often, this gender difference in promotion is attributed to a tendency to shy away from competition on the part of females (Lawless & Fox, 2008; Davies-Netzley, 1998). Our study contributes to this conversation by presenting a controlled study investigating the degree to which male and female individuals engage in a between-group contest against players from the opposite gender identity. By varying the salience of gender identity we can analyse if being reminded of and nudged towards gender identity influences the level of investment into the contest and how this interacts with participants' own gender identity.

We study a contest between groups involving irreversible and costly investments. We start by modelling individuals who maximise utility as a weighted sum of own and others' earnings as a function of their investment into the contest. For this, we define the social-identity parameter α as the weight a player puts on their group mates' payoff.

Our results describe how being in a position of power can drive competitiveness such that advantaged groups tend to invest more into the contest. Throughout, larger, more powerful groups invest more into the contest. Interestingly, male and female participants exhibit distinct responses to the salience of gender identity in this context. In contests where females form the majority, disclosing gender identities prompts minority males to double their investment. Conversely, in male-majority contests, the same disclosure results in a reduction of investments by female minorities. Therefore, being a minority appears to pose a greater disadvantage for females, given the differing reactions to this status between the genders. Speaking to this result we also find that feminine females exhibit lower ingroup bias in a minority status, whereas masculine males display higher in-group bias when they are in minority.

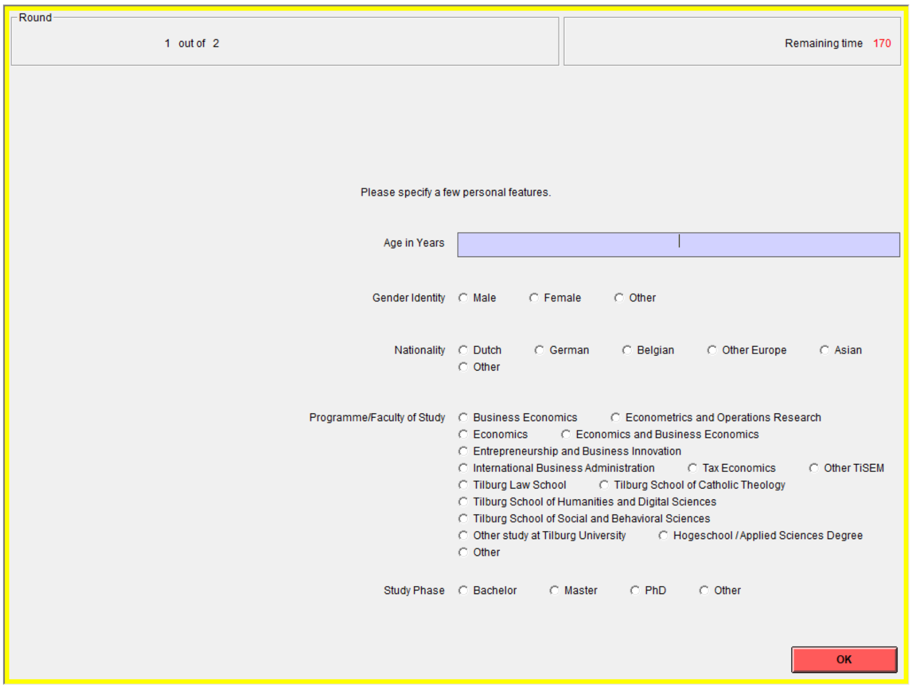
Importantly, this result is not driven via social identity or in-group cohesion. Both our measure for the social-identity parameter α and the Pre-Game and Post-Game SVO measures do not differ by gender identity, and remain stable after the contest.

The design of our study implies a few limitations, some of which may be followed-up by future research. While establishing a solid baseline and maximising the likelihood of triggering in-group bias, our design of a group contest between homogenous social identity groups will probably be an imperfect representation of rent-seeking contests in the field.

Similarly, the induced effort character of our study design in which participants invest abstract points for a contest may represent an imperfect match with many contest situations in the field. Particularly when competition implies the chance for physical harm, prior research suggests that male participants may be more inclined to compete (Hay et al., 2011). Future work may investigate the underlying research question using data from the field. In particular, data on any type of (sports) competition in which male and female groups compete may constitute a valuable extension. Equestrian sports may present a promising application, in which male and female riders compete against each other in various fields (see, e.g., McKenzie, 2013). Other examples are Olympic shooting, in which men and women competed together between 1968-1980, or dog sled racing.

Appendix A Screen Shots

In this section we provide some of the key screen shots for our study. For brevity, we omit minor transition screens, like e.g. a welcome screen and waiting screens.


