Are Men Less Generous to a Smarter Woman? Evidence from a Dictator Game Experiment

Yuki Takahashi[*]

April 5, 2025

Abstract

Although evidence suggests men are more generous to women than to men, it may stem from paternalism and could reverse when women excel in important skills for one's career success, such as cognitive skills. Using a dictator game, this paper studies whether male dictators allocate less to female receivers than to male receivers when these receivers have higher IQs than dictators. By exogenously varying the receivers' IQ relative to the dictators', I do not find evidence consistent with this hypothesis; if anything, male dictators allocate slightly more to female receivers with higher IQs than to male receivers with equivalent IQs. The results hold both in mean and distribution and are robust to the so-called "beauty premium." Also, female dictators' allocations are qualitatively similar to male dictators. These findings suggest that women who excel in cognitive skills may not receive less favorable treatment than equally intelligent men in the labor market.

JEL Classification: D91, C91, J16
Keywords: Gender, IQ, dictator game, laboratory experiment

"All through my life, culturally reinforced signals cautioned me against being branded as too smart or too successful."

— Sheryl Sandberg *Lean In: Women, Work, and the Will to Lead*

1 Introduction

Cognitive skills are important personal attributes that affect one's career success (Herrnstein and Murray 1996). However, individual skills are not the only determinant of one's success: a positive and welcoming environment is also important, especially for jobs based on teamwork. Yet, colleagues' support may also be affected by gender norms, which makes the role of cognitive skills and its interplay with gender possibly more nuanced (Eagly and Karau 2002; Heilman 2001; Ridgeway 2001; Rudman and Phelan 2008). Although the literature finds that men are more generous to women than to men (Dufwenberg and Muren 2006; List 2004), few studies examine its interaction with cognitive skills.

One manifestation of gender norms is paternalism towards women, which may motivate men to be more generous to them. For example, many men (and women) in the US believe that women who fall behind deserve more support than men in similar situations because they attribute women's setbacks to bad luck rather than a lack of effort (Cappelen, Falch, and Tungodden 2023). Similarly, men (and women) are more likely to attribute female leaders' bad outcomes to bad luck rather than poor decisions, compared to how they attribute male leaders' bad outcomes (Erkal, Gangadharan, and Koh 2023). Furthermore, many countries prohibit women from engaging in certain activities, such as close military combat (Fitriani, Cooper, and Matthews 2016).[1]

The main research question addressed in this paper is whether men are less generous to women who have higher cognitive skills than they do. I hypothesize that this is indeed the case because if the paternalism derived from gender norms drives men's higher generosity towards women, their generosity may disappear or even reverse when women violate these norms by excelling in critical skills for their career success. I test this hypothesis via a laboratory experiment.

In the experiment, participants first work on an incentivized IQ test (Raven matrices) that measures their cognitive skills. After the test, participants are randomly assigned to a group of six and receive an IQ rank relative to other group members. Then three of the six members are randomly chosen to be dictators and play three rounds of dictator game with the other three members chosen to be receivers, observing the receivers' facial photos and first names – both of which convey information about their gender – and the relative IQ ranks. I use the dictators' allocation as the measure of generosity, a widely used tool in experimental economics and shown to predict one's generosity outside the laboratory (Franzen and Pointner 2013). The use of photos allows me to inform the dictators of the receivers' gender with minimum experimenter demand

---

1. In fact, most countries exhibit protective attitudes toward women (Glick et al. 2000) and such attitudes are more pronounced in low-income countries. For example, Buchmann, Sullivan, and Meyer (2024) show that men prohibit women from taking risky jobs against their will.

effects.[2] I use dictator IQ fixed effects in the analysis to compare allocations of dictators with the same IQ but assigned different relative IQ ranks due to random group formation to cut the correlation between IQ and baseline generosity.[3,4]

I first confirm that male and female dictators' behaviors are consistent with the literature: their allocation amounts are similar to those in a previous study with a comparable level of social distance from receivers. Male dictators allocate more to female receivers than to male receivers (albeit statistically insignificant), and female dictators allocate more than male dictators. However, I do not find evidence that male dictators are less generous to female receivers with higher IQs. The point estimate is quantitatively negligible, statistically indistinguishable from zero, and has a tight confidence interval. If anything, male dictators are slightly more generous to female receivers with higher IQs. These results are robust to the so-called "beauty premium" and hold across the entire distribution of male dictator allocations. The allocation patterns of female dictators are qualitatively similar to those of male dictators. Taken together, contrary to the hypothesis, men (and women) are no less generous to women, even when women excel in cognitive skills. Thus, the results suggest that women who excel in cognitive skills may not receive less favorable treatment than equally intelligent men in the labor market.[5]

This paper contributes to the literature about people's differential attitudes towards competent women and men. Social psychology and sociology literature find that people perceive and evaluate competent women more negatively than competent men (Heilman et al. 2004; Phelan, Moss-Racusin, and Rudman 2008; Rudman 1998; Rudman and Fairchild 2004; Rudman et al. 2012). Also, Quadlin (2018) finds female college students with a very high grade point average (GPA) receive fewer callbacks in hiring than male students with a similar GPA. However, there are studies that find the opposite: Ceci and Williams (2015) and Williams and Ceci (2015) find well-qualified female candidates for assistant professor positions receive equal or more favorable treatment than equally qualified male candidates in hiring. I show that although intelligence is an important factor of career success and one of the attributes men (and women) care most about (Castagnetti and Schmacker 2022; Eil and Rao 2011; Zimmermann 2020), excelling in intelligence is not a sufficient condition for women to receive less favorable treatments than equally intelligent men, consistent with the latter line of studies.

My paper also contributes to the literature on the role of gender in dictator games. The literature finds that female dictators allocate more than male dictators, but the difference is quantitatively modest at best (Bilén, Dreber, and Johannesson 2021). The difference is also context-dependent (Croson and Gneezy 2009; Doñate-Buendía, García-Gallego, and Petrović 2022) and possibly driven

---

2. For example, Babcock et al. (2017), Coffman (2014), and Isaksson (2018) use photos to inform experimental participants of the other participants' gender. Yet, I show the robustness of the results to the concern that facial photos may be subject to the so-called "beauty premium."

