Essays on Group Identity and Economic Behavior

Daniel Carvajal
NHH Norwegian School of Economics

Acknowledgements

I am sincerely thankful for all the advice, support, and encouragement received over the years. In particular, I am extremely grateful to my supervisors and my amazing Costa Rican and German families. Without them, none of this would have been possible. I feel fortunate that in my academic journey I have been gifted with the most special friendships during my time in Madrid, Bergen, and Boston. Friendships that I am confident will stand, no matter the distance. Additionally, I am grateful to my friends in Costa Rica who have always been there, regardless of the distance or time difference. Lastly, I want to thank all my colleagues – fellow PhDs, administration, and faculty – for their invaluable support and the shared moments.

Best regards,

Daniel Carvajal Zúñiga
Bergen, February 2024

Contents

Introduction

Identity, defined as a person's sense of self derived from perceived membership in social groups, has been highlighted as a primary motivation in economic behavior. Research has demonstrated that an individual's perception of group membership significantly influences a range of decisions, such as food consumption (Atkin, Colson-Sihra and Shayo, 2021), human capital investments (Fryer Jr and Levitt, 2004), labor market choices (Oh, 2023), and social behaviors, including the extent of concern for others (Chen and Li, 2009). These behavioral patterns have substantial implications, particularly in the political landscape. Identity has been at the core of the discussion of policies regarding redistribution (Klor and Shayo, 2010), trade (Grossman and Helpman, 2021), migration (Fouka, Mazumder and Tabellini, 2022), and many others. Understanding the mechanisms behind identity concerns helps us not only to explain the observed phenomena, but also to explore whether these concerns contribute to adverse consequences like discrimination, polarization, and intergroup conflict, and how to address them.

Social identity theory was first formalized by social psychologists Tajfel and Turner (1979), presenting social categorization and group belonging as an explanation for intergroup conflict. Building upon this, Akerlof and Kranton (2000) integrated identity motivations into economic behavior, positioning them alongside pecuniary incentives as a primary determinant of decision-making. This theoretical framework and subsequent developments have underlined two main types of behavioral outcomes stemming from identity concerns. First, ingroup bias, which is the extent to which an individual prioritizes the welfare of individuals with whom they share the same group membership – ingroup – over others with whom they do not share group membership – outgroup (Chen and Li, 2009). Second, conformity in behavior, where an individual derives utility from adopting behaviors that are consistent with the norms and stereotypes associated with that particular group identity (Benjamin, Choi and Strickland, 2010). The presence of both behaviors derives many implications for education, labor supply, work effort, consumption, etc., and has been empirically validated through extensive research over recent decades (Shayo, 2020; Charness and Chen, 2020; Li, 2020).

However, several questions are yet to be understood. For instance, when do identity concerns arise? How do our social environments influence our group identification? Are all individuals equally prone to these concerns? Moreover, in what scenarios do identity considerations affect behavior, and are they harmful?

This dissertation consists of three chapters addressing these questions. The first two chapters explore the emergence of identity concerns, focusing on how social environments affect ingroup bias in social preferences. The third chapter identifies a new setting, technology adoption, in which gender-related identity concerns potentially have adverse effects of increasing inequality in education.

This research provides causal and correlational evidence in the study of these questions, through the use of experimental methods, incentivized measures, and large-scale surveys. Collectively, the dissertation offers valuable insights which inform empirical research, models, and theories in the economics of identity.

The first chapter, titled **Exposure to diversity, social proximity, and ingroup bias**, investigates how varying levels of diversity in one's social context, characterized by the individuals to whom we are exposed, impact preferences for giving. Using a large-scale sample from the U.S., the study examines incentivized allocation decisions towards either fellow U.S. nationals or foreigners, while being exogenously exposed to social contexts with varying levels of diversity in nationalities. The findings reveal that facing a diverse context amplifies ingroup bias, the tendency to favor one's own group, driven by both increased allocations towards U.S. nationals and decreased allocations to foreigners relative to allocations in a homogeneous context. Evidence suggests that changes in perceptions of social proximity are a mechanism behind the effects of context on allocations. Finally, exposure to a diverse context influences political views, in a direction consistent with an increased concern towards U.S. nationals.

These findings contribute to our understanding of how identity concerns arise and their intensity, highlighting the role of social context in shaping ingroup favoritism in giving behavior. Furthermore, this research can inform policy debates in societies experiencing demographic shifts, which involve constant changes in the social environments of individuals.

The second chapter, titled **The role of a majority-minority status and ingroup affinity in shaping social preferences**, extends this research agenda by investigating how an individual's majority-minority status in an immediate social context influences their social preferences. The study analyzes decisions in three incentivized allocation games, using a controlled experiment in a sample of around 1600 participants from the US. Participants are randomly assigned to roles as members of a majority or a minority. The findings reveal that majority-minority status does not significantly affect these social preferences. Moreover, ingroup affinity, defined as whether an individual feels closer to the ingroup relative to the outgroup, arises as a crucial factor in the emergence of ingroup biases in social preferences.

The paper extends our knowledge on the impact of changes in social contexts, driven by demographic shifts, on social preferences. It also identifies which individuals are subject to group identity concerns. In particular, I find that ingroup affinity is predictive of the presence of ingroup bias in allocation.

The final chapter, titled **Will Artificial Intelligence get in the way of achieving gender equality?**, is coauthored with Catalina Franco and Siri Isaksson, and examines differences in AI technology adoption by gender. We conducted a survey at the Norwegian School of Economics collecting use and attitudes towards ChatGPT, a measure of AI proficiency, and responses to policies allowing or forbidding ChatGPT use. Three key findings emerge: first, female students report a significantly lower use of ChatGPT compared to their male counterparts. Second, male students are more skilled at writing successful prompts, which is not explained by their higher ChatGPT usage. Third, imposing university bans on the use of ChatGPT widens the gender gap in intended use substantially. It also provides important insights into potential factors influencing the AI adoption gender gap. The gap in AI use and proficiency is closed when controlling for background characteristics such as confidence in AI use.

The study has increased relevance with the rise of productivity-enhancing AI technology to help solve a wide range of tasks. The differential usage of AI tools between women and men could potentially result in productivity and pay gaps. Therefore, early AI adoption is likely to become crucial in a rapidly-evolving labor market demanding these skills. This paper contributes by identifying potential factors influencing this gap in early adoption, in which identity concerns, stemming from conformity to group norms and stereotypes could be addressed to reduce this gap.

Chapter 1

Exposure to diversity, social proximity and ingroup bias

Abstract: As society becomes increasingly diverse, a key question arises: does a change in our social context – defined by the individuals we are exposed to – influence our interactions with each other? This paper studies this question using an experiment in a large-scale U.S. sample. Participants make an incentivized allocation towards either a fellow U.S. national or a foreigner, when being exogenously exposed to social contexts with varying levels of diversity in nationalities. I find that facing a *diverse* context amplifies ingroup bias, the tendency to favor one's own group, driven by both increased allocations towards U.S. nationals and decreased allocations to foreigners relative to allocations in a *homogeneous* context. Evidence suggests that changes in perceptions of social proximity are a mechanism behind the effects of context on allocations. Finally, exposure to a diverse context influences political views, in a direction consistent with an increased concern towards U.S. nationals.

I am grateful to my supervisors – Bertil Tungodden, Heidi Thysen, Matthew Rabin and Alexander Cappelen – for their advice and discussions. This paper benefited from helpful comments from Christine Exley, Nickolas Gagnon, Uri Gneezy, Boon Han Koh, Marlis Schneider, and seminar audiences at NHH and conferences. Special thanks to Catalina Franco, Akshay Moorthy and Laura Khoury for insightful discussions and feedback in the writing of the paper. Funding for this project was provided by the ERC Advanced Grant 788433. This study was preregistered in the AEA RCT Registry (AEARCTR-0010179).

1.1 Introduction

Our neighborhoods, classrooms, and workplaces are becoming increasingly diverse (Brookings, 2020). Yet even amidst this surge in diversity, segregation persists, potentially causing individuals to experience mere exposure to diverse groups, without actually engaging in genuine contact (Boustan, 2013; Hellerstein and Neumark, 2008). This heightened exposure to diversity raises a crucial question: whether changes in our *social context*, characterized by the individuals to whom we are *exposed*, influence the way we interact with each other. For instance, would the interaction between two white Americans, in a predominantly white community, change following an influx of foreign families?

This paper uses a preregistered large-scale experiment on a U.S. sample ($N \approx 2800$) to study whether a change in a social context has a causal effect on preferences for giving, which captures an important determinant of social cohesion and cooperation in interactions (Fehr and Fischbacher, 2003; Alan et al., 2021). Studying the effects of changes in social context on behavior comes with two main empirical challenges. First, observed changes in social contexts are often intertwined with other significant economic and social shifts, making it complex to pinpoint specific causality. Second, it is often not possible to distinguish in a social context whether an individual is just merely exposed or has genuine contact with others in their society, making it challenging to identify mechanisms and to specify the direction of the effect of a change in social context. This paper addresses these concerns by isolating mere exposure, a predominant situation and an overlooked channel, in an experimental design where social context is exogenously manipulated.

Participants from the U.S. make an incentivized allocation decision towards a receiver, who is either a fellow U.S. national or a foreigner, while exogenously exposed to a social context. To mimic the varying exposure to diversity in society, the decision-maker is exposed to one out of two social contexts: (i) a *homogeneous* context, where she observes only peers from a single country–e.g., she observes only U.S. nationals or only foreigners, and (ii) a *diverse* context, where she observes both fellow U.S. nationals and foreigners simultaneously. By comparing allocations across the two social contexts, I show that the social context a participant is exposed to has a substantial impact on giving. Facing a diverse context increases ingroup bias in favor of U.S. nationals, driven by both increased allocations towards U.S. nationals and decreased allocations towards foreigners relative to a homogeneous context. I provide suggestive evidence of changes in perceived social proximity towards the receiver as a mechanism of the effects social context on giving. Moreover, the increased altruism towards fellow U.S. nationals driven by changes in social context carries over to political preferences, by making U.S. participants care more about redistribution at the national level.

The experimental design introduces a way to exogenously manipulate social context. A decision-maker is exposed to a social context, composed of two other matched individuals. In each context, the two individuals are randomly assigned the roles of a receiver and a non-receiver, where the decision-maker can decide how much to redistribute from her endowment towards the receiver only. The non-receiver is key for the manipulation of social context. A decision-maker is in a homogeneous context when the receiver and non-receiver come from the same country; and the diverse context is the case when the receiver and non-receiver come from different countries. Thus, the experiment follows a between-subject 2x2 design, meaning that a decision-maker faces only one allocation decision which is either in a homogeneous or a diverse context, and directed towards a U.S. national or a foreigner. This design allows me to study the effects of changes in the social context on allocations towards both: fellow U.S. nationals and foreigners. In a diverse context individuals allocate 12% more to fellow U.S. nationals and 15% less to foreigners, relative to the respective allocations in a homogeneous context. This translates in an in-group bias in the diverse context where the decision-maker gives to a fellow U.S. national 28% more than what she gives to a foreigner. Notably, in homogeneous contexts allocations towards the U.S. nationals and to foreigners are the same. Taken together, these results suggest that social context is key for the emergence of in-group bias.