Round
1 out of 2
Remaining time 170

Please specify a few personal features.

Age in Years

Gender Identity: Male, Female, Other

Nationality: Dutch, German, Belgian, Other Europe, Asian, Other

Programme/Faculty of Study: Business Economics, Econometrics and Operations Research, Economics, Economics and Business Economics, Entrepreneurship and Business Innovation, International Business Administration, Tax Economics, Other TiSEM, Tilburg Law School, Tilburg School of Catholic Theology, Tilburg School of Humanities and Digital Sciences, Tilburg School of Social and Behavioral Sciences, Other study at Tilburg University, Hogeschool / Applied Sciences Degree, Other

Study Phase: Bachelor, Master, PhD, Other

OK

Figure 13: Demographics questionnaire at the start of the experiment. The purpose of this stage was also to check the success of the induced gender identity balance.

Appendix B Equilibrium Strategies

This section presents theoretical predictions and equilibrium strategies for the group contest game building upon methods described in Konrad (2009); Zaunbrecher and Riedl (2016). Similar to Charness and Rabin (2002); Y. Chen and Li (2009); R. Chen and Chen (2011) we model individual utility as a weighted average of own and others' payoff. In particular, we consider the utility function of the form $u_g(i) = (1-\alpha)\cdot\pi_g + \alpha\cdot\overline{\pi}_{A\setminus g}$, with π_g as payoff for player g, $\overline{\pi}_{A\setminus g}$ the average payoff of player g's other group members and $\alpha \in [0,1]$ the strength of g's social identity, where a higher α implies a stronger social identity. Without loss of generality, the following analysis holds true both with or without social preferences. Under individualistic preferences, let $\alpha = 0$. A player maximises the following utility function:

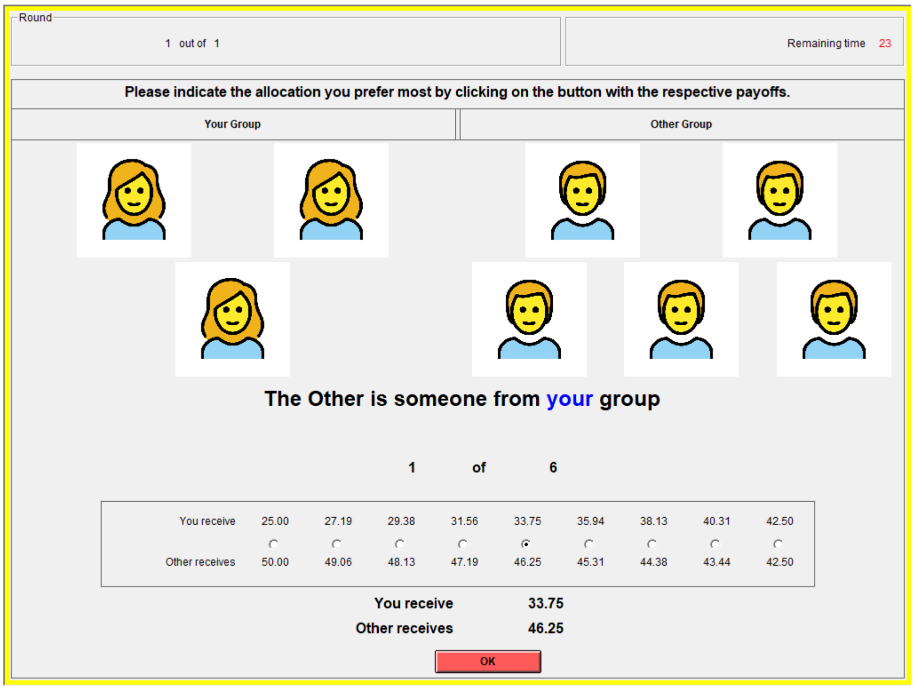


Figure 14: SVO test towards someone from the player's *own* group in the Asymmetric Male Identity treatment. Each player made six allocation decisions towards another player from the *own* group and six comparable decisions towards someone from the *other* group. After the group contest, each player again encountered the same number of decisions towards the *own* and the *other* group.