3. IQ fixed effects were used by Zimmermann (2020), among others.

4. For example, Falk et al. (2021) show that children from families with higher socioeconomic status tend to have higher IQ and are more altruistic than children from families with lower socioeconomic status.

5. A caveat is that the size of the standard error is not very small; thus, the results should be interpreted with caution.

by non-monetary motives (Klinowski 2018). Regarding the gender of the receivers, the literature finds that dictators allocate more to female receivers than to male receivers (Engel 2011), and male dictators may do so more (Dufwenberg and Muren 2006; List 2004). My paper adds to the latter evidence by introducing IQ as an additional dimension to gender and shows that male dictators' allocation patterns are largely unchanged when female receivers have higher IQs than male dictators.

The remainder of the paper proceeds as follows. Section 2 describes the experimental design, procedure, implementation, and data. Section 3 discusses the empirical strategy. Section 4 presents the results. Section 5 concludes.

2 Experiment

I conduct two experiments: the main and the follow-up. The main experiment collects data on dictator game allocation, and the follow-up experiment collects data on the receivers' attractiveness and other facial characteristics. I describe each of them in detail below.

2.1 Main Experiment

The main experiment consists of two parts, and participants receive instructions at the beginning of each part.

Pre-Experiment: Random Desk Assignment & Photo-Taking

Participants are randomly assigned to a partitioned computer desk. Afterward, they have their facial photos taken at a photo booth and enter their first names on their computers. The experimenters then go to each participant's desk to check that their photo and first name are correct, to assure other participants that the checked participants' photos and first names are real, following Isaksson (2018).

Part 1: IQ Test

In Part 1, participants work on nine incentivized IQ questions for nine minutes to measure their IQ. I use Bilker et al. (2012)'s "form A 9-item Raven test," which measures one's IQ 90% as accurately as the full-length Raven test but with fewer questions. Participants receive 0.5€ for each correctly solved IQ test question, but they do not receive information about how many questions they have solved correctly until the end of the experiment.

After the IQ test, participants make an incentivized guess on the number of IQ test questions they have solved correctly, which I use as a measure of their over-confidence level. They receive 0.5€ for a correct guess. They do not receive feedback on their guess accuracy until the end of the experiment.

Following Eil and Rao (2011), six participants are randomly grouped and privately informed about their IQ rank relative to the other five participants in the group. Ties are broken randomly.

They then answer a set of comprehension questions about their IQ rank. They cannot proceed to the next part until they answer the comprehension questions correctly.

Part 2a: Dictator Game (Dictators Only)

Figure 1: Dictator's allocation screen

Round 1 of 3

Giovanna

Rank 5

You have received **7€** for this round.

You have been paired with **Giovanna**.

Please allocate the endowment between yourself and Giovanna. When you click the line below, a cursor appears. You can move the cursor by dragging it. Please move the cursor to your preferred position to determine the allocation.

You ⸺⬭⸺ **Giovanna**

Next

Notes: This figure shows an example of a dictator's allocation screen. In this example, the dictator is playing the first round and paired with a receiver whose first name is Giovanna (a female name) and whose IQ rank is 5. In the experiment, dictators see Giovanna's facial photo instead of the silhouette.

In Part 2, three participants in each group are randomly assigned the role of dictators (which I call "active participant"), and the other three participants are randomly assigned the role of receivers (which I call "passive participant"). Dictators then play dictator games with windfall money three times, each time with a different receiver from their group, one by one, using a perfect stranger-matching protocol. During the dictator game, dictators observe the receivers' facial photos, first names, and IQ ranks; see Figure 1 for an example of a dictator's decision screen. The use of photos follows gender economics literature (Babcock et al. 2017; Coffman 2014; Isaksson 2018) to minimize the experimenter demand effects.

Dictators are told that their decisions are anonymous to the receivers and that their allocation will be paid as a "top-up" to the receivers' earnings. Dictators decide allocations by moving a cursor

on a slider where the cursor is initially hidden to prevent anchoring. I vary the endowment across rounds to make each dictator game less repetitive: 7€ for the 1st and the 3rd rounds, and 5€ for the 2nd round. At the end of the experiment, one of the three allocations is randomly chosen for each participant as earnings for Part 2.[6],[7]

Part 2b: Belief Elicitation (Receivers Only)

I also collect a proxy of dictators' beliefs on how many IQ test questions the matched receivers have solved correctly. To prevent the belief elicitation from affecting/being affected by the dictator game, I exploit the random role assignment, and use the receivers' beliefs. This is a valid proxy because both dictators and receivers are exactly in the same experimental environment up to the role assignment and the role assignment is random. Specifically, while dictators are playing the dictator game, receivers are matched with the other two receivers in the same group one by one with a perfect stranger-matching protocol, and they make incentivized guesses on how many IQ test questions they have solved correctly, observing the receivers' facial photo, first name, and IQ rank, just as dictators do. They receive 0.5€ for each correct guess.

To address the non-anonymity of showing facial photos and first names, I ask both dictators and receivers how well they know the paired participants on a scale of 4, from "Did not know at all" to "Knew very well."[8] I ask this question twice to make sure they do not answer randomly: right after the three dictator games (for dictators) or two guesses (for receivers) and in the post-experimental questionnaire.

Post-Experiment: Questionnaire

After the dictator game and the belief elicitation, participants are told their earnings from the IQ test, the dictator game, and the belief elicitation in private. Before receiving their earnings, participants answer a short questionnaire about their demographics. In addition, receivers are asked whether I could use their photos in another experiment with a gratuity of 1.5€ (only for receivers who agreed; 149 receivers out of 193, or 77% of receivers, agreed).

Implementation

The main experiment was programmed with oTree (Chen, Schonger, and Wickens 2016) and conducted in English in November-December 2019 at the Bologna Laboratory for Experiments in Social Science (BLESS). I recruited 390 students (195 female and 195 male) of the University

---

6. For each dictator and in each round, one of the three receivers in the same group is randomly chosen *without replacement* and the dictator allocates the endowment between themselves and the receiver. Thus, it is possible for two dictators to play the dictator game with the same receiver in the same round. At the end of the dictator games, each participant has three allocations, one of which is randomly chosen for payment.

