Backlash against Women in Competition

Simone Haeckl, Jakob Möller, Anita Zednik

December 3, 2021

Abstract

The gender pay gap remains persistent in most developed countries and a large proportion of the gap cannot be explained by observable factors. The repeated finding that women compete less than men has been suggested as a possible factor contributing to the lower wages of women and policy measures have been implemented to encourage women to enter more highly-competitive, well-paid, and often male-stereotypical careers. However, literature on *backlash* against *agentically* acting women suggests that women who choose to compete in male-typed domains may receive punishment for breaking social norms. With this experiment we investigate how people respond to income inequalities arising from competition between men and women when competition is voluntary and competitors act agentically.

Contents

1 Background and Motivation

A strand of literature from psychology with some more recent additions from economics studies how women receive negative reciprocity for 'acting male', that is, in ways that do not conform with the gender stereotype. Women have been found to receive *backlash* for acting *agentically*, i.e. acting in an assertive, dominant, self-promoting, self-interested or competitive way. It is argued that, while this type of behaviour is expected of men, it is in disagreement with the prevailing social norm for behaviour of women and may be met with a negative response (Rudman 1998; Rudman and Phelan 2008). For instance, women are liked less and hired less when they assert agentic behaviour (Rudman and Glick 2001; Heilman, Wallen, et al. 2004; Heilman and Okimoto 2007), women are punished more for initiating negotiations (Bowles, Babcock, and Lai 2007) and are more likely to receive punishment for misconduct (Egan, Matvos, and Seru 2017).

In this experiment, we test for the existence of backlash against women after agentic behavior by systematically comparing treatments with increasingly agentic behavior.

Since women are punished for acting agentically and competitive behaviour is seen as agentic behaviour, there is reason to believe that women experience backlash when they choose to compete. This punishment for competitive behaviour amongst women may be part of the reason for why women have been found to compete less than men, which may ultimately be contributing to the gender pay gap (Gneezy, Niederle, and Rustichini 2003; Niederle and Vesterlund 2007; Masclet, Peterle, and Larribeau 2015). For this reason, we study how individuals respond to income inequalities arising from competition between a male and a female workers while varying the extent to which workers behave agentically.

2 Research Question and Hypotheses

With this study we aim to address the following research question:

Do women experience backlash when acting agentically?

In this context, we define *backlash* as a harsher punishment for agentic behaviour of women than of men. We study an environment in which impartial spectators redistribute between two workers whose incomes have been determined in a competition where the winning worker receives a higher income than the losing worker. Punishment for agentic behaviour in our design means that less money is redistributed towards a worker who has

behaved agentically than to a worker who has not behaved agentically. In our design, redistributing less to a female loser than to a male loser is equivalent to taking more from the female winner than the male winner.

Based on the findings of previous studies on backlash against women such as Rudman (1998), Rudman and Glick (2001), Heilman, Wallen, et al. (2004), Heilman and Okimoto (2007), and Egan, Matvos, and Seru (2017), we formulate the following hypotheses:

2.1 Primary Hypotheses

1. The reduction in redistribution with more agentic behaviour between treatments is stronger for female losers than for male losers.

    1a. Spectators reduce their redistribution to female losers more than to male losers in a mixed gender setting when competition is voluntary, compared to when competition is mandatory.

    1b. Spectators reduce their redistribution to female losers more than to male losers when losers make use of the option to behave selfishly, compared to when they do not have the option to behave selfishly, in a voluntary competition.

    1c. Spectators reduce their redistribution to female losers more than to male losers when losers send a high-dominance message, compared to when they send a low-dominance message.

    1d. Spectators reduce their redistribution to female losers more than to male losers after they received a high-dominance message form the loser, compared to before they received any message.

Cappelen, Falch, and Tungodden (2019) (henceforth CFT) find that men are punished more than women for losing a competition. They term this the "boy crisis". If the effects specified by hypothesis 1 are large enough, they can reverse the boy crisis effects found by CFT. This leads to the following hypotheses.