To shed light on the mechanisms behind the main results, I examine the relationship between the allocation decisions and perceived social proximity towards the receiver, measured using the Inclusion of the Other in the Self (IOS) scale developed by Aron, Aron and Smollan (1992) and used in recent work in economics (Goette and Tripodi, 2021; Bicchieri et al., 2022; Gächter, Starmer and Tufano, 2022). The analysis follows a simple formal framework, based on two premises. First, drawing from economic studies on identity, a decision-maker places greater weight on the receiver's pay-offs when the receiver's identity is perceived as closer to their own (Chen and Li, 2009; Leider et al., 2009; Robson, 2021). Second, this perception of social proximity can be influenced by the prevailing social context. Following Bordalo et al. (2016*a*), exposure to diversity modifies the decision-maker's comparison group, through the non-receiver. By contrasting the receiver's group with the new and distinct comparison group, perceptions of social proximity shift in a manner that magnifies group differences, enhancing social proximity to U.S. nationals. The framework predicts that higher perceived social proximity to the ingroup leads to a rise in ingroup-biased allocations.

Therefore, I hypothesize that diversity increases ingroup favoritism in both social proximity and, as a consequence, in allocations relative to a homogeneous society, which I test using the collected measure of perceived social proximity. My findings provide support for the two premises of the model. First, as allocation decisions and perceived social proximity exhibit a strong positive correlation. Second, in diverse contexts, decision-makers feel closer to the ingroup relative to homogeneous contexts. Finally, I relate the results on social proximity with allocations, where I find that the effects on allocations are partially driven by the individuals whose proximity was affected by the social context. Collectively, the evidence suggests that the effects of context on allocations are substantially driven by shifts in perceived social proximity.

I demonstrate the robustness of the results by analyzing the effects of social context on allocations and social proximity separately for each of the foreign countries used in the experiment: China and Canada. I find that the results are consistent regardless of the foreign country used, which suggests that the effects are not driven by country-specific beliefs or attitudes. Furthermore, taking advantage of the large-scale U.S. sample, I study heterogeneity in the effects of social context on allocations and perceived social proximity exploiting a set of sociodemographic characteristics. I find no heterogeneity across education, age, sex, political affiliation, or race, and the results are consistent across these dimensions. The consistency of my findings across subgroups of the population suggests that the study potentially identifies a general feature in human behavior.

Finally, I explore whether the increase in preferences for giving towards fellow anonymous U.S. nationals, following a change in social context, carries over to political preferences over redistributive policies. Participants indicated, in a hypothetical scenario, whether they agree with prioritizing nationwide over local redistribution of a tax raised at a local community level, in the form of welfare payments. My findings are twofold: first, participants giving higher amounts towards other anonymous U.S. nationals exhibit a higher level of agreement with prioritizing nationwide redistribution. Second, exposure to international diversity not only increases giving towards fellow anonymous U.S. nationals, but also enhances support over prioritizing nationwide distribution. This suggests that exposure to diversity reshapes the boundaries of the ingroup from local to national, which is consistent with work that shows how changes in social context reshape ingroup boundaries over intergroup interactions (Fouka and Tabellini, 2022).

This paper contributes to our knowledge in several strands of literature. First, this paper enriches the interdisciplinary research that delves into the implications of diversity and context on intergroup relations (Allport, 1954; Rao, 2019; Mousa, 2020; Lowe, 2021; Nathan and Sands, 2023). While previous studies have predominantly centered on intergroup interactions, e.g., interactions of natives and foreigners; this paper shows that diversity also affects intragroup interactions, e.g., how natives interact with each other. In a recent paper, Anderberg et al. (2023) study natives' decisions in a trust game towards other natives and immigrants in a classroom context, using survey data in Germany. Similar to my findings, they show that classroom levels of diversity affect ingroup bias.

My paper differs from their approach by using an experiment that allows me to additionally assess: first, how exposure to diversity affects altruism, using a non-strategic setting; second, how mere exposure to diversity affects behavior, using a controlled setting that isolates this channel; and finally, the role of perceived social proximity as a driver of the effects of social context.

This research is grounded in theoretical work from economics and psychology on how group identity is affected by contextual factors and plays a role in economic decisions (Tajfel and Turner, 1979; Turner et al., 1987; Heidhues, Kőszegi and Strack, 2020; Grossman and Helpman, 2018; Bonomi, Gennaioli and Tabellini, 2021). I contribute to this work by providing empirical evidence of the effect of social context on a key economic behavior, which is altruism. Moreover, drawing from theoretical research on the effects of context on beliefs and perceptions (Esponda, Oprea and Yuksel, 2023; Bordalo et al., 2016*a*), I provide evidence of a novel mechanism through which social context affects altruism, which is changes in perceived social proximity.

Relatedly, this paper enriches the literature on ingroup bias in social preferences (Luttmer, 2001; Chen and Li, 2009; Charness and Chen, 2020; Shayo, 2020; Kranton et al., 2020) and discrimination (Bertrand and Mullainathan, 2004; Lane, 2016; Achard and Suetens, 2023). While extensive research has documented the existence of ingroup bias in different settings, less work has been directed on when this bias is present or what determines its strength (Hett, Mechtel and Kröll, 2020). My findings highlight the importance of social context on the intensity and presence of ingroup favoritism.

Finally, my findings contribute to the literature on the effects of social context on policy views (Condra and Linardi, 2019; Hangartner et al., 2019; Steinmayr, 2020; Fouka, Mazumder and Tabellini, 2022; Alesina and Tabellini, 2022). Most work focuses on understanding the effects of diversity or migrants on policy views regarding the outgroup, e.g., anti-immigrant sentiment or immigration policies. My paper instead focuses on the effects of social context on policies associated with the ingroup and links the novel mechanism of changes in perceived social proximity towards the ingroup as an additional driver of political preferences.

The remainder of the paper is structured as follows. In Section 1.2, I propose the experimental setup that allows me to identify social context effects on willingness to help. Section 1.3 defines the empirical strategy and indicates the hypotheses. In Section 1.4, the main results are presented, as well as suggestive evidence of the mechanism and implications on policy views. Finally, in Section 1.5, I discuss the potential extensions of the paper and future directions.

1.2 Experimental Design

I first describe how social context is defined in my experimental framework. Subsequently, I explain the main task of the decision maker (DM): the allocation decision. Then, I illustrate the manipulation of social context in the experimental setting. I proceed to provide an overview of the secondary outcomes collected in the experiment. Finally, I describe the treatment conditions and summarize the sample and procedures.

Social context. In the experiment, the group membership that characterizes social context is given by nationality, where the DM is always from the USA.[1] The DM can be exposed to either a U.S.

[1]The exact definition of the group membership would be the country where

national, with whom she shares the same group membership, or a foreigner, who differs from the DM in group membership. The study focuses on two types of social contexts a DM can be exposed to. First, a *homogeneous* context, where the DM is exposed to a set of peers from a single country, e.g., the DM observes only Americans or only foreigners. Second, a *diverse* context, where the DM is exposed to a heterogeneous set of peers, where different countries contrast with each other, e.g., the DM observes both Americans and foreigners simultaneously. The foreign countries used in the experiment correspond to China and Canada.

Allocation decision. I measure prosocial behavior using a redistributive allocation decision. Each DM is matched with two other participants of the study (A and B) and informed about their group membership. All three, together, share a total of \$50, initially distributed between the three as follows: \$40 for the respondent, and \$5 for each of the matched participants.[2] After providing the information about the initial distribution to the DM, one of the two matched participants is randomly selected with equal probability, who I refer to as the *receiver*. The DM is informed about the selection process. To determine the final pay-offs, the DM must decide how much of her initial endowment of \$40 she would allocate *only* towards the receiver. Note that final pay-offs of the DM and the receiver depend on the allocation choice of the DM, whereas the final pay-off of the unselected participant, which I refer to as the *non-receiver*, is \$5 regardless of the choice of the DM. Nonetheless, the presence of the non-receiver is key for manipulating social context.

the participant of the study was sampled from. However, for simplification in exposition, I refer to it as nationality.

[2] All participants are asked to perform the task, however, they are informed that only a randomly selected subset of the participants will receive the total amount.

Importantly, the respondent is informed that she is the only participant in their group that is able to change the initial distribution of the pay-offs, and that the decision is anonymous. This is done as a means to withdraw potential signaling and social image concerns in determining the allocation decision (Andreoni and Bernheim, 2009).

Figure 1.1 shows a screen capture of two potential situations a DM may be subject to. In each panel, the DM is matched with two participants and is informed about the country of each participant. In red, it is indicated which participant was randomly selected as a receiver. In the first situation (Panel 1.1a), the DM (USA) is matched to participants A and B, who are both fellow U.S. nationals. Participant B was selected, and therefore the DM must decide how much to allocate towards a fellow U.S. national. In the second situation (Panel 1.1b), the DM is matched with participant A, who is a foreigner, and participant B, who is a fellow U.S. national. Participant B was selected, and therefore the DM must decide on how much to allocate towards a fellow U.S national. This example highlights the two key features of the design. First, in both situations the decision and choice set is the same, corresponding to an allocation towards a U.S. national. Second, across situations the only aspect that differs is social context: the first situation corresponds to a homogeneous context (A and B are both U.S. nationals), whereas the second situation corresponds to a diverse context (A is a U.S. national and B is a foreigner). Note that social context is manipulated by only modifying the non-receiver.

Social proximity elicitation. To understand the effects of social context on perceptions in the social identity domain, I elicit a proxy measure of social proximity between individuals. Using the Inclusion of Other in the Self (IOS) scale developed by Aron, Aron and

Smollan (1992), I ask the respondents to indicate *how close they feel* towards the receiver using two overlapping circles, where no overlap indicates not close at all, and the greater the overlap between the circles, the closer the respondent feels. This measure is unincentivized; however, it has been validated as a reliable measure for social proximity in comparison to other more sophisticated survey methods (Gächter, Starmer and Tufano, 2015) and has seen increased use in recent experiments in psychology and economics (Goette and Tripodi, 2021; Bicchieri et al., 2022; Gächter, Starmer and Tufano, 2022).[3]

Policy views. Based on the work of Enke, Rodríguez-Padilla and Zimmermann (2022) and Cappelen, Enke and Tungodden (2023), I elicit preferences over policy views, where respondents indicate their support towards a redistributive policy. The participants are asked to indicate their agreement in a 5 points scale from "Strongly Disagree" to "Strongly Agree" (-2 to 2), towards the following policy:

"The government should redistribute local tax revenues as welfare payments across all communities nationwide, rather than only within the local communities they were raised."

Other outcomes. In the survey, key demographic information such as education, political affiliation, age, ethnicity, and gender are collected.

[3]The order of presentation of the allocation task and the social proximity elicitation is randomized.

1.2.1 Treatment conditions

Figure 1.2 provides an overview of the experimental design. I use a 2x2 between-subject design, where subjects are randomized across two dimensions. First, whether the receiver is a U.S. national or a foreigner, from either China or Canada. Second, whether the allocation takes place in a homogeneous context, where both matched participants come from the same country (e.g., A and B are both from China) or a diverse context, where one matched participant is from the U.S., and the other is from a foreign country.[4] The randomization procedure defines the following four treatments:

1. Homogenous - U.S. National. $N \approx 400$.

2. Homogenous - Foreigner. $N \approx 800$.

3. Diverse - U.S. National. $N \approx 800$.

4. Diverse - Foreigner. $N \approx 800$.

In the latter treatments that involve the foreigners, I split the subsamples according to which country was used as the foreign country: China or Canada. The decision of selecting China and Canada as the foreigners emerged from the interest in analyzing the context effects from two different types of outgroups, defined in the preregistration. First, the *strong outgroup* (China) which corresponds to a foreign group that is dissimilar in observable characteristics to the USA group. Second, I define the *weak outgroup* (Canada), which corresponds to a foreign group that is more similar to the USA group. I expected, accordingly, that the social prox-

[4]This setting considers societies with members only from the USA and/or Canada, or societies with members only from the USA and/or China. Other situations that combine both of the foreigners together (Canada and China) are outside the scope of this paper.

imity towards participants from China and Canada would be different. Nonetheless, Figure A1 in Appendix 1.B shows that there is no difference in social proximity towards the two selected groups of foreigners, in the homogeneous setting. I interpret this finding as if both Canada and China are perceived as an equal outgroup. As a result, I focus on the analysis of the pooled sample, using both countries together as foreigners.