7. Note that, as with other information provision experiments, the results I show later are intention-to-treat effects because some dictators may not have taken into account the receivers' IQ rank when deciding the allocation. Yet, dictators solved the comprehension questions about their IQ rank relative to other group members, as in Appendix Figure D1.

8. See the experimental instructions in Appendix E for the exact wording.

of Bologna via ORSEE (Greiner 2015) who (i) were born in Italy, (ii) had not participated in gender-related experiments in the past (as far as I could trace), and (iii) were available to participate in experiments in English. The first condition was to reduce the chance that receivers' photos and first names would signal ethnicity, race, or cultural background. The second condition was to minimize the experimenter demand effects. The third condition was to run the experiment in English. As a further attempt to make the data cleaner, I excluded receivers with non-Italian-sounding names and allocations in which the dictators answered at least once that they knew the paired receivers "very well."[9] These data screenings leave me with 388 participants, 195 dictators, and 558 dictator allocations.

The number of participants is based on the power simulation in the pre-analysis plan to achieve 80% power.[10] The experiment is pre-registered with the OSF.[11] Appendix A explains deviations from the pre-analysis plan.

I ran 24 sessions in total, and the number of participants in each session was a multiple of 6 (12 to 30). The average session length was 70 minutes, including registration and payment. The average payment per participant was about 10€ including the participation fee of 2.5€ and a gratuity of 1.5€ for photo use in another experiment (only for receivers who agreed).

2.2 Follow-Up Experiment

I conduct a follow-up experiment to collect a measure of the attractiveness and other facial characteristics of the receivers in the main experiment. It consists of one part, and participants earn a flat fee of 10€.

Participants are randomly assigned to a partitioned computer desk. Afterward, participants see photos of 100 receivers in the main experiment one by one and rate the attractiveness of each of the photos on a 5-point Likert scale. I also ask them to rate the photos in terms of how kind they look (on a 5-point Likert scale), to which extent they look Italian (on a 3-point Likert scale), and whether they know the person (yes or no) – all in one screen for each photo. The photos are randomly drawn from the pool of all receivers who agreed to show their photos in another experiment, and the order of the photos is randomized. After rating all the 100 photos, participants provide their gender and age.

The follow-up experiment was programmed with oTree and conducted in Italian in October 2023 at the Bologna Laboratory for Experiments in Social Science (BLESS). I recruited 28 students (14 female and 14 male) who (i) were born in Italy and (ii) were available to participate in experiments in English to make the subject pool as close as possible to that of the main experiment. I also

---

9. Although it is easy to distinguish Italian and non-Italian-sounding names, to make sure not to misclassify I asked the laboratory manager, who was a native Italian, to check the participants' first names after each session.

10. I exclude the first-session data because of the problem discussed in Appendix A.

11. The pre-analysis plan and the R code for power calculation are available at the OSF registry: https://osf.io/r6d 8f/files.

restricted participants to those who had not participated in the main experiment.[12,13,14]

2.3 Data

Appendix B describes the data in detail; I briefly summarize its key aspects here. First, most male dictators (95%) do not know the receiver at all, so it is unlikely that the relationship outside the laboratory affects dictator allocation. Second, consistent with the literature, male dictators allocate less (6.3 percentage points less, the p-value < 0.05 with a two-sided t-test) to receivers than female dictators do. Third, again somewhat consistent with the literature, dictators, especially male dictators, allocate slightly more (2.4 percentage points more; male dictators allocate 3.1 percentage points more) to female receivers than to male receivers, although the difference is statistically insignificant. Fourth, both men and women consider women to look more attractive than men, regardless of their IQ level.

3 Empirical Strategy

A naive way to test whether male dictators allocate less to higher-IQ female receivers than to higher-IQ male receivers is to compare male dictators' allocations to those receivers. However, this simple difference may be biased upward because male dictators may allocate more to more attractive receivers (Rosenblat 2008), and men rate women more highly than men on attractiveness (see the discussion in Section 2.3). In fact, male dictators allocate more to more attractive receivers (see Appendix Figure D2, Panels A and B).[15]

A difference-in-differences would eliminate the bias from the attractiveness differences, using the difference between male dictators' allocations to lower-IQ female receivers and to lower-IQ male receivers as a comparison group. However, it introduces two other biases because low-IQ dictators are more likely to face higher-IQ receivers. First, because low-IQ dictators earned less on the IQ test than high-IQ dictators, higher-IQ receivers and dictators' earnings are negatively correlated, potentially biasing the dictators' allocation to higher-IQ receivers due to the wealth effects. Also, if low-IQ dictators think the differences between their earnings and the higher-IQ receivers' earnings are larger than high-IQ dictators think, it also induces inequality aversion (Fehr and Schmidt 1999). Second, because one's IQ is positively associated with one's socioeconomic background, and one's socioeconomic background influences one's social preferences (Falk et al. 2021), the dictators' allocation to higher-IQ receivers may be biased upward or downward. In my sample, male dictators' IQs (or their earnings on the IQ test) and allocations are negatively correlated, and male dictators'

---

12. Participants know the people in the photos in 10 out of 2800, or 0.4% of total ratings.

13. Although I recruited participants born in Italy, 5 people in the photos were considered completely non-Italian (4 by male raters and 4 by female raters with overlaps of 3 receivers).

14. I use ratings by male raters for male dictators and by female raters for female dictators. For female and male raters, the median ratings per receiver is 9, the mean is 8.66, the minimum is 4, the maximum is 14, and the standard deviation is 1.84.

15. Interestingly, female dictators do not allocate more to more attractive receivers (see Appendix Figure D2, Panels C and D).