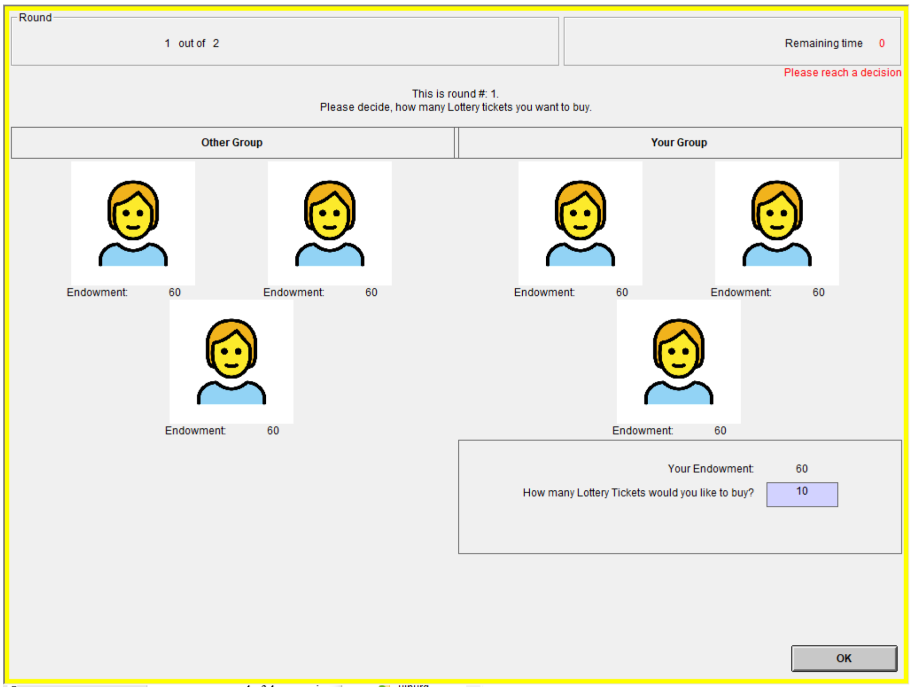


Figure 15: Decision Stage in the Symmetric Control treatment. In this phase, each player decides, how many lottery tickets to buy for the group contest.

Figure 16: Results Stage of a winning group in the Asymmetric Male Identity treatment.

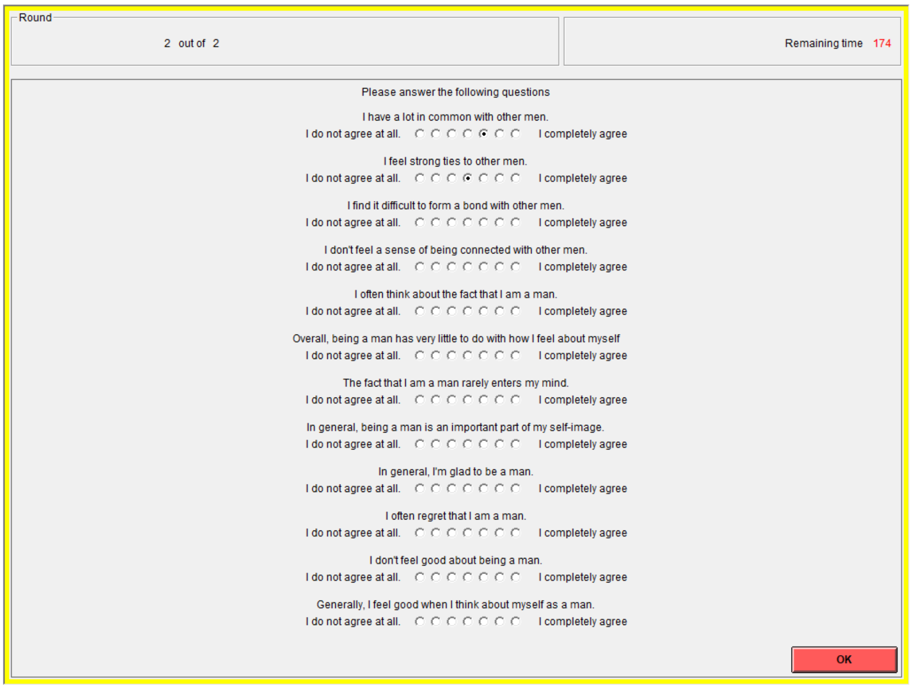


Figure 17: Gender Identity Survey, displayed either at the start or end of the experiment.

$$
\begin{aligned}
u_g\left(\sum_{i\in A} a_i, \sum_{j\in B} b_j\right) =& (1-\alpha)\left[T_i + \frac{\sum_{i\in A} a_i}{\sum_{i\in A} a_i + \sum_{j\in B} b_j}\cdot z_i - a_g\right] + \\
& \frac{\alpha}{N_A - 1}\left[(N_A - 1)\left(T_i + \frac{\sum_{i\in A} a_i}{\sum_{i\in A} a_i + \sum_{j\in B} b_j}\cdot z_i\right) - \sum_{i\in A\backslash g} a_i\right]
\end{aligned}
\tag{7}
$$