1.2.2 Experimental procedures

The survey was conducted online by survey provider Dynata. The company recruited a representative sample of the US population, between October 14th-31st, 2022. There was a total of 2808 valid participants who passed an initial mandatory attention check. Respondents were stratified to match the adult population by age, sex, and geography. The average response duration was around 5 minutes. The experimental design and the hypotheses were preregistered at the AEA RCT Registry (AEARCTR-0010179). The full instructions are made available in Appendix 1.C.

Table A1 in Appendix 1.B provides a summary of the outcomes collected in the experiment. The sample consists of around 60% women, around a third identifying as Republican, a mean age of 50, mostly white (87%), and a majority college-educated (66%). I observe a successful randomization, where individual characteristics are balanced across treatments.

1.3 Empirical strategy

Following my 2×2 experimental design, here I outline the main specification for the analysis, which has as the baseline group the

Homogenous - U.S. National treatment:

$$y_i = \beta_0 + \beta_1 Diverse_i + \beta_2 Foreigner_i + \beta_3 Diverse_i \times Foreigner_i + X_i + \varepsilon_i \quad (1.1)$$

where y_i corresponds to the allocation decision, $Diverse_i$ corresponds to an indicator variable with value 1 if the allocation decision is in the diverse context, $Foreigner_i$ is an indicator variable with value 1 if the allocation is towards a foreigner, and X_i is a set of collected covariates at the individual level.

I am interested in two main effects captured by this specification. The first corresponds to the effect of a change in social context on the allocation towards U.S. nationals, which is captured by β_1 in equation 1.1. Second, I am interested in the difference across social contexts of the ingroup bias, defined as the differences in the mean allocation towards a U.S. national vs towards a foreigner. The 2x2 experimental design allows for a difference-in-difference approach to estimate the effect of a change in social context on the ingroup bias, which is captured by $-\beta_3$ in equation 1.1. Additionally, β_2 provides the estimated difference between the allocations towards a U.S. national and a foreigner in the homogeneous treatment.

1.3.1 Hypotheses

The experiment represents a society of $N = 3$, with a DM, a receiver, and a non-receiver. The DM is always from the USA and is exposed to two individuals that can be either a fellow U.S. national, or a foreigner (China or Canada). The analysis follows a simple formal framework, described in Appendix 1.A, and is based on two premises, based on work in economics and psychology.

The first premise is that a decision-maker places greater weight on the receiver's pay-offs when the receiver is perceived as closer to their own. This draws from research in economics on how individuals behave more altruistically to people who are perceived as part of their ingroup (Chen and Li, 2009), who are socially closer (Leider et al., 2009), and who are perceived as closer (Robson, 2021). The second premise indicates that this perception of social proximity can be affected by the social context, by changing the reference groups. This insight draws from extensive research in psychology which suggests that context exerts a significant influence on perception (Gold, 2014).[5] Recent research in economics has incorporated these insights into the social domain. For instance, Esponda, Oprea and Yuksel (2023) shows that inference when assessing new information about an individual, pertaining to a specific group, can be biased by the context, which is represented by the reference group the individual is compared to, in a phenomenon closely associated with Kahneman and Tversky (1972)'s "representativeness heuristic".

Under the two premises, a change in social context from homogeneous to diverse would affect the reference group of the decision-maker when assessing social proximity. To assess the impact of the reference group, I follow Bordalo et al. (2016*a*), where by contrasting the country of the receiver and the country of the non-receiver, the decision-maker would bias her perceived social proximity to-

[5]A very influential example corresponds to the Ebbinghaus-Titchener illusion (Titchener, 1901). Two circles of the same size are surrounded by a different context each: the first circle is surrounded by small circles and the second circle is surrounded by big circles. When most observers view these figures, the context affects perceptions of size, through contrasting with the surrounding circles and makes individuals perceive as if they have a different size. This has sparked an extensive literature on the effects of context on perceptions.

wards the receiver in a manner that magnifies group differences. This heightened perceived social proximity to the U.S. nationals driven by contrast with the foreigner leads to a rise in ingroup-biased allocations. The model provides the following preregistered hypotheses:

Hypothesis 1.3.1. *Allocations towards U.S. nationals (foreigners) are higher (lower) in a diverse context with respect to a homogeneous context* ($\beta_1 > 0$)*.*

Hypothesis 1.3.2. *The ingroup bias in allocations is higher in a diverse context with respect to a homogeneous context* ($-\beta_3 > 0$)*.*

Hypothesis 1.3.3. *The mechanism behind the effects in hypotheses 1.3.1 and 1.3.2, correspond to changes in perceived social proximity as a consequence of changing from a homogeneous to a diverse context.*

1.4 Results

In the subsequent sections, I delve into the impact of transitioning from a homogeneous to a diverse social context on the incentivized allocation decision. Subsequently, I provide suggestive evidence pointing to perceived social proximity as a mechanism of the effect of social context on allocations. Leveraging the large-scale sample, I explore variations across population subgroups and find the effects documented are consistent across demographic groups. Finally, I analyze the political implications of the main experimental findings.

1.4.1 Diversity and allocations

First, I describe how participants in the experiment responded the allocation task descriptively. A significant majority, over 80% of participants, designated a positive amount to the receiver (see Figure A2 in Appendix 1.B). The average allocation stands at approximately 12 USD out of a possible 40 USD, aligning with previous studies on giving (Engel, 2011). I now present evidence of the impact on social context on giving.

Figure 1.3 summarizes the allocations of the decision-makers. The patterns underline the influence of social context on altruism towards either a fellow U.S. National (orange circle) or a foreigner (blue triangle). In a diverse setting, the mean allocation to a fellow U.S. national surpasses that in a homogeneous context. On the other hand, allocations to foreigners decrease in diverse settings. Thus, the ingroup bias in allocations, defined as the difference between the allocations towards fellow U.S. Nationals and towards foreigners, in a diverse context is broader than in a homogeneous context. Notably, the bias is negligible in a homogeneous setting. Collectively, these findings suggest that diversity skews allocations in favor of U.S. nationals, and that social context is key for the emergence of ingroup bias.

I estimate equation 1.1 over the full sample, shown in Columns 1 and 2 of Table 1.1 with allocations in USD as the dependent variable. The coefficient on the variable "Diverse" indicates the treatment effect of exposure to diversity on allocations to U.S. nationals. The average allocation to an American in a diverse context exceeds its counterpart in a homogeneous context by 1.4 USD ($p=0.03$), representing an approximate 12% increase (column 1). This increase remains robust even when controlling by the set of collected individual-level characteristics. On the other hand, the allocation towards a foreigner, given by the sum of the coefficients "Diverse" and "Diverse x Foreigner", is 1.9 USD lower in a diverse context relative the homogeneous context ($p<0.01$), representing a decrease of around 15%. The results on allocations are robust to controlling by demographic characteristics and state fixed effects.

The shift in mean allocation towards U.S. nationals is driven more by distribution changes in the amount given rather than by the number of participants allocating positively (see Figures A3 of Appendix 1.B). A chi-squared test shows that there is no difference in the proportion of respondents that give a positive amount across contexts.

Both, the increase in the allocations to the ingroup and the decrease in the allocations to the outgroup after exposure to diversity, imply an increase in the ingroup bias. The 2x2 experimental design facilitates a difference-in-differences approach that allow us to study the effect of exposure to diversity in the ingroup bias. A negative coefficient of the interaction term "Diversity x Foreigner" indicates a widening of the ingroup bias in diverse context. The regressions estimates show a significant widening of the ingroup bias in diverse contexts. This results in an ingroup bias in a diverse society where allocations towards fellow U.S. nationals are 28% higher than towards foreigners (column 1 of Table 1.1). This confirms that diversity not only bolsters allocation towards U.S. nationals but also accentuates disparities in behaviors towards U.S. nationals and foreigners.

A country-wise breakdown in Table 1.1, analyzing the data separately for each of the foreign countries: China and Canada, confirms the consistency of this result across both countries. Moreover, the differences in the point estimates for China (column 3) and Canada (column 5) are not statistically significant ($p>0.3$). This provides evidence that country-specific beliefs or attitudes regarding participants from China or Canada are not the drivers of the results, e.g., beliefs about income or attitudes towards specific nationalities.

Finally, I analyze the ingroup bias the homogeneous context. The coefficient "Foreigners" reflects the difference between the allocation towards a U.S. national and towards a foreigner in the homogeneous context alone. Notably, there is no ingroup bias in the homogeneous context ($p=0.54$). This holds true even when examining the bias specifically for Canada or China. A joint-F test for null hypothesis of both countries as well as individual null hypothesis tests for each country show no positive ingroup bias in a homogeneous context ($p>0.3$). Furthermore, I analyze the bias for each partition in the set of collected characteristics, where I find no positive ingroup bias (see Figure A5 in Appendix 1.B).

Result 1. *Exposure to diversity increases ingroup bias, where individuals allocate more to U.S. nationals and less to foreigners relative to a homogeneous context. Notably, decision-makers in a homogeneous contexts do not exhibit such bias.*

These findings highlight that social context is key for the emergence and strength of ingroup bias. The results do not contradict previous work on identity effects, as in most studies that investigate ingroup bias, to the best of my knowledge, the experimental design is either within-subject, or in a setting that highlights the diversity with competing identities (Charness and Chen, 2020; Shayo, 2020). However, the fact that informing individuals about group membership alone in this set up does not generate ingroup bias, highlights the importance on social context in the origins of ingroup favoritism.

Appendix 1.B introduces various robustness checks of the analysis of treatment effects for on the allocation task. To address the concern of whether the effects on the mean allocation are driven by extreme outcomes, I analyze the outcomes using non-parametric Mann-Whitney tests and Randomization Inference tests (Young, 2019), where I find that the results are robust ($p<0.04$) for a two-sided test in the pooled sample (see A3 for the histogram of allocations by treatment) over the effects over the allocations towards both, U.S. nationals and foreigners. Moreover, I run the same specification 1.1 on a subsample of the population after excluding the individuals that allocated all their pot (40 USD) and the results persist (see Table A2).

1.4.2 Perceived social proximity as a mechanism

First, I describe the perceived social proximity elicitation across the full sample. More than 40% of the participants indicated the lowest level on the IOS Scale, which is anticipated since the matched participants are unfamiliar to the decision-makers (DMs). Second, I present evidence of the role of perceived social proximity as a mechanism in the impact on social context on giving.