2. Spectators redistribute less to female losers than to male losers.

    2a. Spectators redistribute less to female losers than to male losers in a mixed gender setting when competition is voluntary.

    2b. Spectators redistribute less to female losers than to male losers in a mixed gender setting when the loser behaves selfishly.

2c. Spectators redistribute less to female losers than to male losers in a mixed gender setting when the loser sends a high-dominance message.

2.2 Secondary Hypotheses

There is evidence for the existence of a social norm in Western society that people should be held accountable for their own choices (Cappelen, Fest, et al. 2016). Based on this, we expect that spectators redistribute less towards low-income individuals who chose to compete for their income, than to those who did not have a choice. Further, there is evidence that people have concerns for fairness which can be in conflict with rational, self-interest maximizing behaviour (Fehr and Schmidt 1999; Charness, Rustichini, and Van de Ven 2018; Buser, Cappelen, and Tungodden 2021). Accordingly, we expect there to be negative reciprocity in the form of less redistribution towards low-income individuals who have acted selfishly in a competition. Based on this, we formulate the following hypotheses.

3. Workers get punished for agentic behavior.

    3a. Spectators redistribute less to losers when they have chosen competition compared to when they have been forced to compete.

    3b. Spectators redistribute less to losers when they have behaved selfishly compared to when they have not behaved selfishly in a voluntary competition.

The methods we use to test these hypotheses are described below in section 3.2.

3 Research Strategy

3.1 Experiment Design

Our experiment design is closely aligned with CFT, and builds upon a previous experiment which was preregistered in December 2020 (Haeckl, Möller, and Zednik 2020).

3.1.1 Setup

Our study consists of two experiments: the *workers experiment* and the *spectators experiment*. For logistical reasons the spectators experiment will be conducted before the workers experiment.

In the *spectators experiment* the participant (spectator) is randomly assigned a pair of workers who both completed a knowledge task about sports and games. Of the two workers, the one who answered more questions correctly has earned £6 and the other has earned £0. The spectator learns the gender of the two workers and their earnings from the competition. Then the spectator chooses how much $R \in [0, 6]$ of the winner's earnings to redistribute to the loser.

In the *workers experiment* participants (workers) answer a quiz with 20 multiple choice questions about sports and games. Subsequently each worker makes choices about how his/her payoff is determined and has the opportunity to send one of two messages to the spectator who will decide about the workers' payoffs. The choice set for how the payoff is determined available to the workers varies based on the treatment.

The final payoff of the workers who competed is determined by the decision of a spectator. We therefore first collect the decisions of spectators. Subsequently, we conduct the workers experiment and match each pair of workers with a randomly selected spectator decision from the respective treatment. Spectators are told in advance that a subset of all spectator decisions is randomly selected to be applied to the payoff of participants from the workers experiment. All workers are told that, depending on their decisions, their final income may be determined by the decision of an impartial spectator.

3.1.2 Treatments

Each spectator is randomly assigned a pair of workers who competed against one another. Half of the spectators will be assigned a pair in which the loser is a woman; the other half will be assigned a pair in which the loser is a man. We conduct the following three between-subject treatments:

In the *Mandatory Competition (MC) treatment* we replicate CFT. Spectators redistribute incomes of workers who do not have a choice but are always matched with another worker for a competition which determines the workers' payoffs.

In the *Voluntary Competition (VC) treatment* spectators redistribute incomes of workers who have chosen the competition over a piece-wise payment scheme.

In the *Selfish Competition (SC) treatment* spectators also redistribute incomes of workers who have chosen the competition over a piece-wise payment scheme. In addition, the losing worker has decided to buy a *sabotage coin* for 10 Pence. This sabotage coin gave the worker a 50% chance to win the competition irrespective of his or her performance.

However, the worker has lost nevertheless.