Taking the derivative with respect to a_g delivers the first order condition:

$$
\begin{aligned}
& \frac{\partial u_g\left(\sum_{i\in A} a_i, \sum_{j\in B} b_j\right)}{\partial a_g} = 0 \\
\Leftrightarrow\quad & \frac{\sum_{j\in B} b_j}{\left(\sum_{i\in A} a_i + \sum_{j\in B} b_j\right)^2} = \frac{1-\alpha}{z_i} && (8) \\
\Leftrightarrow\quad & \frac{\sum_{j\in B} b_j}{1-\alpha} = \frac{\left(\sum_{i\in A} a_i + \sum_{j\in B} b_j\right)^2}{z_i} && (9)
\end{aligned}
$$

Consider the second derivative to assess if u_g is concave, i.e. whether the first order condition delivers a maximum.

$$\frac{\partial^2 u_g\left(\sum_{i\in A} a_i, \sum_{j\in B} b_j\right)}{\partial a_g^2} = \frac{-2z_i \sum_{i\in A} a_i}{(1-\alpha)\left(\sum_{i\in A} a_i + \sum_{j\in B} b_j\right)^3} < 0 \;\forall\; \sum_{i\in A} a_i + \sum_{j\in B} b_j > 0$$

The function is concave and the extreme point will be a maximum, except for the case when both groups invest zero. $\sum_{i\in A} a_i + \sum_{j\in B} b_j = 0$ cannot be a maximum, though, as it would be individually optimal to deviate from this point, invest one point into the contest and win the prize with certainty. We solve the first order condition (Equation 9) for group contributions in group A:

$$\sum_{i\in A} a_i = \sqrt{\frac{z_i \cdot \sum_{j\in B} b_j}{1-\alpha}} - \sum_{j\in B} b_j \tag{10}$$

Similarly, we derive the best response for an individual b_j from group B with $\beta \in [0,1]$ as measure for social identity, equivalent to α for group A:

$$\frac{\sum_{i\in A} a_i}{1-\beta} = \frac{\left(\sum_{i\in A} a_i + \sum_{j\in B} b_j\right)^2}{z_i} \tag{11}$$

Equating the left-hand sides of Equations 9 and 11 we find that in equilibrium $\sum_{j\in B} b_j = \frac{(1-\alpha)\sum_{i\in A} a_i}{1-\beta}$. Using this, we can solve Equation 10 for

$$\sum_{i\in A} a_i = \frac{z_i (1-\beta)}{(2-\alpha-\beta)^2} \tag{4}$$

and

$$\sum_{j \in B} b_j = \frac{z_i (1 - \alpha)}{(2 - \alpha - \beta)^2} \tag{5}$$

For individualistic players, let $\alpha = 0$ to see that the equilibrium prediction will be $\sum_{i \in A} a_i = \frac{z_i}{4}$. Our model assumes constant marginal costs of investment and a homogeneous social-identity parameter for a given group. The model does allow, though, for different social-identity parameter between the two competing groups, i.e. α may or may not be equal to β. For our result, no further symmetry assumptions are required (Abbink et al., 2010; Konrad, 2009). However, this does not deliver a unique solution for individual contributions as all combinations of $\sum_{i \in A} a_i$ that sum up to $\frac{z_i}{4(1-\alpha)}$ constitute an equilibrium.

Note that equilibrium group contribution α is contingent on the *individual* prize for winning the contest z_i despite being an equilibrium prediction at the group level. As the individual prize remains unchanged between the symmetric and asymmetric treatments, (standard) equilibrium predictions remain the same, irrespective of group size.

B.1 Sensitivity for Social Identity

We next analyse how the social-identity parameter towards the own group α influences contribution decisions, before turning to how the level of social identity β in the competing group B influences investment decisions in group A. Deriving Equation 4 with respect to α delivers

$$\frac{\partial \sum_{i \in A} a_i}{\partial \alpha} = \frac{2}{(2 - \alpha - \beta)^3} \geq 0 \quad \forall \quad 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1. \tag{12}$$

As $\frac{\partial \sum_{i \in A} a_i}{\partial \alpha} \geq 0$ for the levels of α and β considered here, a higher level of social identity towards the own group increases the amount of contest spending. Similarly, we derive Equation 4 with respect to β to get

$$\frac{\partial \sum_{i \in A} a_i}{\partial \beta} = \frac{-z_i (2 - \alpha - \beta) + 2}{(2 - \alpha - \beta)^3}. \tag{13}$$