As posited in section 1.3, the first premise of the formal framework theorized that allocation decisions relate positively with perceived social proximity. Figure 1.4 shows the distribution and mean allocation of a respondent towards an individual with a fixed level i in the IOS scale, for all levels of the IOS scale $i = 1, ..., 7$, where 1 is not close and 7 is very close. The figure establishes a clear correlation: the closer respondents felt to a participant, the more they allocated. The statistical relationship is positive and compelling ($p<0.01$). Moreover, non-parametric Mann-Whitney tests show, for each consecutive comparison of levels of social proximity, that the closer the selected participant is perceived the higher is the distribution of allocations, ($p<0.05$).

The second premise suggests that changes in the social context affects perceived social proximity. The model predicted that perceived social proximity is higher in a diverse than in a homogeneous context. Figure 1.5 suggests that perceived social proximity exhibits a similar pattern as allocation decisions (Figure 1.3). The perceived social proximity towards fellow U.S. nationals increases in a diverse context (orange circle), widening the ingroup bias in proximity. However, the proximity towards foreigners seems to be unaffected (blue triangle).

Table 1.2 reports the coefficients from the estimation of equation 1.1 where social proximity is used as the dependent variable (columns 1 and 2). The coefficient "Diverse" reveals an increase in perceived proximity towards U.S. nationals in diverse contexts of 28%. Moreover, the opposite of the interaction effect between "Diverse" and "Foreigner" demonstrates that showcasing diverse identities increases the ingroup bias in social proximity. Therefore, the treatment triggers the same effect for the U.S. nationals in both outcomes, the allocation and the social proximity. However, the social proximity to the foreigners is not influenced by exposure to diversity. Table A3 in Appendix 1.B show that a country-wise breakdown of the treatment effects show the same pattern, regardless of my choice of foreign country.

I now turn my attention to the role of perceived social proximity as a mediator of the treatment effect in the allocation decision. Columns 3 to 6 in Table 1.2 display the regression estimates based on specification 1.1, with the allocation serving as the dependent variable. Interestingly, when perceived social proximity is included as a control variable, the effect of diversity on the allocations towards U.S. nationals vanishes. Additionally, the coefficient capturing the interaction effect diminishes considerably. While this coefficient becomes significantly smaller, it retains its significance. The explanatory power of the linear regression also sees a notable boost, which aligns with strong positive relationship between perceived social proximity and allocation.

Result 2. *Evidence suggests that the increase in ingroup bias in allocations is driven by changes in perceived social proximity caused by exposure to diversity.*

The influence of social proximity as a mediator of the effect of social context on prosocial behavior stands strong under different types of analysis, presented in Appendix 1.B. Table A3 provides a country-wise breakdown of the analysis, and the results are consistent when separating by foreign country. Table A4 presents mediation analysis following Imai, Keele and Yamamoto (2010) and Heckman and Pinto (2015). This analysis leverages exogenous variation in two dimensions: the group membership of the receiver and the social context exposed to. The aim is to determine if social proximity can account for a portion of the treatment effect on allocation towards the ingroup. Both the indirect and direct effects of the treatment remain significantly positive after introducing social proximity as a mediator.

Discussion and other mechanisms

While this paper presents compelling evidence for social proximity influencing ingroup allocations, it is vital to consider other possible drivers. In the experimental framework, every dollar allocated to the receiver creates a disparity between the receiver and the non-receiver. It is plausible that individuals display increased aversion to inequality when faced with two members of the same country, compared to when they are distinct. Although Figure 1.5 and table 1.2 reveal a trend in line with this hypothesis for U.S. nationals, this is not the case for the allocations towards foreigners, where a DM appears more tolerant of inequality in homogeneous contexts than in diverse ones. Moreover, if this alternate mechanism were solely responsible for the effects on U.S. nationals, I would expect a a positive and significant residual of the treatment effect in the coefficient "Diverse" in table 1.2, after controlling for perceived social proximity, which does not hold.

As depicted in Figure 1.5, social context has no influence on the social proximity towards foreigners. Figure A4b highlights that most participants indicated the lowest level of perceived social proximity, 1, towards foreigners in the homogeneous context. This suggests that a potential explanation of the null effect of context on social proximity might be associated to reaching the lower bound in the choice set. Regardless of the explanation for the negligible effect, changes in social proximity cannot account for the observed effects on allocations towards foreigners. This leaves a segment of the ingroup bias in allocations unexplained.

There are two alternative explanations. First, in the allocation decision, the weight that the DM puts on the receiver's pay-offs might be dependent on the relative proximity, defined as the difference between proximity between the U.S. nationals and foreigners, instead of the absolute proximity to the receiver. This relative proximity seems to surge in diverse settings, as evident in Figure 1.5. This intuition relates to the work of social psychology on social categorization theory, where individuals when classifying others as ingroup or outgroup, they use a *meta-contrast principle* which maximizes the difference across groups and minimize distance within groups (Tajfel and Turner, 1979; Turner et al., 1987). An adaptation of the conceptual framework presented in Appendix 1.A, where absolute proximity to the receiver is substituted for relative proximity, provides us with predictions consistent with the empirical findings of the experiment. However, individual-level data on relative proximity is missing, limiting the ability to discern any patterns.

A second theory proposes variable levels of inequality aversion based on the social context and which group benefits from the inequality. My research indicates that allocations shift in favor of the ingroup, creating disparities. Moreover, allocations towards U.S. nationals are driven by the individuals with changes in social proximity, leaving no residual of the effect to be solely attributed to inequality aversion. Testing this for allocations towards foreigners is not feasible.

1.4.3 Heterogeneity analysis

By leveraging the representative sample in this study, I look into the heterogeneity of effects across various population subgroups. Figure 1.6 shows the coefficients of the effects of diversity on (i) the allocations towards U.S. nationals, (ii) the ingroup bias in allocations, and (iii) the perceived social proximity towards U.S. nationals, which serves as a mechanism in the effects, across subgroups formed from the collected set of demographic characteristics. While there are no differences accross groups in the coefficients presented, they are systematically consistent with the hypotheses, and present across any subgroup of the set of covariates. This is suggestive evidence that the results seem to reflect a general feature in decision-making.[6]

Result 3. *The effects of diversity on allocations and perceived social proximity are consistent and robust, irrespective of population subgroups.*

1.4.4 Implications: policy views

I conduct an exploratory analysis to assess policy relevant implications of the experimental findings, focusing on preferences over redistributive policies. For this analysis, I am interested in understanding whether the effects in social context over behavior and perceptions towards U.S. nationals carry over policy outcomes targeted to Americans.

Policy views. I elicited preferences over a redistributive policy, where participants were asked to indicate their agreement with a statement prioritizing nationwide over local redistribution of a tax that was collected locally, in the form of welfare payments. The agreement scale is a 5-point scale with a value of -2 for "Strongly Disagree" and a value of 2 for "Strongly Agree".

Around 37% of participants agree with prioritizing nationwide over local redistribution. To relate the behavioral measures with agreement on this policy, column 1 of Table 1.3 shows how agreement with prioritizing nationwide redistribution relates to the allocation in USD towards a U.S. national. The more a U.S. decision-maker gives towards an anonymous fellow U.S. national, the more she supports prioritizing nationwide redistribution ($p<0.01$).

The aim is evaluating whether the increased preferences for giving towards anonymous fellow U.S. nationals generated by ex-

[6] Note that the analysis by subgroup splits the sample, which eventually reduces power to detect statistical impacts, however, in all three variables I find at least one subgroup to have a significant positive effect ($p<0.05$) with correction for multiple hypothesis testing.

posure to diversity carries over into the political domain. As a reminder, exposure to a diverse context corresponds to a DM facing a fellow U.S. national and a participant living in a different country (China or Canada), who does *not* benefit from the hypothetical policy.

Table 1.3 reports the coefficients from the regression of the level of agreement (5 points scale) on an indicator variable "Diverse" that takes value 1 if the context is diverse. The effect of prior exposure to diversity over the average agreement to the statement that prioritizes nationwide over local redistribution. The table reveals that a diverse context bolsters support for nationwide policies relative to a homogeneous context. The effect size corresponds to an increase in the average agreement of 0.11, on a 5-point scale. These results are robust to alternative specification choices, such as employing an ordered probit or assessing the effect on the percentage that agrees, as shown in Table A7 in Appendix 1.B. Furthermore, there is no heterogeneity in the effects by subgroups of the population, and the effect is consistent across subgroups (see Figure A6 in the Appendix 1.B). Finally, Table 1.3 also incorporates demographic controls, where the relationship of the level of agreement with the demographic characteristics is consistent with previous research on communal versus national support for policies (Enke, Rodríguez-Padilla and Zimmermann, 2022; Cappelen, Enke and Tungodden, 2023).

Collectively, these results provide evidence of the treatment effects of exposure to diversity on preferences for giving carrying over into policy views, in a manner consistent with increased ingroup favoritism towards U.S. nationals. An increase in giving to anonymous U.S. nationals generated by exposure to diversity translates into higher support for national policies. These findings align with empirical work in political science and economics on the effects of migration or war in shaping ingroup boundaries and nation-building, where the emergence of an opposing third party amplifies nationalistic sentiments (Fouka and Tabellini, 2022).

Result 4. *The increase in giving towards anonymous fellow U.S. nationals caused by a change in social context carries into the political domain, making participants support more national redistribution.*

1.5 Conclusion

In a world increasingly marked by diversity in neighborhoods, schools, and workplaces, understanding the effects of changes in a social context on behavior becomes crucial. This paper explored how changes in social context influenced perceptions of social proximity and consequent prosocial behavior. Using a novel experimental approach, I highlight role of social context on the emergence of ingroup bias with implications over policy preferences.

Four main findings emerge. First, the exposure to a diverse context intensifies ingroup bias in allocations favoring U.S. nationals. Second, shifts in perceived social proximity –caused by changes in the social context– correspond to a mechanism in the effects on allocations. Third, the effects of social context on both allocations and perceived social proximity are robust and consistent across subgroups of the population. Finally, the increase in giving towards anonymous U.S. nationals triggered by exposure to diversity carries over political attitudes, increasing support for nationwide redistributive policies. Altogether, these findings offer crucial insights into how social contexts can affect behavior in political and economic decisions.

The fact that mere exposure to diversity can increase ingroup bias in prosocial behavior, highlights that if a social planner's goal is to increase social cohesion or intergroup cooperations, it is important to avoid situations in which social contexts correspond to mere exposure to diversity, as my findings suggest biases might emerge. The social planner should generate conditions such that the proposed perceptual mechanism has a diminished role in exacerbating differences in behavior towards ingroup or outgroup members.

However, more exploration regarding different ways in which context can affect prosocial behavior is yet to be explored. This paper abstracts from social signaling concerns in the allocation decisions. However, whether the receiver or non-receiver is aware or not about the allocation decision made by the decision-maker, might affect their behavior, and affect the intensity of the results, which corresponds to an important extension to explore.

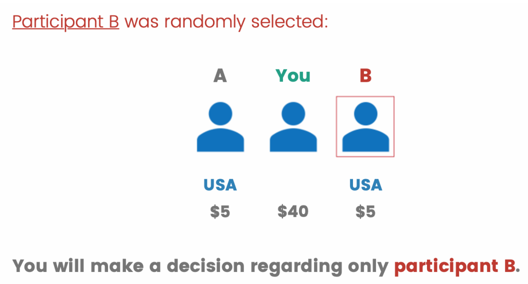
In this paper, I isolate the effect of mere exposure to diversity on interactions, and provide a mechanism and measure through which prosocial behavior is affected by social context. The experimental design focuses on immediate and short term exposure. Future work should address how the mechanism proposed evolves when the exposure is long term. Moreover it is also key to understand, how perceived social proximity towards both: ingroups and outgroups, changes when contact in the society happens.