On top of these three between-subject treatments, we also conduct a within-subject treatment. After the initial redistribution decision described above, spectators are randomly sorted into either the *High-Dominance Message (HDM) treatment* or the *Low-Dominance Message (LDM) treatment*. In the HDM treatment, spectators receive a message from the loser which signals high dominance (and thereby is agentic). In the LDM treatment, the message sent by the loser signals low dominance. The message is displayed to the spectator after he/she has made an initial redistribution decision. After having received the message, the spectator has the chance to change the redistribution decision previously made. This way, we collect two redistribution decisions from each spectator, one made before and one made after having received a message from the losing worker. Table 1 gives an overview of the structure of the treatments in the experiment.

Finally, we implement a *veil of risk* (Exley 2016; Coffman, Exley, and Niederle 2021) across all treatments. This means that all spectators are informed that, for any amount of money they choose to transfer from the winning worker to the losing worker, there is a 99% chance that the loser receives the money and a 1% chance that the money is lost, i.e. neither the winner nor the loser receives it.

3.2 Hypothesis Tests

After data collection, we will use the following methods to test the hypotheses specified above.

To test hypothesis 1, we run the following regression with an interaction term for treatment and gender of the loser on the full sample of this experiment. The dependent variable is the amount redistributed by spectator i. R_{1i} denotes the amount redistributed before the spectator i received any message from the loser, and R_{2i} is the amount redistributed after having received a message. voluntary competition$_i$ is a dummy variable that equals 1 if i is in a treatment with voluntary competition, i.e., VC or SC. selfish$_i$ is an indicator for being in the SC treatment. X_i denotes a vector of control variables such as gender of the spectator, socio-demographic variables and attitudes. We will run all regressions both

with and without control variables.

$$\begin{aligned} \mathrm{R}_{1i} = \beta_0 + \beta_1 \text{female loser}_i + \beta_2 \text{voluntary competition}_i + \beta_3 \text{selfish}_i \\ + \beta_4 \text{voluntary competition}_i * \text{female loser}_i + \beta_5 \text{selfish}_i * \text{female loser}_i \\ + \beta_6 \mathrm{X}_i + \epsilon_i \end{aligned} \quad (1)$$

For hypothesis 1a, we run the following regression on a sample only including subjects in VC and MC. VC_i is a dummy variable indicating that a individual i is in the VC treatment.

$$\mathrm{R}_{1i} = \beta_0 + \beta_1 \text{female loser}_i + \beta_2 \mathrm{VC}_i + \beta_3 \text{female loser}_i * \mathrm{VC}_i + \beta_4 \mathrm{X}_i + \epsilon_i \quad (2)$$

This allows to isolate the effect of women entering into competition voluntarily on the amount redistributed. According to hypothesis 1a we expect coefficient β_3 to be negative.

For hypothesis 1b, we run regression 3 on a sample only including subjects in VC and SC. SC_i is a dummy variable indicating individual i is in the SC treatment.

$$\mathrm{R}_{1i} = \beta_0 + \beta_1 \text{female loser}_i + \beta_2 \mathrm{SC}_i + \beta_3 \text{female loser}_i * \mathrm{SC}_i + \beta_4 \mathrm{X}_i + \epsilon_i \quad (3)$$

This allows to isolate the effect of selfish behaviour by women on the amount redistributed when competition is voluntary. Hypothesis 1b predicts coefficient β_3 to be negative.

For hypothesis 1c we run regression 4 on the full sample (including all three treatments). HDM_i is a dummy variable indicating that individual i is in the HDM treatment.

$$\mathrm{R}_{2i} = \beta_0 + \beta_1 \text{female loser}_i + \beta_2 \mathrm{HDM}_i + \beta_3 \text{female loser}_i * \mathrm{HDM}_i + \beta_4 \mathrm{X}_i + \epsilon_i \quad (4)$$

Here, coefficient β_3 captures the effect on the amount redistributed of female workers sending a high-dominance rather than a low-dominance message to spectators. Regressions including control variables, also include an indicator of the treatment.