IQ ranks and allocations are positively correlated (see Appendix Figure D2, Panels A and B), both of which bias the allocations to higher-IQ receivers upwards.[16]

To address these concerns, I estimate the following difference-in-differences equation via OLS using male dictators' allocation data:

$$
\begin{aligned}
Allocate_{ij} =& \beta_1 HigherIQReceiver_{ij} \times FemaleReceiver_j \\
& + \beta_2 HigherIQReceiver_{ij} + \beta_3 FemaleReceiver_j + X'_{ij}\gamma + \mu_i^{IQ} + \epsilon_{ij}
\end{aligned}
\tag{1}
$$

where each variable is defined as follows:

- $Allocate_{ij} \in [0,1]$: dictator i's allocation to receiver j as a fraction of endowment.
- $HigherIQReceiver_{ij} \in \{0,1\}$: an indicator variable equals 1 if receiver j's IQ is higher than that of dictator i.
- $FemaleReceiver_j \in \{0,1\}$: an indicator variable equals 1 if receiver j is female.
- X_{ij}: a set of additional covariates to increase statistical power. Appendix C provides a full description of the covariates.
- ϵ_{ij}: the error term.

and $\mu_i^{IQ} \equiv \sum_{k=1}^{9} \theta_k^{IQ} \mathbb{1}[\text{i's IQ} = k]$ is fixed effects for the dictators' IQ (the number of IQ test questions they have solved correctly), where $\mathbb{1}$ is the indicator function. Standard errors are clustered at the dictator level with Pustejovsky and Tipton (2018)'s small cluster bias adjustment. The IQ fixed effects are included following Zimmermann (2020), so that the coefficients in equation 1 capture allocation differences due to the receivers' IQ, not that of the dictators. As discussed above, without the IQ fixed effects, the coefficients related to higher-IQ receiver can capture the wealth effects, inequality aversion, and the difference in the preference for allocation between high IQ and low IQ dictators.

The coefficient of interest is β_1, which captures the difference between the dictator allocations to higher-IQ female receivers and higher-IQ male receivers relative to the difference between the dictator allocations to lower-IQ female receivers and lower-IQ male receivers. As discussed above, because men rate women more highly than men on attractiveness, the estimates related to the female receiver can capture the effects of attractiveness differences without taking the double differences. On the other hand, as discussed in Section 2.3, men's ratings on high-IQ women's attractiveness and low-IQ women's attractiveness are very similar. Also, men's rating on high-IQ men's attractiveness and low-IQ men's attractiveness is very similar, justifying the difference-in-differences.

4   Results

4.1   Regression Results

Table 1 presents the regression results of equation 1 with data for male dictators. Column 1 shows that the coefficient estimate on higher-IQ receivers (β_2) is 0.093 and statistically marginally

---

16. We do not observe this correlation patterns for female dictators (see Appendix Figure D2, Panels C and D).

Table 1: Dictator allocations to higher-IQ female receivers – OLS, male dictators

Outcome:	Dictator's allocation (fraction of endowment)			
Sample:	Male dictators			
	(1)	(2)	(3)	(4)
Higher IQ receiver x Female receiver (β_1)	0.018	0.017	0.031	0.047
	(0.060)	(0.060)	(0.061)	(0.075)
	[-0.101, 0.136]	[-0.101, 0.134]	[-0.089, 0.151]	[-0.101, 0.195]
Higher IQ receiver (β_2)	0.093*	0.054	0.048	0.007
	(0.048)	(0.053)	(0.055)	(0.059)
	[-0.001, 0.188]	[-0.050, 0.159]	[-0.060, 0.156]	[-0.109, 0.124]
Female receiver (β_3)	0.038	0.031	0.014	-0.021
	(0.035)	(0.035)	(0.034)	(0.044)
	[-0.031, 0.107]	[-0.038, 0.100]	[-0.053, 0.081]	[-0.108, 0.066]
Dictator IQ FE	-	✓	✓	✓
Round FE	-	-	✓	✓
Social distance FE	-	-	✓	✓
Dictator demographics	-	-	✓	✓
Receiver demographics	-	-	✓	✓
Receiver attractiveness FE	-	-	-	✓
Higher IQ receiver x Female receiver +Higher IQ receiver ($\beta_1 + \beta_2$)	0.111**	0.071	0.079	0.054
	(0.055)	(0.048)	(0.052)	(0.070)
	[0.004, 0.219]	[-0.024, 0.166]	[-0.023, 0.181]	[-0.084, 0.193]
Higher IQ receiver x Female receiver +Female receiver ($\beta_1 + \beta_3$)	0.056	0.048	0.045	0.026
	(0.049)	(0.046)	(0.047)	(0.061)
	[-0.040, 0.151]	[-0.043, 0.139]	[-0.048, 0.138]	[-0.094, 0.146]
Baseline Mean	0.305	0.305	0.305	0.327
Baseline SD	0.269	0.269	0.269	0.270
Adj. R-squared	0.032	0.052	0.080	0.088
Observations	260	260	260	211
Clusters	91	91	91	91

Notes: This table presents the regression results of equation 1. The standard error (in parenthesis) and the 95% confidence interval (in bracket) are reported below each coefficient estimate. The standard errors are clustered at the dictator level with Pustejovsky and Tipton (2018)'s small cluster bias adjustment. Baseline mean and standard deviation are that of lower-IQ male receivers. Significance levels: * 10%, ** 5%, and *** 1%.

significant at 10%, suggesting that male dictators allocate 9.3 percentage points more to higher-IQ male receivers than to lower-IQ male receivers. Also, the sum of the coefficient estimates on the interaction term between higher-IQ receivers and the female receivers, and on the higher-IQ receivers $(\beta_1 + \beta_2)$ is 0.111 and statistically significant at 5%, suggesting that male dictators allocate 11.1 percentage points more to higher-IQ female receivers than to lower-IQ female receivers. However, as discussed in Section 3, the coefficient estimates related to higher-IQ receivers are biased because low-IQ dictators are more likely to face higher-IQ receivers. Hence, we move to Column 2, which includes dictator IQ fixed effects. Comparing the coefficient estimates of Columns 1 and 2, the only difference is the coefficient estimate on higher-IQ receivers (β_2), which is in line with the discussion in Section 3. Indeed, the dictator IQ fixed effects in Column 2 are jointly statistically significant with a p-value of 0.07.

Column 2 shows that the coefficient estimate on female receivers (β_3) is 0.031, suggesting that male dictators may allocate slightly more to lower-IQ female receivers than to lower-IQ male receivers, albeit the difference is statistically insignificant. Also, the coefficient estimate on higher-IQ receivers (β_2) is 0.054, suggesting that male dictators may allocate more higher-IQ male receivers than to lower-IQ male receivers, albeit the difference is statistically insignificant. The sum of the coefficient estimates on the interaction term between higher-IQ receivers and female receivers, and on higher-IQ receivers $(\beta_1 + \beta_2)$ is 0.071, suggesting that male dictators may allocate slightly more to higher-IQ female receivers than to lower-IQ female receivers, albeit the difference is statistically insignificant. The sum of the coefficient estimates on the interaction term between higher-IQ receivers and female receivers, and on female receivers $(\beta_1 + \beta_3)$ is 0.048, albeit statistically insignificant, suggesting that male dictators may allocate more to higher-IQ female receivers than to higher-IQ male receivers.