Further explorations that attempt to generalize the effects can be performed. For instance, studying whether the effects differ or persist across other types of identities: e.g., ethnicity, religion, etc. It is possible to make use of lab-in-the-field tools, in schools or universities to understand whether the results are robust to settings where identity is not explicitly indicated. Furthermore, another direction is to extend the scope of study, and evaluate the effects of social context on other types of prosocial behavior, such as trust and cooperation. Implications of a more complete set of choices can speak to a wider set of real life settings, and give a better understanding over the relevance of social context on prosocial behavior.

Finally, this study provides a building stone for future research on the effects on context on decision making. The experimental design allows the researcher to cleanly study context effects in a wide array of decision sets, particularly within the framework of contrast biased evaluations.

1.6 Main Figures

Figure 1.1: Example of a change in social context from homogeneous (above) to diverse (below) in the allocation toward U.S. nationals.



(a) Treatment **Homogeneous - U.S. National**

(b) Treatment: **Diverse - U.S. National**

Figure 1.2: Overview of the randomization in the experiment

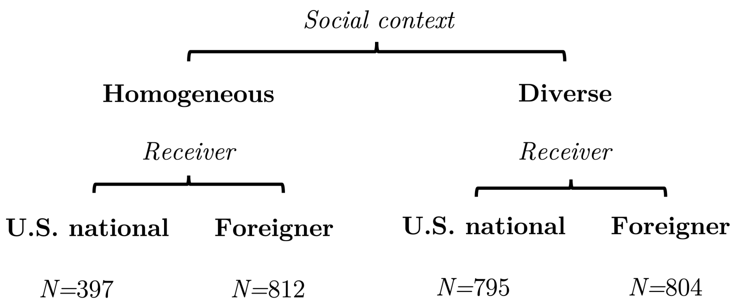
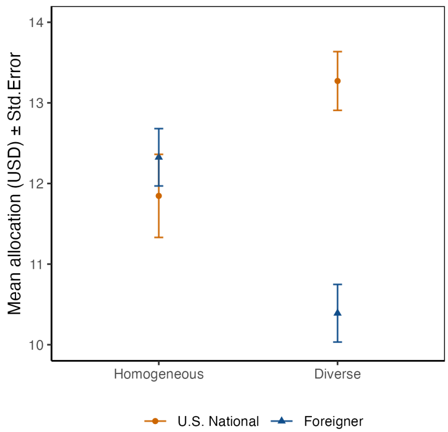
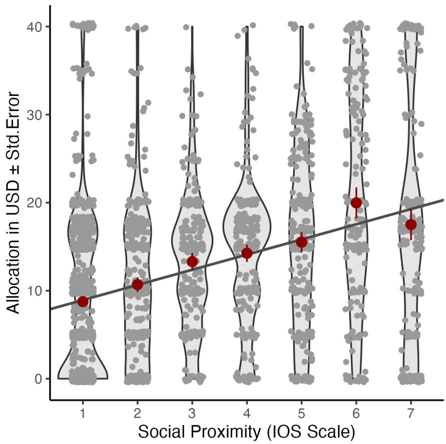


Figure 1.3: Mean allocations towards U.S. nationals and foreigners in the homogeneous and diverse contexts.



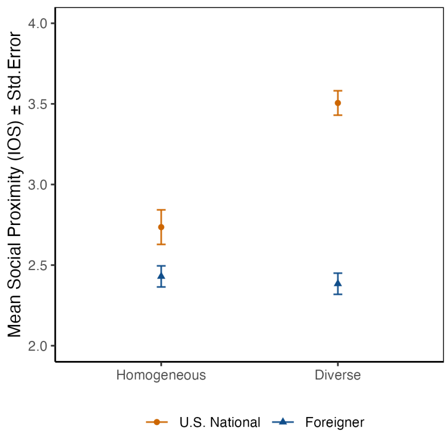
Notes: In this figure I show the mean Allocation in USD (values 0 to 40) $\pm$ Robust S.E., for each treatment in the 2x2 design: an allocation in either a homogeneous or diverse context, directed towards either a U.S. national or a foreigner.

Figure 1.4: Distribution and mean allocations for each level of social proximity using the IOS Scale (from 1-7).



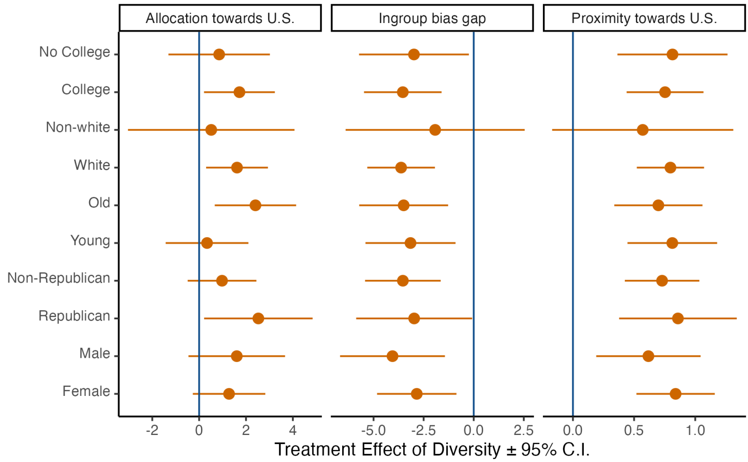
Notes: This figure shows the distribution of the allocation in USD, for each level of social proximity (scale from 1 to 7). I provide the mean allocation for each level of social proximity $\pm$ robust standard error, and a fitted line representing the correlation between social proximity and allocations.

Figure 1.5: Mean level of perceived social proximity by treatment group using the IOS Scale (from 1 to 7).



Notes: In this figure I show the mean perceived social proximity (values 1 to 7) $\pm$ Robust S.E., for each treatment in the 2x2 design: elicitation in either a homogeneous or diverse context, directed towards either a U.S. national or a foreigner.

Figure 1.6: Coefficients of the effect of diversity on the three main outcomes of analysis: allocations towards U.S. nationals, ingroup bias (negative coefficient means increase) and proximity to U.S. nationals, by subgroup of the population.



Notes: In this figure I show the point estimates of an OLS regression on the main specificacion when the dependent variable is Allocation, for the first two graphs, and when the dependent variable is Social Proximity for the third graph, for each subsample of covariates collected. I focus on coefficients: "Diverse" and "Diverse x Foreigner". Coefficient "Diverse" reflects the effect on allocation towards U.S. nationals when the dependent variable is Allocation (first graph), and reflects the effect over social proximity towards U.S. nationals when the dependent variable is Social Proximity (third graph). The interaction "Diverse x Foreigner" shows the opposite of the effect of diversity over the ingroup bias in allocations (second graph). This means that a negative coefficient is an increase in the ingroup bias. Old and Young correspond to above or below 50 years old (median age). I show the point estimates $\pm$ Robust S.E., for the regression without controls.

1.7 Main Tables

Table 1.1: OLS estimates of the regression on the allocation.

	Allocation in USD					
	Pooled sample		Foreigner: China		Foreigner: Canada	
	(1)	(2)	(3)	(4)	(5)	(6)
Diverse	1.425**	1.717***	1.223*	1.482**	1.631**	1.801**
	(0.613)	(0.623)	(0.726)	(0.749)	(0.712)	(0.740)
Foreigner	0.479	0.695	−0.122	0.070	1.059	1.204*
	(0.608)	(0.613)	(0.723)	(0.738)	(0.698)	(0.716)
Diverse x Foreigner	−3.360***	−3.567***	−3.992***	−3.960***	−2.675***	−2.812***
	(0.794)	(0.802)	(1.027)	(1.056)	(0.994)	(1.018)
Intercept	11.85	11.79	11.85	12.99	11.85	10.25
Controls	No	Yes	No	Yes	No	Yes
Observations	2,808	2,808	1,604	1,604	1,601	1,601
R^2	0.012	0.044	0.021	0.074	0.005	0.053

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table presents the results from an OLS regression with robust S.E. The variable Diverse takes value 1 when the participant is in the diverse context and 0 otherwise, and the variable Foreigner takes value 1 when the allocation is towards a foreigner or 0 otherwise. Diverse x Foreigner shows the estimate of the interaction. Columns 1 and 2 shows the estimates by using the entire sample. Column 3 and 4 estimate the coefficients using the subsample where the foreign country was specifically Canada, and columns 5 and 6 when the foreign country was China. The even columns show specifications that control for all covariates measured, except for the elicited social proximity. The controls include: gender, age, political party, education, race and state fixed effects.

Table 1.2: OLS estimates of the regression on the social proximity and analysis of mechanism of effects on allocation.

	Social Proximity (IOS)		Allocation in USD			
	(1)	(2)	(3)	(4)	(5)	(6)
Diverse	0.770^{***}	0.753^{***}	1.425^{**}	1.717^{***}	0.102	0.428
	(0.126)	(0.126)	(0.613)	(0.623)	(0.587)	(0.600)
Foreigner	-0.306^{***}	-0.298^{**}	0.479	0.695	1.004^{*}	1.205^{**}
	(0.119)	(0.118)	(0.608)	(0.613)	(0.572)	(0.580)
Diverse x Foreigner	-0.816^{***}	-0.816^{***}	-3.360^{***}	-3.567^{***}	-1.958^{***}	-2.171^{***}
	(0.157)	(0.156)	(0.794)	(0.802)	(0.753)	(0.764)
Proximity					1.719^{***}	1.712^{***}
					(0.107)	(0.109)
Intercept	2.74	4.13	11.85	11.79	7.14	4.73
Controls	No	Yes	No	Yes	No	Yes
Observations	2,808	2,808	2,808	2,808	2,808	2,808
R^2	0.055	0.117	0.012	0.044	0.122	0.146

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table presents the results from an OLS regression with robust S.E., over perceived social proximity, which takes value 1 to 7, and allocation in USD. The variable Diverse takes value 1 when the participant is in the diverse context and 0 otherwise, and the variable Foreigner takes value 1 when the allocation is towards a foreigner or 0 otherwise. Diverse x Foreigner shows the estimate of the interaction, what I call the diff-in-diff estimate. Columns 1 and 2 shows the coeficients of the regression with dependent variable Social Proximity. Column 3 to 6 estimate the coefficients with dependent variable Allocation and studies the role of social proximity in explaining the effects on the allocation. The controls include: gender, age, political party, education, race and state fixed effects.

Table 1.3: OLS estimates of the regression on the level of agreement of prioritizing nationwide over local redistribution of a tax.

	Support for Nationwide Welfare payments		
	(1)	(2)	(3)
Allocation (USD)	0.016***		
	(0.003)		
Diverse		0.101*	0.112**
		(0.052)	(0.050)
Male			0.194***
			(0.053)
Republican			−0.659***
			(0.055)
Older than 50			−0.324***
			(0.050)
College			−0.017
			(0.051)
Controls	No	No	Yes
Observations	2,808	2,808	2,808
R^2	0.014	0.001	0.085

Note: *p<0.1; **p<0.05; ***p<0.01

This table presents the results from an OLS regression with robust S.E., where the dependent variable corresponds the level of agreement over prioritizing nationwide over local redistribution of a tax collected locally, which is a 5-point scale ranging from -2 to 2. The variable Diverse takes value 1 when the participant is in the diverse context and 0 otherwise.