To test hypothesis 1d, we use a within-subject design. For each spectator, we elicit a redistribution decision before the spectator has received a message from the loser and after the spectator has received a message from the loser. We define

$$\Delta \mathrm{R}_i = \mathrm{R}_{1i} - \mathrm{R}_{2i}$$

to be the change in the amount redistributed by spectator i upon receiving a message from

the loser. According to hypothesis 1d, we expect ΔR_i to be more negative for female losers than for male losers. We use a two-sample t-test on the full sample to test this hypothesis.

To test hypotheses 2 and 3, we use a Wilcoxon rank-sum test to test for differences in the distributions of the amount redistributed (R_i) to losers between treatments[1]. In addition, we use regressions with R_i as the dependent variable and a dummy for treatment.

As a robustness check we will also perform all analyses with the proportion of spectators who redistributed a positive amount as the outcome variable[2].

3.3 Heterogeneous Effects

We expect heterogeneous effects based on the spectator's gender and the strength of the social norms on competitive behaviour of women held by spectators. We will elicit the strength of social norms for spectators and test for different magnitudes of effects based on social norms using a median split.

3.4 Sampling and Power

With this study, we build upon data collected in a previous experiment preregistered in December 2020 (Haeckl, Möller, and Zednik 2020). With the previous experiment, we collected data on 2400 spectators who all received the VC treatment. 1600 of the 2400 spectators were men. The previous experiment did not include a veil of risk and will serve as the control group for comparison to the data collected in this experiment (which includes the veil of risk). As we restrict participants of this experiment to a narrower age group, we will only compare workers of a similar demographic. If we find no effect of the veil of risk on spectator decisions, we will pool the data for the voluntary competition treatment from the previous and the current experiment.

In this experiment, we will collect data on another 5880 spectators using the online-study platform Prolific. We only include participants from the UK aged between 30-60 years and stratify treatments by spectator age and gender. The spectators will be randomly assigned either a pair of workers with a male or a female loser and to either the MC, VC or SC treatment; see Table 1. This way, there are 1960 participants receiving each treatment, with an even distribution of male and female losers within each treatment. In addition, the

[1]For hypotheses 3a and 3b we will perform the test once on the full sample and once on the sample including only female losers, whom we expect to be driving the result.

[2]That is, we will estimate binary response regressions with the outcome variable of whether individual i chose to redistribute or not.

spectators will be assigned to either the HDM treatment (2940 observations) or the LDM treatment (2940 observations). All spectators in the current experiment are treated with the veil of risk.

	MC		VC		SC	
	LDM	HDM	LDM	HDM	LDM	HDM
Male loser	490	490	490	490	490	490
Female loser	490	490	490	490	490	490

Table 1: Between-subject treatments

The power analysis below was performed using Optimal Design Software and simulations based on the data of Haeckl, Möller, and Zednik (2020). For direct tests within treatments (hypotheses 2a - 2b) ($N = 1600$), we can detect an effect of 0.13 standard deviations with a power of 80% and $\alpha = 0.05$. We expect this minimal detectable effect size to be further reduced when we include control variables in the regressions. For hypothesis 2c, we can use the full sample ($N = 5880$) and have a minimal detectable effect size of 0.08 standard deviations using the same assumptions as before.

For the interaction terms in hypotheses 1a - 1b the power calculation requires more assumptions. For simplicity, we assume that acting agentically does not affect the amount redistributed to male losers. In addition, we take the average amount redistributed in Haeckl, Möller, and Zednik (2020) as the baseline amount redistributed in the voluntary competition treatment, i.e., 1.61 pounds for female losers and 1.41 for female losers. Given these assumptions and a sample size of 3920, we can detect a β_3 of -0.18 with a power of 80% and an alpha of 0.05. To test hypotheses 1c and 1d, we can use the full sample ($N = 5880$) and can therefore detect an effect size of 0.15 using the same assumptions as for hypotheses 1a - 1b. We expect the minimal detectable effect size to be further reduced when we include control variables in the regressions.