However, as discussed in Section 3, the comparison between allocations to higher-IQ female receivers and allocations to lower-IQ female receivers is confounded because higher-IQ receivers earn more in the IQ test than lower-IQ receivers, which may induce dictators' inequality aversion. Also, the comparison between allocations to higher-IQ female receivers and allocations to higher-IQ male receivers is confounded because men consider women to be more attractive than men and thus may allocate more to female receivers than to male receivers.

To address these concerns, we turn to our coefficient of interest: The coefficient estimate on the interaction between higher-IQ receivers and female receivers (β_1), which is 0.017 and statistically insignificant, suggesting that male dictators' allocation differences towards higher-IQ female and male receivers and towards lower-IQ female and male receivers are statistically indistinguishable. Column 3, which controls for dictator and receiver demographics, as well as for round and dictator-receiver social distance to increase statistical power, presents quantitatively similar empirical patterns to Column 2.[17,18,19] Thus, the results are inconsistent with the hypothesis that men are less generous

17. Appendix Table D1 presents results where I gradually add controls and show that the main results are not driven by specific controls.

18. Appendix Table D2 presents results where the standard errors are neither cluster-robust nor heteroskedasticity-robust to see how the standard errors change by clustering and making it robust to heteroskedasticity.

19. Appendix Table D3 presents the same results as Table 1 but with female dictators. While the coefficient estimates on higher-IQ receivers and on female receivers are slightly negative albeit statistically insignificant, the coefficient

to higher-IQ women.

A caveat is that the confidence interval is not very tight. Looking at the 95% confidence intervals of the interaction between higher-IQ receivers and female receivers (β_1), our coefficient of interest, we can reject at the 5% significance level that the effect size is no smaller than -8.9 to -10.1 percentage points and no larger than 13.4 to 15.1 percentage points. As references, Engel (2011) finds via meta-analysis that male dictators allocate 14.3 percentage points more to female receivers than to male receivers, and Dufwenberg and Muren (2006) find using a university-student sample that male dictators allocate 7.6 to 8.9 percentage points more to female receivers than to male receivers. To halve the confidence intervals, we need approximately 1560 participants (=390*4).[20]

4.2 Distribution Results

Figure 2: Dictator allocations to higher-IQ female receivers – Distribution, male dictators
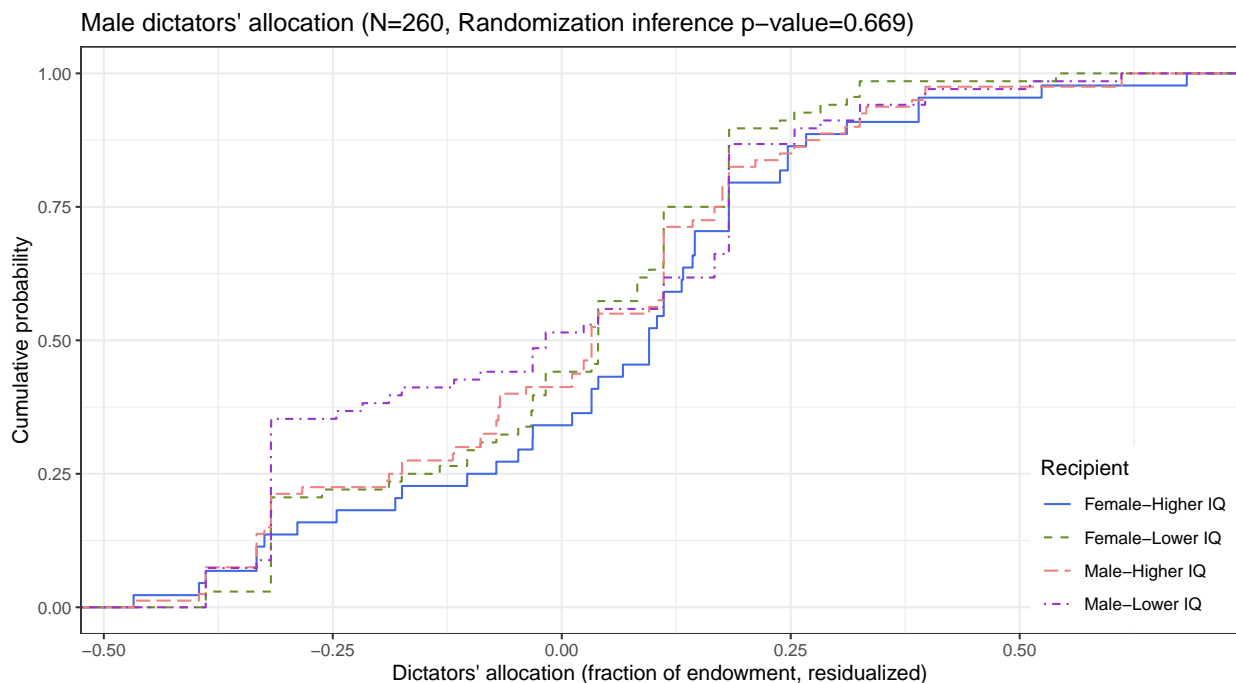
Notes: The figure presents the empirical CDF of dictator allocations by receiver types, residualized with the dictator-IQ fixed effects, to give a causal interpretation to the differences. The randomization inference p-value (Young 2019) is calculated with the Kruskal-Wallis test with 2000 random draws. I use randomization inference to address arbitrary dependency among allocations. The null hypothesis is that all CDFs coincide.

Although we did not find support for the hypothesis in mean, we may find support in distribution. To see this, Figure 2 presents the empirical CDF of dictator allocations by receiver types, residualized with the dictator-IQ fixed effects to give a causal interpretation to the differences,[21] and shows that

---

estimate of the interaction between higher-IQ receivers and female receivers remains positive, albeit statistically insignificant.