This is not a monotonous function for large z_i as in our experiment, as illustrated in Figure 18. The graph depicts the equilibrium group contribution $\sum_{i \in A} a_i$ (Equation 4) on the z-axis as a function of own-group (α) on the x-axis and other-group (β) social identity on the y-axis within the range defined by the experiment. In particular, the range is $0 \leq \sum_{i \in A} a_i \leq 300$ for large groups. The upper boundary for small groups at $\sum_{i \in A} a_i = 180$ for small groups is represented by the grey coloured surface. The graph visualises the strictly positive relationship between $\sum_{i \in A} a_i$ and α along the x-axis and the non-monotonous relationship between $\sum_{i \in A} a_i$ and β along the y-axis. Note that for very high levels of α and β, the equilibrium is in a corner solution at $\sum_{i \in A} a_i = 180$ for small groups and $\sum_{i \in A} a_i = 300$ for large groups, respectively, as depicted in the graph.

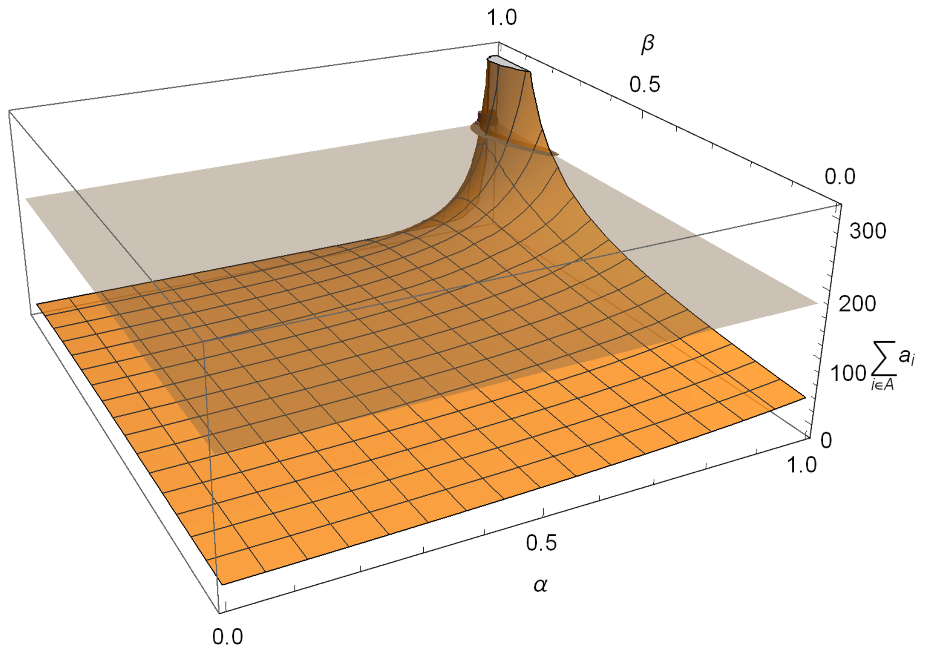


Figure 18: Equilibrium Group Contribution ($\sum_{i \in A} a_i$) as a function of Own-Group (α) and Other-Group (β) Social-Identity Parameter if $z_i = 40$ as in our experiment. The plot range corresponds to the limits defined by the calibrations of the experiment, i.e. $\sum_{i \in A} a_i \in [0, 300]$ for a large group of $n = 5$. The semi-transparent grey surface indicates the upper bound for a small group of $n = 3$.

Appendix C Power Analysis

We follow guidelines formulated by Athey and Imbens (2017); Vasilaky and Brock (2019). In specific, we assume there exists a null hypothesis μ_0 when there is no treatment effect (i.e. in the control group) and an alternative hypothesis μ_1 when there is a treatment effect (in the treatment group). We then investigate the true treatment effect, being $\theta = \mu_1 - \mu_0$ under the null hypothesis (θ_0) that $\mu_1 = \mu_0$. In this section our focus is on understanding the potential Type II error associated with this investigation in the context of our experimental design. We will use the results of this analysis to make an informed decision on the sample size required for reliably investigating our research questions.[19]

Using the result from Chowdhury et al. (2016) we calculate the standardised effect size at the group level between 0.6401-0.8463. We target at a significance level of $\alpha = 0.05$ and statistical power of 0.8. Using the Optimal Design Software (Raudenbush et al., 2011) we calculate that the total number of small groups pairs should be between 12-18 and larger group pairs should be between 6-10. Therefore, in total we require 168-268 participants.