Lastly, for the secondary hypotheses (hypotheses 3a and 3b), we make between treatment comparisons and use again Optimal Design Software. When comparing two treatments we have a total N of 3920 and can detect an effect size of 0.09 standard deviations with a power of 80% and $\alpha = 0.05$.

References

Bowles, Hannah Riley, Linda Babcock, and Lei Lai (2007). "Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask." In: *Organizational Behavior and human decision Processes* 103.1, pp. 84–103.

Buser, Thomas, Alexander W Cappelen, and Bertil Tungodden (2021). "Fairness and Willingness to Compete." In: *NHH Dept. of Economics Discussion Paper* 08.

Cappelen, Alexander W, Ranveig Falch, and Bertil Tungodden (2019). "The boy crisis: Experimental evidence on the acceptance of males falling behind." In: *NHH Dept. of Economics Discussion Paper* 06.

Cappelen, Alexander W, Sebastian Fest, Erik Ø Sørensen, and Bertil Tungodden (2016). "Choice and personal responsibility: What is a morally relevant choice?" In: *The Review of Economics and Statistics*, pp. 1–35.

Charness, Gary, Aldo Rustichini, and Jeroen Van de Ven (2018). "Self-Confidence and Strategic Behavior." In: *Experimental Economics* 21.1, pp. 72–98.

Coffman, Katherine B, Christine L Exley, and Muriel Niederle (2021). "The role of beliefs in driving gender discrimination." In: *Management Science*.

Egan, Mark L, Gregor Matvos, and Amit Seru (2017). *When Harry fired Sally: The double standard in punishing misconduct*. Tech. rep. National Bureau of Economic Research.

Exley, Christine L (2016). "Excusing selfishness in charitable giving: The role of risk." In: *The Review of Economic Studies* 83.2, pp. 587–628.

Fehr, Ernst and Klaus M Schmidt (1999). "A theory of fairness, competition, and cooperation." In: *The quarterly journal of economics* 114.3, pp. 817–868.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini (2003). "Performance in competitive environments: Gender differences." In: *The quarterly journal of economics* 118.3, pp. 1049–1074.

Haeckl, Simone, Jakob Möller, and Anita Zednik (2020). *Participation in voluntary competitions and gender discrimination*. AEA RCT Registry. DOI: `https://doi.org/10.1257/rct.6845-1.2000000000000002`.

Heilman, Madeline E and Tyler G Okimoto (2007). "Why are women penalized for success at male tasks?: the implied communality deficit." In: *Journal of applied psychology* 92.1, p. 81.

Heilman, Madeline E, Aaron S Wallen, Daniella Fuchs, and Melinda M Tamkins (2004). "Penalties for success: reactions to women who succeed at male gender-typed tasks." In: *Journal of applied psychology* 89.3, p. 416.

Masclet, David, Emmanuel Peterle, and Sophie Larribeau (2015). "Gender differences in tournament and flat-wage schemes: An experimental study." In: *Journal of Economic Psychology* 47, pp. 103–115.

Niederle, Muriel and Lise Vesterlund (2007). "Do women shy away from competition? Do men compete too much?" In: *The quarterly journal of economics* 122.3, pp. 1067–1101.

Rudman, Laurie A (1998). "Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management." In: *Journal of personality and social psychology* 74.3, p. 629.

Rudman, Laurie A and Peter Glick (2001). "Prescriptive gender stereotypes and backlash toward agentic women." In: *Journal of social issues* 57.4, pp. 743–762.

Rudman, Laurie A and Julie E Phelan (2008). "Backlash effects for disconfirming gender stereotypes in organizations." In: *Research in organizational behavior* 28, pp. 61–79.