Appendices

1.A Conceptual framework

Consider a DM from the U.S. facing a society composed of herself (i) and two other individuals (j and $-j$). A DM must allocate resources with only one randomly selected individual: receiver j. In this society, each individual $k \in \{i, j, -j\}$ has group membership $g_k \in \{N, F\}$, which correspond to U.S. national (N) and foreigner (F). I can define G_s to be the set of existing groups that the DM is exposed to in a society s. I consider two types of social contexts characterized by G_s: a homogeneous context, where either $G_s = \{N, N\}$ or $G_s = \{F, F\}$, and a diverse context, where $G_s = \{N, F\}$.

The conceptual framework is based on two premises established in the literature. The first premise corresponds to the fact a decision-maker places a greater weight on the receiver's pay-offs when the receiver is perceived as closer to their own, which draws from research in economics on how individuals behave more altruistic to people that are perceived as part of their ingroup (Chen and Li, 2009), that are socially closer (Leider et al., 2009), and that are perceived as closer (Robson, 2021). The second premise indicates that this perception of social proximity can be affected by the social context, by changing the reference groups. This insight draws from extensive research in psychology suggests that context exert a significant influence on perception (Gold, 2014). In what follows, I present a formalization of each premise, and subsequently, I will present a series of hypothesis building from the premises.

1.A.1 Premise 1. Allocations and perceived social proximity are positively correlated

In the experiment, the selected players will earn together a total of \$50 dollars, which are initially distributed as \$40 for DM i, \$5 for receiver j and \$5 for non-receiver $-j$. The DM (i) chooses the amount that she wants to redistribute from her initial endowment towards the selected receiver (j). To inform the analysis, I assume that the DM is maximizing a utility function that depends on her own pay-off and what the selected individual j receives. The pay-off that the other individual $k \neq i, j$ in the society receives is not relevant for the DM, as the concerns over the earnings of any k are assumed to be separable, and the amount k receives is fixed. Following Cappelen et al. (2007) the optimization problem is represented as follows:

$$\max_{x_j} U_i(x_i, x_j) = x_i - \frac{\omega(\hat{\gamma}_j)}{2}(x_j - F_j)^2 \qquad \text{s.t.} \quad x_i + x_j = 45, \quad x_j \geq 5$$

where x_i and x_j correspond to the allocations that i and j receive, respectively; $\omega(.)$ represents the weight player i puts on allocating the fair outcome towards player j, relative to her selfishness; and $\hat{\gamma}_j$ is the perceived social proximity towards participant j, which I define as the probability of j being a part of the ingroup. Finally, F_j represents what the DM considers the fair income to j, which is assumed to be independent of the treatment manipulation. Therefore, if the solution to the problem above is interior, the optimal allocation will be given by the following equation:

$$x_j^* = F_j - \frac{1}{\omega(\hat{\gamma}_j)} \tag{1.2}$$

which indicates that the optimal allocation depends on what i consider fair to give to j and the weight i gives on the fair allocation. Premise 1 is represented under the following assumption:

P1. $\frac{\partial x^*}{\partial \hat{\gamma}_j} > 0$.

This assumption indicates that the weight player i puts on allocating the fair outcome towards j depends positively on how close i feels towards j. This imply that an increase in perceived social proximity $\hat{\gamma}_j$ generates an increase in allocation x_j^*.

1.A.2 Premise 2. Social context affects perceived social proximity

From premise 1, the DM behaves more prosocial to j, if the DM considers j to be an ingroup. However, the DM does not know with certainty whether j is an ingroup (in) or not. Instead, the DM takes into account in her decision the probability of individual j to be an ingroup, denoted by $\gamma_j = f(in|g_j)$, which I define as the social proximity j, and I assume that $f(in|N) > f(in|F)$, as the DM is from the U.S.[7] In my set-up, for a DM to determine how much to give to others, she must assess the perceived social proximity of j within a society where she is exposed to other individuals, where each individual of the society has group membership $g_k \in G_s$.

I consider a DM that might suffer from contrast-driven biases in perceptions in their assessment of perceived social proximity of j, which will depend on the group distribution of the society G_s. The framework draws from Kahneman and Tversky's (1972)

[7]The decision of interpreting social proximity as a probability comes from simplicity in adapting Esponda, Oprea and Yuksel (2023) framework into this setting. The rational resonates with the assumption that individuals care more about people that they consider an ingroup, relative to unknown others (Chen and Li, 2009).

"representativeness heuristic", where a decision-maker can form distorted beliefs about a target group by overweighting its representative types. I incorporate this formally following Bordalo et al. (2016*a*) representativeness measure given by the likelihood ratio: $R(in, g_j, g_{-j}) := \frac{f(in|g_j)}{f(in|g_{-j})}$, which captures how representative is being an ingroup (*in*) to group g_j of the receiver relative to group g_{-j} of the non-receiver.

Thus, the DM's *perceived* social proximity of j will be given in the following way:

$$\hat{\gamma}_j = \kappa f(in|g_j)(R(in, g_j, g_{-j}))^{\alpha} \tag{1.3}$$

where κ is a normalization factor, and $\alpha \geq 0$ is a parameter that reflects to which extent is the DM affected by the bias, where if $\gamma_p = 0$, social context does not distort social proximity. However, if $\gamma_p > 0$, then G_s will distort perceived social proximity. Premise 2 is represented under the following assumption:

P2. $\alpha > 0$.

1.A.3 Hypotheses

Following the premises P1 and P2 and equations (1.2) and (1.3) I can derive the following observations for the effect of social context on the allocation decisions of decision-maker i towards receiver j.

Hypothesis 1.A.1. *If premises P1 and P2 are satisfied, allocation x_j^* from the DM towards a U.S. national receiver in the diverse context is higher than in the homogeneous context.*

Proof. A change from a homogeneous to a diverse context decreases the value of $f(in|g_{-j})$, making being type *in* more representative for group g_j. Given P2, this generates an increase of perceived social proximity $\hat{\gamma}_j$. Given P1, the increase in perceived social proximity generates an increase in x_j^*. □

Hypothesis 1.A.2. *If premises P1 and P2 are satisfied, allocation x_j^* from the DM towards a foreigner receiver in the diverse context is lower than in the homogeneous context.*

Proof. A change from a homogeneous to a diverse context increases the value of $f(in|g_{-j})$, making being type *in* less representative for group g_j. Given P2, this generates an decrease of perceived social proximity $\hat{\gamma}_j$. Given P1, the decrease in perceived social proximity generates an decrease in x_j^*. □

1.A.4 Relative proximity as opposed to absolute proximity

Social categorization theory in social psychology, defined by Turner et al. (1987), argues that individuals might use a *meta-contrast principle* in determining which individual is a member of the ingroup and which is an outgroup. The principle indicates that the individual make use of a rule that maximizes intergroup (across groups) differences and minimizes intragroup (within group) differences. A way to represent this in this conceptual framework of the allocation decisions, is to make the weight an individual puts on being fair towards the receiver as $\omega(\hat{\gamma}_j - \hat{\gamma}_{-j})$ as opposed to $\omega(\hat{\gamma}_j)$.

When incorporating this modification, The results obtained from the data in the experiment can be consistent with the experiment. The new model specification allows for the presence of effects on allocations on both, the ingroup and the outgroup, even in the absence of effects of diversity over the social proximity towards foreigners.

1.B Additional Figures and Tables

Figure A1: Mean level of perceived closeness towards a person from the USA, from China and from Canada in the no-contrast context.

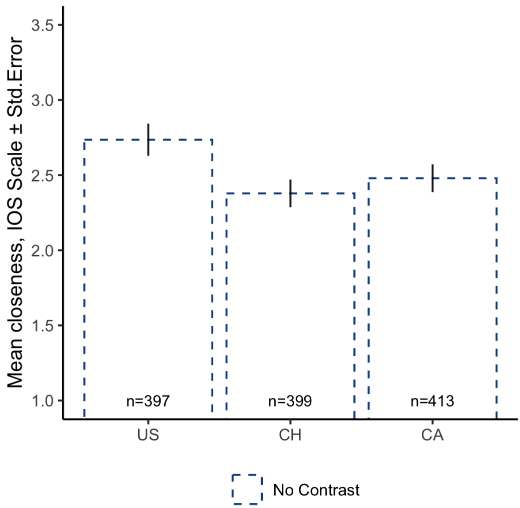


Figure A2: Histograms of the allocation task (left) and the social proximity elicitation (right).

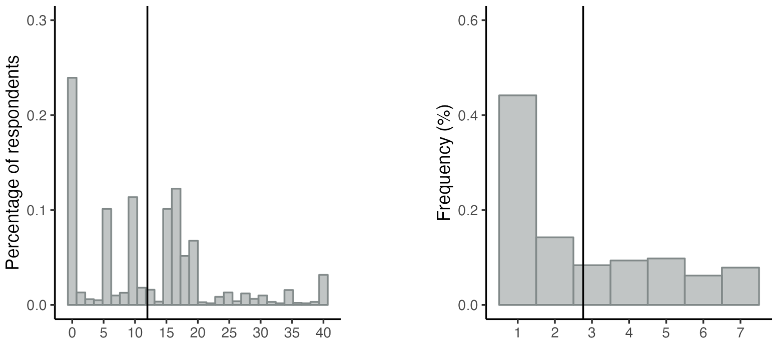
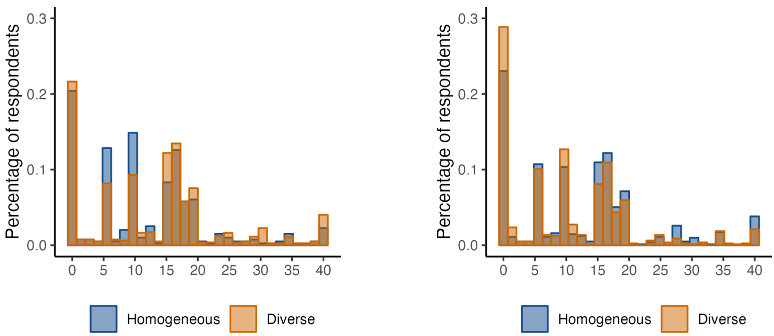


Figure A3: Histogram of the allocations towards U.S. nationals and foreigners in each social contexts: homogeneous and diverse.



(a) U.S. nationals

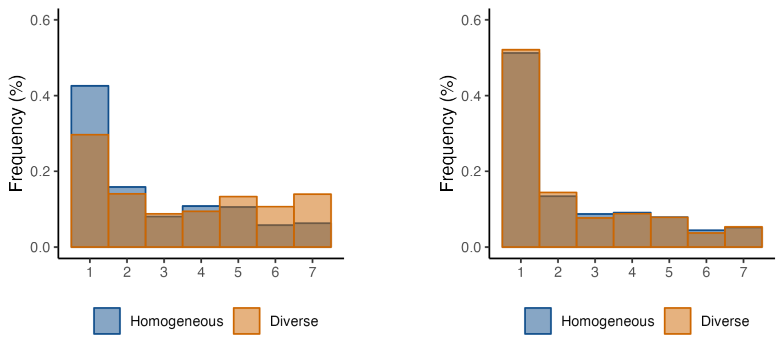
(b) Foreigners

Table A1: Descriptive statistics: number of observations and means across collected outcomes.