Figure 19 shows the total number of group contest groups required for a given level of power. The n in the legend is the number of participants in a group ($n = 6$ for small group and $n = 8$ for the large group). δ is the standardized effect size. α_δ^2 is effect size variability.

[19]Executing and reporting the results of a cogent power analysis also contributes substantially at qualifying potential null effects (Nikiforakis & Slonim, 2015). As part of the scientific process, well-designed studies with null effects deserve consideration for publication when part of a well-powered study. Ignoring null results in the body of scientific evidence would feed the publication bias.

To be conservative we took R^2 as zero.

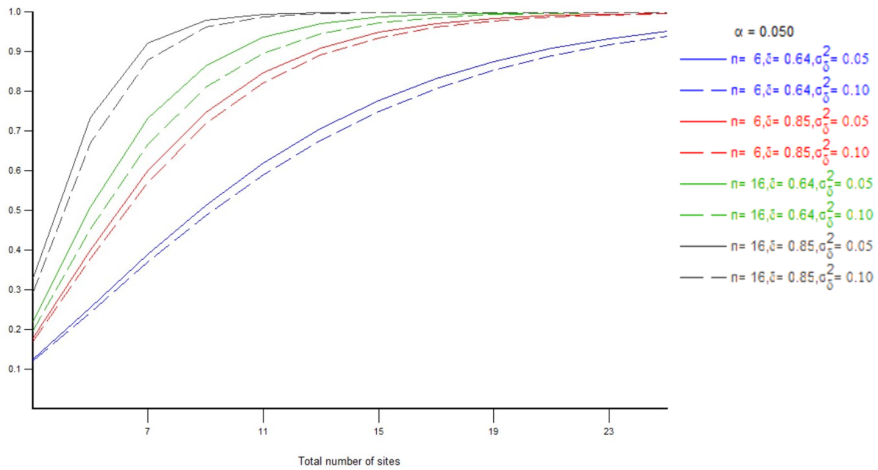


Figure 19: Plot of the Power Analysis. Number of groups (called "sites") on the x-axis, power on the y-axis. The lower four curves (in red and blue colour) represent the power for small groups, the upper four curves (in green and black colour) represent the large groups. Solid lines depict power calculations for an effect size variability of $\alpha_\delta^2 = 0.05$, dashed lines depict power calculations for an effect size variability of $\alpha_\delta^2 = 0.10$

.

Appendix D Results Tables

This appendix complements the results from Section 4 providing tabular representations of the data. Table 4 provides data corresponding to Sub-Figure 6a. While the boxplots depict the median (50th percentile), the 25th and 75th percentiles, as well as the minimum (0th) and maximum (100th percentile) excluding outliers, the following table provides the mean, standard deviation and number of independently distributed observations, i.e. group pairs for each treatment.

Table 4: Group contest investment per group pair averaged over all rounds.

	Average	Standard Deviation	N
Symmetric Control	48.811	28.227	9
Asymmetric Female Control	56.693	8.727	7
Asymmetric Male Control	61.219	23.876	8
Symmetric Identity	54.000	26.283	8
Asymmetric Female Identity	69.287	30.383	8
Asymmetric Male Identity	58.806	11.890	8
Total	57.972	23.117	48

Table 5 presents results from an OLS regression with error terms clustered at group-pair level regressing individual contest investment averaged over the ten rounds of the group contest on the individual gender identity survey score and other factors. This analysis complements the discussion in Subsection 4.6 and shows that the gender identity score does not influence contest investment decisions. The regression does reproduce the stationarity with respect to the investment level of the own (that is investment of *other* group members, i.e. excluding i) and the other group.

Table 5: OLS regression with error terms clustered at group-pair level regressing individual contest investment averaged over all 10 rounds on the gender identity score and other factors.

	(1)	(2)	(3)	(4)
	Average Individual Contribution			
Gender Identity	0.692	0.200	−0.532	−0.195
Survey Score	(0.68)	(1.14)	(1.14)	(1.51)
Identity		−3.059		
		(6.63)		
Identity × Gender		0.897		
Identity Survey Score		(1.47)		
Female			−2.892	
			(7.88)	
Female × Gender			1.819	
Identity Survey Score			(1.91)	
Alpha				3.290
				(12.57)
Alpha × Gender				1.214
Identity Survey Score				(2.91)
Average Contribution	0.075^{**}	0.074^{**}	0.054^{*}	0.075^{**}
Other Groupmates	(0.03)	(0.03)	(0.03)	(0.03)
Average Contribution	0.063^{*}	0.062^{*}	0.083^{**}	0.060^{*}
Other Group	(0.03)	(0.03)	(0.03)	(0.03)
Constant	6.103^{*}	7.915	8.560^{*}	6.116
	(3.36)	(5.20)	(4.95)	(6.96)
Number of observations	350	350	350	350
R-squared	0.050	0.052	0.084	0.087

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Clustered standard errors in parentheses.