20. The number is based on the OLS standard error formula.

21. Residualized allocation is residual from a regression of dictator allocations on dictator-IQ fixed effects.

the results in Section 4.1 also hold in distribution. It shows that across almost the whole dictators' allocations, male dictators allocate slightly more to higher-IQ female receivers than to lower-IQ male receivers, higher-IQ male receivers, and lower-IQ female receivers, albeit the differences are statistically insignificant (the randomization inference p-value is 0.669, which is robust to arbitrary dependency among allocations, calculated with the Kruskal-Wallis test with 2,000 random draws).[22]

Thus, the results are inconsistent with the hypothesis that men are less generous to higher-IQ women even in distribution; rather, if anything, men are slightly more generous to higher-IQ women than to higher-IQ men.

4.3 Robustness Check

"Beauty Premium"

In the main results, I control for the possibility that dictators allocate more to more attractive receivers – the so-called "beauty premium" (Rosenblat 2008) – by taking double differences. Yet, I further address this concern in Column 4 of Table 1. In addition to all the controls in Column 3, Column 4 controls for the receivers' attractiveness via fixed effects (adding dummies for each level of attractiveness ratings, which are jointly significant with a p-value of 0.001), and shows that the magnitude of the coefficient of interest, the interaction term between higher-IQ receivers and female receivers (β_1), remains essentially the same as in Column 3, albeit the standard errors are larger due to a drop in the sample size by 23%.[23]

Inequality Aversion due to Differential Earnings in the IQ Test

I control for dictators' inequality aversion (Fehr and Schmidt 1999) due to the differential earnings in the IQ test between higher-IQ receivers and dictators via IQ fixed effects. Yet, it assumes that male dictators believe higher-IQ female and male receivers solved the same number of IQ test questions. Thus, I examine the validity of this assumption in Column 5 of Appendix Table D1. Instead of allocation data for male dictators, Column 5 uses data for male receivers' beliefs on the IQ levels of the other receivers.[24] It shows that male receivers believe that higher-IQ female receivers solve about 0.82 more IQ test questions (=1.109-0.291) or earn about 0.41€ more (about 6.5% of the dictator endowment) than higher-IQ male receivers. Thus, if anything, the main results are underestimated in the absence of inequality aversion.

IQ Rank Differences

Dictators see the difference between the receivers' and their own IQ rank. Yet, in the main results, I only consider whether the receivers' IQ rank is higher than the dictators' IQ rank. To investigate

---

22. Appendix Figure D3 presents the same results as Figure 2, but with female dictators and shows essentially the same empirical patterns.

23. 44 out of 193 receivers, or 23% of all receivers, refused to show their photos in the follow-up experiment.

24. See Section 2 for the justification that it is a valid proxy of the male dictators' beliefs about the receivers' IQ level.

the heterogeneity by the difference in the IQ rank, I estimate the following equation:

$$Allocate_{ij} = \sum_{k=-5,\neq 1}^{5} \beta_1^k \mathbb{1}[ReceiverIQRank_j - DictatorIQRank_i = k] \times FemaleReceiver_j$$

$$+ \sum_{k=-5,\neq 1}^{5} \beta_2^k \mathbb{1}[ReceiverIQRank_j - DictatorIQRank_i = k] + \beta_3 FemaleReceiver_j$$

$$+ X_{ij}'\gamma + \mu_i^{IQ} + \epsilon_{ij}$$

$$(2)$$

where $ReceiverIQRank_j$ is receiver j's IQ rank, $DictatorIQRank_i$ is dictator i's IQ rank, and other variables are as defined in equation 1. In the main specification, $HigherIQReceiver_{ij}$ is equal to $\sum_{k=-5}^{-1} \mathbb{1}[ReceiverIQRank_j - DictatorIQRank_i = k]$ and the omitted category is $\sum_{k=2}^{5} \mathbb{1}[ReceiverIQRank_j - DictatorIQRank_i = k]$.

Figure 3: Dictator allocations to higher-IQ female receivers – IQ rank differences, male dictators
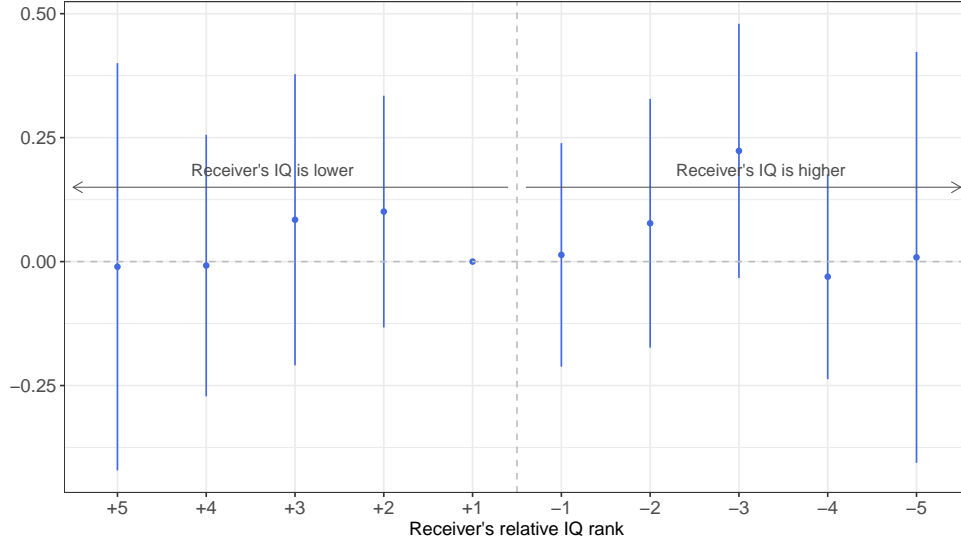
Notes: This figure plots the OLS estimates of β_1s in equation 2, along with the 95% confidence intervals. The omitted category is the receiver's relative IQ rank is 1. The standard errors are clustered at the dictator level with Pustejovsky and Tipton (2018)'s small cluster bias adjustment.