	Outgroup: China				Outgroup: Canada		
	Hom.U.S.	Div.U.S.	Div.For	Hom.For.	Div.U.S.	Div.For	Hom.For.
Age	51.65	49.84	50.77	51.05	49.86	50.98	51.34
	(15.08)	(15.46)	(15.16)	(14.54)	(15.23)	(15.07)	(14.94)
Female	0.65	0.62	0.65	0.60	0.63	0.65	0.61
	(0.48)	(0.49)	(0.48)	(0.49)	(0.48)	(0.48)	(0.49)
College	0.68	0.64	0.64	0.66	0.66	0.65	0.69
	(0.47)	(0.48)	(0.48)	(0.48)	(0.47)	(0.48)	(0.46)
White	0.89	0.86	0.86	0.88	0.86	0.87	0.86
	(0.32)	(0.34)	(0.34)	(0.33)	(0.35)	(0.34)	(0.34)
Repub.	0.28	0.28	0.34	0.30	0.32	0.35	0.29
	(0.45)	(0.45)	(0.47)	(0.46)	(0.47)	(0.48)	(0.45)
Altruism	7.81	7.75	7.75	7.54	7.87	7.77	7.79
	(2.50)	(2.32)	(2.48)	(2.56)	(2.35)	(2.52)	(2.37)
Allocation	11.85	13.07	8.96	11.72	13.48	11.86	12.91
	(9.66)	(10.79)	(9.89)	(10.70)	(10.33)	(9.52)	(10.17)
N	397	401	407	399	394	397	413

Notes: The table shows the mean value and standard deviation of each collected variable for each treatment group. Female, College, White and Republican are indicator variables. Age is measured in years. Altruism is a scale from 0 to 10, were 0 is low willigness to give, and 10 is high willingness to give. Finally, Allocation is the amount of USD given to the selected participant, with the choices ranging from 0 to 40 USD.

Figure A4: Histogram of the social proximity towards U.S. nationals and foreigners in each social contexts: homogeneous and diverse.



(a) U.S. nationals (b) Foreigners

Table A2: Effects of exposure to diversity on social proximity under different subsamples and specifications.

	Allocation in USD		Positive giving	
	(1)	(2)	(3)	(4)
Diverse	0.957*	0.971*	−0.012	−0.011
	(0.556)	(0.542)	(0.025)	(0.024)
Foreigner	0.033	0.139	−0.026	−0.021
	(0.550)	(0.536)	(0.025)	(0.025)
Diverse x Foreigner	−2.433***	−2.476***	−0.046	−0.049
	(0.718)	(0.702)	(0.033)	(0.033)
Intercept	11.19	5.07	0.8	0.55
Controls	No	Yes	No	Yes
Observations	2,719	2,719	2,808	2,808
R^2	0.010	0.048	0.006	0.036
Note:		*p<0.1; **p<0.05; ***p<0.01		

Notes: This table presents a robustness analysis of the allocation decision, using an OLS regression with Robust Standard Errors. Columns 1 and 2 show the results restricting the sample to exclude the individuals that made an allocation of 40 USD (all their endowment). Columns 3 and 4 show the regression results over an indicator variable that takes value 1 if the allocation decision was positive. The controls include: gender, age, political party, education, race and state fixed effects.

Table A3: Effects of exposure to diversity on social proximity, and social proximity as a mediator by country, OLS regression with Robust Standard errors.

	Proximity (IOS)		Allocation (USD)		Proximity (IOS)		Allocation (USD)	
	Foreigner: China				Foreigner: Canada			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Diverse	0.928***	0.904***	1.223*	−0.462***	0.610***	0.604	1.631**	0.679
	(0.149)	(0.152)	(0.726)	(0.148)	(0.149)	(0.712)	(0.679)	(0.675)
Foreigner	−0.357***	−0.357***	−0.122	0.527***	−0.256*	−0.267	1.059	1.459**
	(0.136)	(0.136)	(0.723)	(0.136)	(0.137)	(0.698)	(0.657)	(0.664)
Diverse x Foreigner	−1.090***	−1.066***	−3.992***	−2.012***	−0.532***	−0.556	−2.675***	−1.844*
	(0.198)	(0.202)	(1.027)	(0.198)	(0.201)	(0.994)	(0.945)	(0.946)
Proximity				1.817				1.561***
								(0.119)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1,604	1,604	1,604	1,604	1,601	1,601	1,601	1,601
R^2	0.075	0.149	0.021	0.141	0.029	0.107	0.005	0.101

Note: *p<0.1; **p<0.05; ***p<0.01

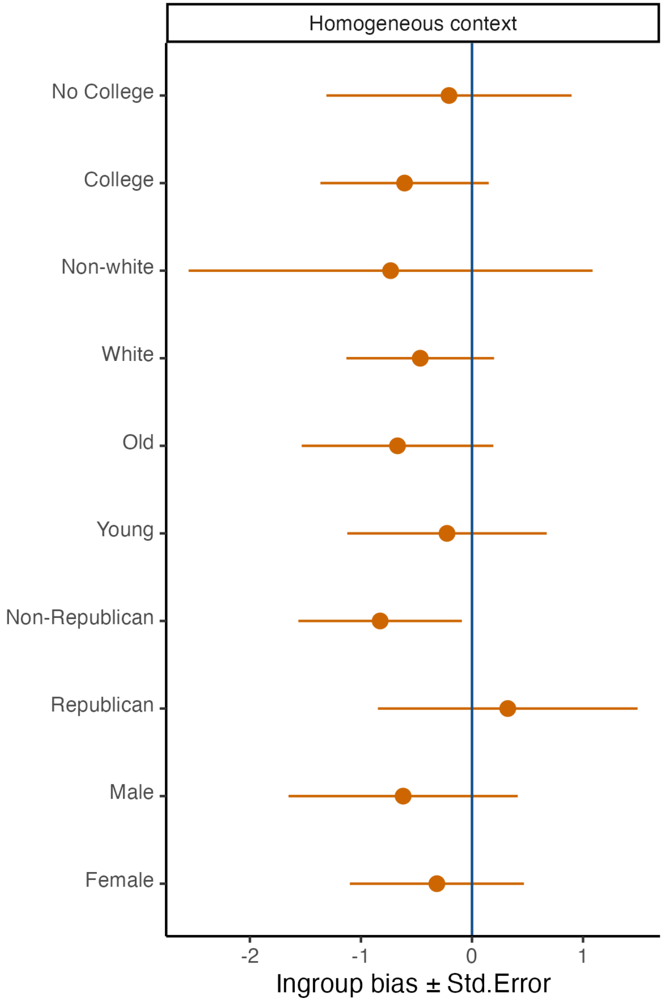
Notes: This table presents a country-wise breakdown of the results on the mechanism, using an OLS regression with Robust Standard Errors. Columns 1 to 4 represent the results focusing on China as the foreign country, and columns 5 to 8 use Canada as the foreign country. Columns 1 to 2 and 5 to 6 show the main specification using Social Proximity (scale 1-7) as the dependent variable. Columns 3 to 4 and 7 to 8 study the role of Proximity as the mechanism over the Allocation decision in USD. The controls include: gender, age, political party, education, race and state fixed effects.

Table A4: Mediation Analysis for social proximity as a mediator in the allocation towards the ingroup decision.

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	1.101	-0.717	1.52	0.000
ADE	0.315	-0.812	1.45	0.566
Total Effect	1.415	0.202	2.59	0.026
Prop. Mediated	0.763	0.374	3.10	0.026

Notes: This table presents a mediation analysis following Imai, Keele and Yamamoto (2010), using the full sample size (2808) over 100 simulations. The analysis focuses on the comparison between ingroup allocation on the homogeneous context and ingroup allocation on the diverse context. ACME corresponds to the average causal mediation effect and ADE corresponds to the average direct effect.

Figure A5: Ingroup bias in the homogeneous context, for each subsample of covariates collected.



Notes: This figure displays the point estimates of an OLS regression of the allocation on an indicator that takes value 1 if the allocation was for the a U.S. national and 0 if it was for the foreigner, for restricting the sample of participants facing a homogenous context, for each subsample of covariates collected. I show the point estimates $\pm$ Robust S.E., for the regression without controls.

Table A5: The additional treatment effect of diversity on the allocation by respondents who first indicated Social Proximity.

	Allocation in USD		
	Pooled	For.: China	For.: Canada
	(1)	(2)	(3)
Diverse	1.150	1.360	0.898
	(0.853)	(0.994)	(0.989)
Second	−0.055	−0.055	−0.055
	(0.973)	(0.973)	(0.973)
Foreigner	0.687	0.094	1.224
	(0.857)	(1.031)	(0.976)
Diverse x Second	0.560	−0.313	1.383
	(1.229)	(1.460)	(1.422)
Foreigner x Second	−0.416	−0.416	−0.342
	(1.219)	(1.450)	(1.399)
Diverse x Foreigner	−2.851**	−4.532***	−1.303
	(1.113)	(1.431)	(1.392)
Diverse x Foreigner x Second	−1.043	1.085	−2.812
	(1.590)	(2.062)	(1.982)
Intercept	11.88	11.88	11.88
Controls	No	Yes	No
Observations	2,808	1,604	1,601
R^2	0.013	0.022	0.008
Note:	*p<0.1; **p<0.05; ***p<0.01		

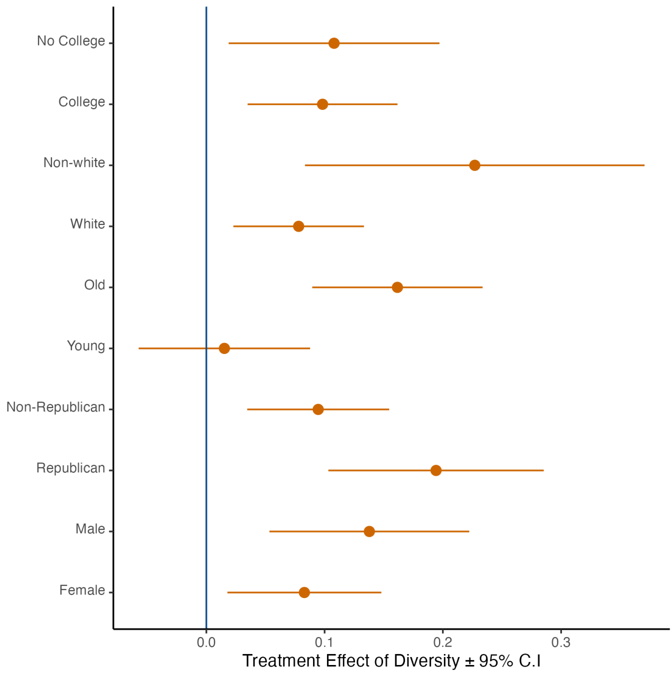
Notes: This table studies whether the order of presentation of the allocation decision matters, using an OLS regression with Robust Standard Errors. The table follows the main regression specification, with Allocation (in USD) as the dependent variable, and the interaction with an indicator variable that takes value 1 if allocation decision followed the social proximity elicitation. Column 1 show the results using the pooled foreign countries, and Column 2 and 3 separates the analysis for China and Canada respectively.

Table A6: The additional treatment effect of diversity on the social proximity by respondents who first indicated the allocation decision.