References

Abbink, K., Brandts, J., Herrmann, B., & Orzen, H. (2010). Intergroup conflict and intra-group punishment in an experimental contest game. *American Economic Review*, *100*(1), 420–47.

Abdelal, R., Herrera, Y. M., Johnston, A. I., & McDermott, R. (2006). Identity as a variable. *Perspectives on Politics*, *4*(4), 695–711.

Ahn, T., Isaac, R. M., & Salmon, T. C. (2011). Rent seeking in groups. *International Journal of Industrial Organization*, *29*(1), 116–125.

Akerlof, G. A., & Kranton, R. (2010). Identity economics. *The Economists' Voice*, *7*(2).

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, *115*(3), 715–753.

Akerlof, G. A., & Kranton, R. E. (2002). Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature*, *40*(4), 1167–1201.

Akerlof, G. A., & Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, *19*(1), 9–32.

Athey, S., & Imbens, G. W. (2017). The econometrics of randomized experiments. In *Handbook of economic field experiments* (Vol. 1, pp. 73–140). Elsevier.

Barone, C., & Assirelli, G. (2020). Gender segregation in higher education: an empirical test of seven explanations. *Higher Education*, *79*(1), 55–78.

Basu, K. (2005). Racial conflict and the malignancy of identity. *The Journal of Economic Inequality*, *3*(3), 221–241.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

Cadsby, C. B., Servátka, M., & Song, F. (2013). How competitive are female professionals? a tale of identity conflict. *Journal of Economic Behavior & Organization*, *92*, 284–303.

Cameron, J. E. (2004). A three-factor model of social identity. *Self and Identity*, *3*(3), 239–262.

Cason, T. N., Sheremeta, R. M., & Zhang, J. (2012). Communication and efficiency in competitive coordination games. *Games and Economic Behavior*, *76*(1), 26–43.

Chakravarty, S., Fonseca, M. A., Ghosh, S., & Marjit, S. (2016). Religious fragmentation, social identity and cooperation: Evidence from an artefactual field experiment in india. *European Economic Review*, *90*, 265–279.

Charness, G., Gneezy, U., & Halladay, B. (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior & Organization*, *131*, 141-150. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167268116301779 doi: https://doi.org/10.1016/j.jebo.2016.08.010

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, *117*(3), 817–869.

Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, *101*(6), 2562–89.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431–57.

Chowdhury, S. M., Jeon, J. Y., & Ramalingam, A. (2016). Identity and group conflict. *European Economic Review*, *90*, 107–121.

Chowdhury, S. M., Sheremeta, R. M., & Turocy, T. L. (2014). Overbidding and overspreading in rent-seeking experiments: Cost structure and prize allocation rules. *Games and Economic Behavior*, *87*, 224–238.

Costa-Font, J., & Cowell, F. (2015). Social identity and redistributive preferences: a survey. *Journal of Economic Surveys*, *29*(2), 357–374.

Cotter, D. A., Hermsen, J. M., & Vanneman, R. (2000). Gender inequality at work. *The American People: Census*, 107–138.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–74.

Cuzick, J. (1985). A wilcoxon-type test for trend. *Statistics in Medicine*, *4*(4), 543–547.

Davies-Netzley, S. A. (1998). Women above the glass ceiling: Perceptions on corporate mobility and strategies for success. *Gender & society*, *12*(3), 339–355.

Dechenaux, E., Kovenock, D., & Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, *18*(4), 609–669.

Delgado, M. R., Schotter, A., Ozbay, E. Y., & Phelps, E. A. (2008). Understanding overbidding: using the neural circuitry of reward to design economic auctions. *Science*, *321*(5897), 1849–1852.

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, *118*(3), 1049–1074.

Griesinger, D. W., & Livingston Jr, J. W. (1973). Toward a model of interpersonal motivation in experimental games. *Behavioral Science*, *18*(3), 173–188.

Hay, D. F., Nash, A., Caplan, M., Swartzentruber, J., Ishikawa, F., & Vespo, J. E. (2011). The emergence of gender differences in physical aggression in the context of conflict between young peers. *British Journal of Developmental Psychology*, *29*(2), 158–175.

Heine, F., & Sefton, M. (2018). To tender or not to tender? deliberate and exogenous sunk costs in a public good game. *Games*, *9*(3), 41.