Figure 3 plots the OLS estimates of β_1s of equation 2, along with the 95% confidence intervals. We do not see any statistically or quantitatively significant heterogeneity due to the IQ rank differences – all $\hat{\beta}_1$s are close to each other. While the estimate of β_1^{-3} is larger than the other estimates, it is not statistically significantly different from the omitted category, and there is no consistent pattern when the receiver's IQ is higher than that of the dictator's as the estimates of β_1^{-4} and β_1^{-5} are very small.
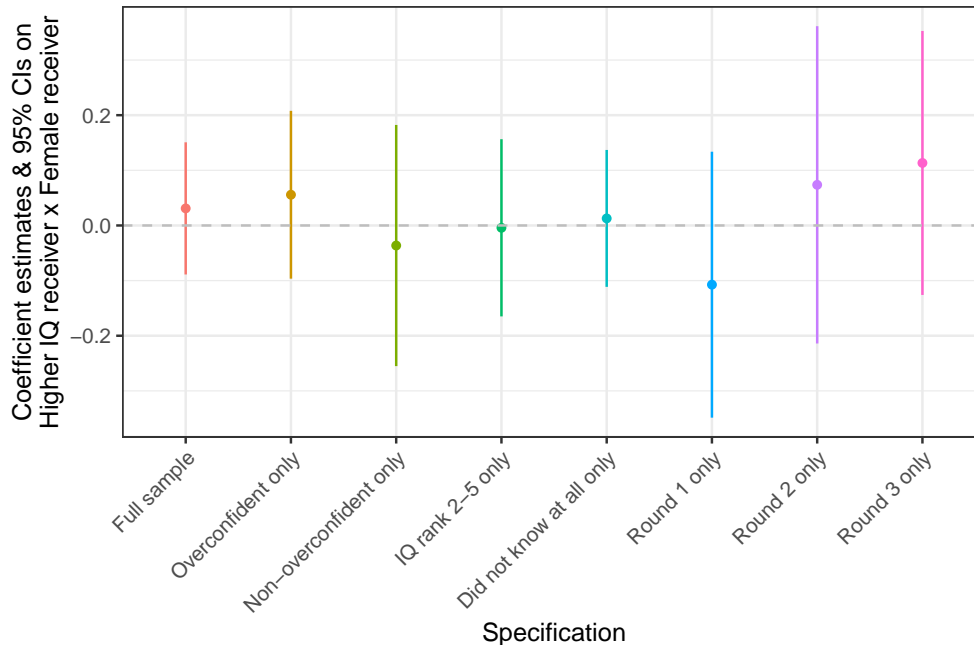
Figure 4: Dictator allocations to higher-IQ female receivers – Sub-sample analysis, male dictators

Notes: This figure presents the OLS estimates of β_1 and their 95% confidence intervals of equation 1 with the same controls in Column 3 of Table 1 but with sub-samples of male dictators. "Full sample" is the same estimate as in Column 3 of Table 1, provided as a reference. The standard errors are clustered at the dictator level with Pustejovsky and Tipton (2018)'s small cluster bias adjustment for specifications from "Full sample" to "Did not know at all only" and heteroskedasticity-robust with Bell and McCaffrey (2002)'s small sample bias adjustment for specifications "Round 1 only," "Round 2 only," and "Round 3 only."

Other Concerns

Figure 4 presents the OLS estimates of β_1 and their 95% confidence intervals of equation 1 with the same controls in Column 3 of Table 1 but with sub-samples of male dictators. "Full sample" is the same estimate as in Column 3 of Table 1, provided as a reference. As we see, the point estimates are quantitatively very close to those of the full sample, although the confidence intervals are wider due to a drop in the sample size. One noticeable heterogeneity is round effects, where male dictators allocate less in round 1 and more in rounds 2 and 3. This heterogeneity could possibly be due to the experimenter demand effect, as dictators may have realized that the experiment was about gender and IQ.

5   Conclusions

The literature finds that men are more generous to women than to men. Still, there is scarce evidence about how men's higher generosity towards women interacts with skills important for one's career success, such as cognitive skills. The literature suggests that men's higher generosity towards women may be due to paternalism and thus be reversed if women excel in those skills. To fill the

gap in the literature, this paper uses a dictator game to examine whether men are less generous to women when those women excel in cognitive skills. Although male dictators allocate more to attractive receivers, consistent with the so-called "beauty premium," and male dictators with high IQs (or those who earn more on the IQ test) allocate less, I exogenously vary the receivers' IQ relative to that of the dictators and use difference-in-differences to control for these associations.

The results, however, are inconsistent with the hypothesis; if anything, male dictators give more to higher-IQ female receivers than to higher-IQ male receivers relative to the difference in giving between lower-IQ female and male receivers. The results hold both in mean and distribution, and female dictators' allocation patterns are qualitatively similar to male dictators. The caveat is that the confidence intervals are not very tightly estimated. Taken together, men are no less generous to women even when women excel in cognitive skills, and thus, women who excel in cognitive skills may not receive less favorable treatment than equally intelligent men in the labor market.

One possible explanation for the results would be that the gender norm violations are offset by something else; for example, it is possible that women are more likely to appreciate other people's work, women with high IQs do so more effectively, and men (and women) reciprocate indirectly. Indeed, Folke and Rickne (2023) show that workplaces with a higher female-worker ratio have a higher incidence of appreciation for others' work. Another potential explanation is that gendered attitudes develop as people get older, and university students – my sample – have not formed gendered attitudes yet. If this is the case, intelligent women may still receive less favorable treatment than equally intelligent men in the labor market. These two explanations are not mutually exclusive, but testing these explanations is beyond the scope of this paper. Nevertheless, my paper showed that, at least among the university student population in Italy, intelligent women would not receive less favorable treatment than equally intelligent men.

References

Babcock, Linda, María P. Recalde, Lise Vesterlund, and Laurie Weingart. 2017. "Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability." *American Economic Review* 107 (3): 714–747.

Bell, Robert M., and Daniel F. McCaffrey. 2002. "Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples." *Survey Methodology* 28 (2): 169–181.

Bilén, David, Anna Dreber, and Magnus Johannesson. 2021. "Are Women More Generous than Men? A Meta-Analysis." *Journal of the Economic Science Association* 7 (1): 1–18.

Bilker, Warren B., John A. Hansen, Colleen M. Brensinger, Jan Richard, Raquel E. Gur, and Ruben C. Gur. 2012. "Development of Abbreviated Nine-Item Forms of the Raven's Standard Progressive Matrices Test." *Assessment* 19 (3): 354–369.

Buchmann, Nina, Colin D. Sullivan, and Carl Meyer. 2024. *Paternalistic Discrimination.* Working Paper.

Cappelen, Alexander W., Ranveig Falch, and Bertil Tungodden. 2023. *Experimental Evidence on the Acceptance of Males Falling Behind.* Working Paper.

Castagnetti, Alessandro, and Renke Schmacker. 2022. "Protecting the Ego: Motivated Information Selection and Updating." *European Economic Review* 142:104007.

Ceci, Stephen J., and Wendy M. Williams. 2015. "Women Have Substantial Advantage in STEM Faculty Hiring, except When Competing against More-Accomplished Men." *Frontiers in Psychology* 6:1532.

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree–An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance* 9:88–97.

Coffman, Katherine Baldiga. 2014. "Evidence on Self-Stereotyping and the Contribution of Ideas." *The Quarterly Journal of Economics* 129 (4): 1625–1660.

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

Doñate-Buendía, Anabel, Aurora García-Gallego, and Marko Petrović. 2022. "Gender and Other Moderators of Giving in the Dictator Game: A Meta-Analysis." *Journal of Economic Behavior & Organization* 198:280–301.

Dufwenberg, Martin, and Astri Muren. 2006. "Generosity, Anonymity, Gender." *Journal of Economic Behavior & Organization* 61 (1): 42–49.

Eagly, Alice H., and Steven J. Karau. 2002. "Role Congruity Theory of Prejudice toward Female Leaders." *Psychological Review* 109 (3): 573–598.

Eil, David, and Justin M. Rao. 2011. "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself." *American Economic Journal: Microeconomics* 3 (2): 114–138.

Engel, Christoph. 2011. "Dictator Games: A Meta Study." *Experimental Economics* 14 (4): 583–610.

Erkal, Nisvan, Lata Gangadharan, and Boon Han Koh. 2023. "Do Women Receive Less Blame than Men? Attribution of Outcomes in a Prosocial Setting." *Journal of Economic Behavior & Organization* 210:441–452.

Falk, Armin, Fabian Kosse, Pia Pinger, Hannah Schildberg-Hörisch, and Thomas Deckers. 2021. "Socioeconomic Status and Inequalities in Children's IQ and Economic Preferences." *Journal of Political Economy* 129 (9): 2504–2545.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *The Quarterly Journal of Economics* 114 (3): 817–868.

Fitriani, Randolf G. S. Cooper, and Ron Matthews. 2016. "Women in Ground Close Combat." *The RUSI Journal* 161 (1): 14–24.

Folke, Olle, and Johanna Rickne. 2023. *Workplace Sex Composition and Appreciation at Work.* Working Paper.

Franzen, Axel, and Sonja Pointner. 2013. "The External Validity of Giving in the Dictator Game." *Experimental Economics* 16 (2): 155–169.

Glick, Peter, Susan T. Fiske, Antonio Mladinic, José L. Saiz, Dominic Abrams, Barbara Masser, Bolanle Adetoun, et al. 2000. "Beyond Prejudice as Simple Antipathy: Hostile and Benevolent Sexism across Cultures." *Journal of Personality and Social Psychology* 79 (5): 763–775.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association* 1 (1): 114–125.

Heilman, Madeline E. 2001. "Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder." *Journal of Social Issues* 57 (4): 657–674.

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." *The Journal of Applied Psychology* 89 (3): 416–427.

Herrnstein, Richard J., and Charles Murray. 1996. *The Bell Curve: Intelligence and Class Structure in American Life.* New York, NY: Free Press.

Isaksson, Siri. 2018. *It Takes Two: Gender Differences in Group Work.* Working Paper.

Klinowski, David. 2018. "Gender Differences in Giving in the Dictator Game: The Role of Reluctant Altruism." *Journal of the Economic Science Association* 4 (2): 110–122.

List, John A. 2004. "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field." *The Quarterly Journal of Economics* 119 (1): 49–89.

Phelan, Julie E., Corinne A. Moss-Racusin, and Laurie A. Rudman. 2008. "Competent yet Out in the Cold: Shifting Criteria for Hiring Reflect Backlash Toward Agentic Women." *Psychology of Women Quarterly* 32 (4): 406–413.

Pustejovsky, James E., and Elizabeth Tipton. 2018. "Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models." *Journal of Business & Economic Statistics* 36 (4): 672–683.

Quadlin, Natasha. 2018. "The Mark of a Woman's Record: Gender and Academic Performance in Hiring." *American Sociological Review* 83 (2): 331–360.

Ridgeway, Cecilia L. 2001. "Gender, Status, and Leadership." *Journal of Social Issues* 57 (4): 637–655.

Rosenblat, Tanya S. 2008. "The Beauty Premium: Physical Attractiveness and Gender in Dictator Games." *Negotiation Journal* 24 (4): 465–481.

Rudman, Laurie A. 1998. "Self-Promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management." *Journal of Personality and Social Psychology* 74 (3): 629–645.

Rudman, Laurie A., and Kimberly Fairchild. 2004. "Reactions to Counterstereotypic Behavior: The Role of Backlash in Cultural Stereotype Maintenance." *Journal of Personality and Social Psychology* 87 (2): 157–176.

Rudman, Laurie A., Corinne A. Moss-Racusin, Julie E. Phelan, and Sanne Nauts. 2012. "Status Incongruity and Backlash Effects: Defending the Gender Hierarchy Motivates Prejudice against Female Leaders." *Journal of Experimental Social Psychology* 48 (1): 165–179.

Rudman, Laurie A., and Julie E. Phelan. 2008. "Backlash Effects for Disconfirming Gender Stereotypes in Organizations." *Research in Organizational Behavior* 28:61–79.

Williams, Wendy M., and Stephen J. Ceci. 2015. "National Hiring Experiments Reveal 2:1 Faculty Preference for Women on STEM Tenure Track." *Proceedings of the National Academy of Sciences* 112 (17): 5360–5365.

Young, Alwyn. 2019. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *The Quarterly Journal of Economics* 134 (2): 557–598.

Zimmermann, Florian. 2020. "The Dynamics of Motivated Beliefs." *American Economic Review* 110 (2): 337–361.