Hurst, C., Gibbon, H. F., & Nurse, A. (2016). *Social inequality: Forms, causes, and consequences*. Routledge.

Katz, E., Nitzan, S., & Rosenberg, J. (1990). Rent-seeking for pure public goods. *Public Choice*, *65*(1), 49-60. Retrieved from http://dx.doi.org/10.1007/BF00139290 doi: 10.1007/BF00139290

Kolmar, M., & Wagener, A. (2019). Group identities in conflicts. *Homo Oeconomicus*, *36*(3-4), 165–192.

Konrad, K. A. (2009). *Strategy and dynamics in contests*. Oxford University Press.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621.

Lang, I. (2010). Targeting inequity: The gender gap in us corporate leadership. *Statement made before http://www.jec.senate.gov/public/index.cfm*.

Lawless, J. L., & Fox, R. L. (2008). Why are women still not running for public office?

Li, S. X. (2020). Group identity, ingroup favoritism, and discrimination. *Handbook of Labor, Human Resources and Population Economics*, 1–28.

Liebrand, W. B. (1984). The effect of social motives, communication and group size on behaviour in an n-person multi-stage mixed-motive game. *European Journal of Social Psychology*, *14*(3), 239–264.

Lim, W., Matros, A., & Turocy, T. L. (2014). Bounded rationality and group size in tullock contests: Experimental evidence. *Journal of Economic Behavior & Organization*, *99*, 155–167.

Mago, S. D., Samak, A. C., & Sheremeta, R. M. (2016). Facing your opponents: Social identification and information feedback in contests. *Journal of Conflict Resolution*, *60*(3), 459–481.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50-60.

Masiliunas, A., Mengel, F., & Reiss, J. P. (2014). *Behavioral variation in tullock contests* (Tech. Rep.). KIT working paper series in economics.

McKenzie, S. (2013). Racing's battle of the sexes–on four legs. *CNN international edition*.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision Making*, *6*(8), 771–781.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1–16.
Nikiforakis, N., & Slonim, R. (2015). *Editors' preface: statistics, replications and null results.* Springer.
OpenMoji. (2020). *Open source emojis for designers, developers and everyone else!* Retrieved 2020-08-13, from https://openmoji.org/
Powlishta, K. K., Serbin, L. A., Doyle, A.-B., & White, D. R. (1994). Gender, ethnic, and body type biases: The generality of prejudice in childhood. *Developmental Psychology*, *30*(4), 526.
Price, C. R., & Sheremeta, R. M. (2015). Endowment origin, demographic effects, and individual preferences in contests. *Journal of Economics & Management Strategy*, *24*(3), 597–619.
Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X., Martinez, A., Bloom, H., & Hill, C. (2011). *Optimal design software for multi-level and longitudinal research (version 3.01)[software].*
Sen, A. (2007). *Identity and violence: The illusion of destiny.* Penguin Books India.
Sheremeta, R. M. (2010). Experimental comparison of multi-stage and one-stage contests. *Games and Economic Behavior*, *68*(2), 731–747.
Sheremeta, R. M. (2011). Contest design: An experimental investigation. *Economic Inquiry*, *49*(2), 573–590.
Sheremeta, R. M. (2018). Behavior in group contests: A review of experimental research. *Journal of Economic Surveys*, *32*(3), 683–704.
Shor, E., Van De Rijt, A., Miltsov, A., Kulkarni, V., & Skiena, S. (2015). A paper ceiling: Explaining the persistent underrepresentation of women in printed news. *American Sociological Review*, *80*(5), 960–984.
Sutter, M., & Strassmair, C. (2009). Communication, cooperation and collusion in team tournaments—an experimental study. *Games and Economic Behavior*, *66*(1), 506–525.
Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational Identity: A Reader*, *56*, 65.
Tullock, G. (1980). Efficient rent seeking. In J. Buchanan, R. Tollison, & G. Tullock (Eds.), *Toward a theory of the rent-seeking society* (p. 97 - 112). Texas A & M University Press.
US Federal Glass Ceiling Commission. (1995). A solid investment: Making full use of the nation's human capital, final report of the commission. *Washington, DC: US Government Printing Office. Downloaded September*, *10*, 2007.
Vasilaky, K., & Brock, J. M. (2019). Power (ful) guidelines for experimental economists. *EBRD Working Paper No. 239*.
Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), pp. 80-83. Retrieved from http://www.jstor.org/stable/3001968
Zaunbrecher, H., & Riedl, A. (2016). Social identity and group contests. *Available at SSRN 2816038*.