	Social proximity (IOS)		
	Pooled	For.: China	For.: Canada
	(1)	(2)	(3)
Diverse	1.190***	1.559***	0.872***
	(0.173)	(0.207)	(0.199)
Second	0.139	0.139	0.139
	(0.198)	(0.198)	(0.198)
Foreigner	−0.239	−0.176	−0.305*
	(0.163)	(0.193)	(0.182)
Diverse x Second	−0.830***	−1.163***	−0.555*
	(0.252)	(0.295)	(0.296)
Foreigner x Second	−0.137	−0.377	0.087
	(0.238)	(0.271)	(0.271)
Diverse x Foreigner	−1.150***	−1.745***	−0.573**
	(0.217)	(0.281)	(0.271)
Diverse x Foreigner x Second	0.660**	1.208***	0.134
	(0.313)	(0.393)	(0.397)
Intercept	2.67	2.67	2.67
Controls	No	Yes	No
Observations	2,808	1,604	1,601
R^2	0.064	0.092	0.033
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: This table studies whether the order of presentation of the social proximity elicitation matters, using an OLS regression with Robust Standard Errors. The table follows the main regression specification with Proximity (scale 1 to 7) as the dependent variable, and the interaction with an indicator variable that takes value 1 if social proximity elicitation followed the allocation decision. Column 1 show the results using the pooled foreign countries, and Column 2 and 3 separates the analysis for China and Canada respectively. The controls include: gender, age, political party, education, race and state fixed effects.

Figure A6: Effect of exposure to diversity on support for national welfare payments, for each subsample of covariates collected.



Notes: This figure displays the point estimates of an OLS regression of the agreement level on a dummy that takes value one if the participant was exposed to diversity, for each subsample of covariates collected. Old and Young correspond to above or below 50 years old (median age). I show the point estimates $\pm$ Robust S.E., for the regression without controls.

Table A7: Effects of exposure to diversity on policy views under different specifications, Ordered Probit and OLS regression with Robust Standard errors.

	Prioritize Nationwide Welfare payments					
	Average agreement			Percentage that agrees		
		ordered probit			*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)
Diverse	0.075^{*}	0.087^{**}	0.054	0.039^{**}	0.043^{**}	0.029
	(0.040)	(0.040)	(0.041)	(0.019)	(0.018)	(0.018)
Male		0.150^{***}	0.113^{***}		0.104^{***}	0.089^{***}
		(0.042)	(0.042)		(0.019)	(0.019)
Republican		-0.541^{***}	-0.547^{***}		-0.169^{***}	-0.168^{***}
		(0.044)	(0.045)		(0.020)	(0.019)
Older than 50		-0.257^{***}	-0.214^{***}		-0.085^{***}	-0.066^{***}
		(0.041)	(0.041)		(0.018)	(0.018)
College		-0.014	-0.018		0.031	0.029
		(0.042)	(0.043)		(0.019)	(0.019)
Proximity			0.086^{***}			0.034^{***}
			(0.010)			(0.004)
Intercept	0.08	0.09	0.05	0.36	0.23	0.16
Controls	No	Yes	Yes	No	Yes	Yes
Observations	2,808	2,808	2,808	2,808	2,808	2,808
R^2				0.002	0.063	0.083

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Notes: This table presents a robustness analysis of the effects of social context on policy preferences, using an OLS regression with Robust Standard Errors. Columns 1 and 2 show the results using an ordered probit, given the setting with an ordered categorical variable. Columns 3 and 4 show the regression results over an indicator variable that takes value 1 if the participant indicated agreement.

1.C Survey Questionnaire

Figure A7: Consent.

Welcome to this research project! We very much appreciate your participation.

Procedures
The study consists of a decision that might affect your final payoffs and a series of questions. Please make sure to always read the instructions carefully and answer truthfully.

Participation
Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to future participation in other studies conducted by us.

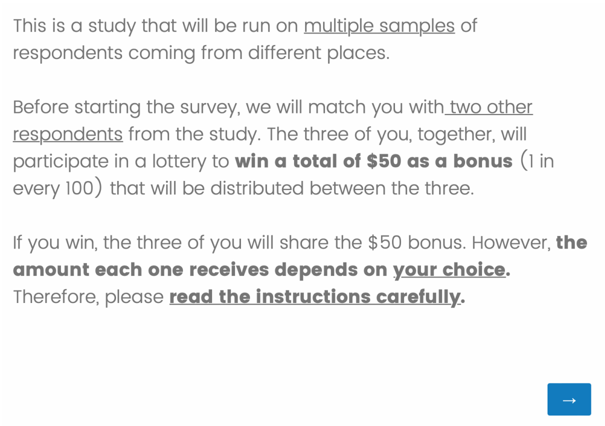
Confidentiality
Your answers to the questions will be anonymous. That is, your name or any information that could identify you will not be asked or attached to your answers. Data may be shared in anonymized form in open science repositories.

Payment
For completing this survey, you will be automatically enrolled in a lottery by the research team where you can win up to 40 USD (equivalent in points), where 1 in every 100 participants will be selected.

Questions about the Research
If you have questions regarding this study, you may contact: thechoicelab@nhh.no

Consent
Please write 'ACCEPT' in the box below if you have understand the above and wish to participate in this study.

Figure A8: Page 1.



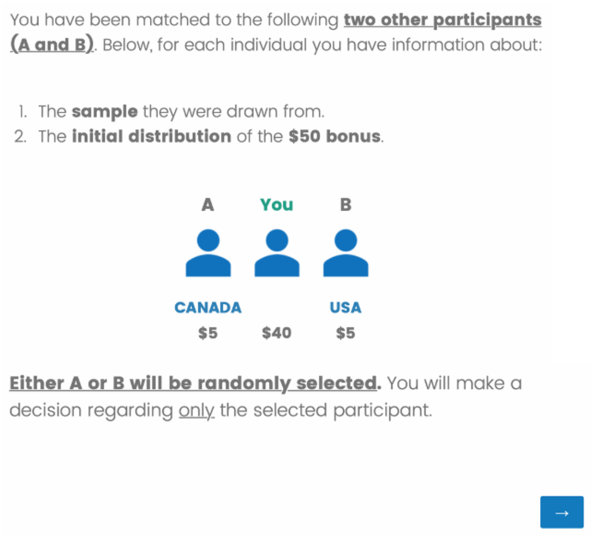
This is a study that will be run on multiple samples of respondents coming from different places.

Before starting the survey, we will match you with two other respondents from the study. The three of you, together, will participate in a lottery to **win a total of $50 as a bonus** (1 in every 100) that will be distributed between the three.

If you win, the three of you will share the $50 bonus. However, **the amount each one receives depends on your choice.** Therefore, please **read the instructions carefully.**

→

Figure A9: Page 2.



You have been matched to the following **two other participants (A and B)**. Below, for each individual you have information about:

1. The **sample** they were drawn from.
2. The **initial distribution** of the **$50 bonus**.

Either A or B will be randomly selected. You will make a decision regarding only the selected participant.

→

Figure A10: Page 3.

Figure A11: Closeness elicitation

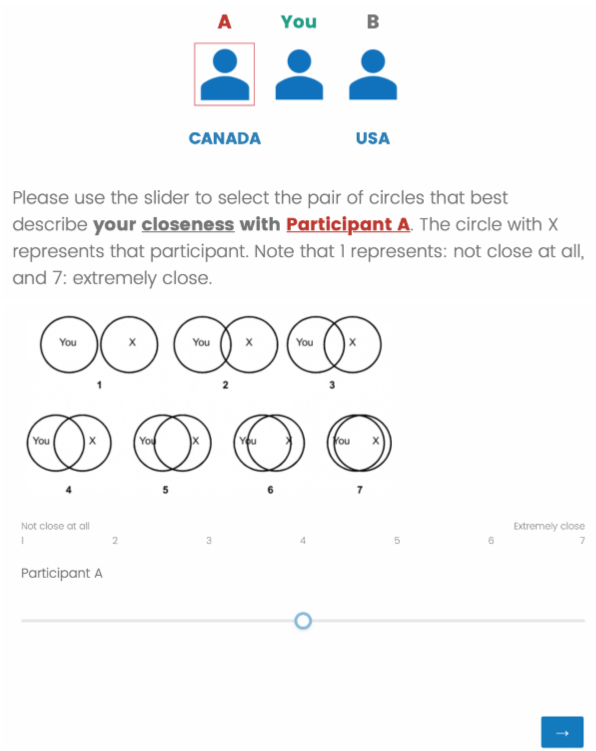
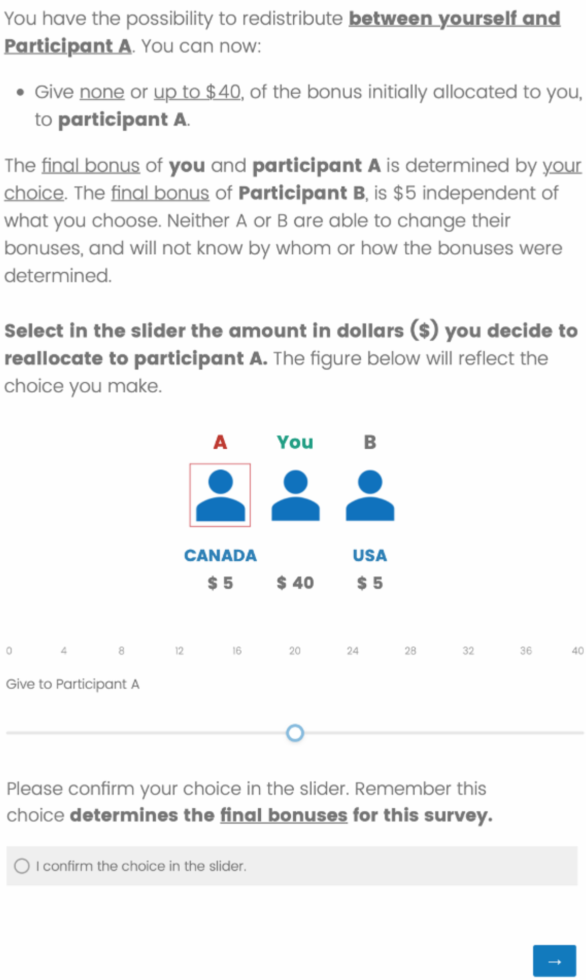


Figure A12: Allocation Decision



You have the possibility to redistribute **between yourself and Participant A**. You can now:

- Give none or up to $40, of the bonus initially allocated to you, to **participant A**.

The final bonus of **you** and **participant A** is determined by your choice. The final bonus of **Participant B**, is $5 independent of what you choose. Neither A or B are able to change their bonuses, and will not know by whom or how the bonuses were determined.

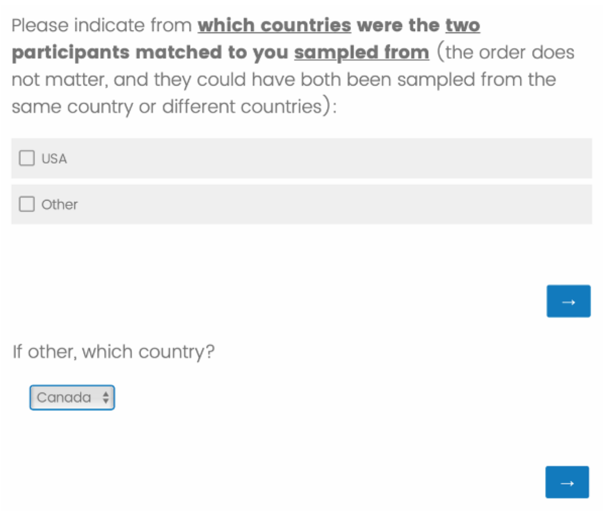
Select in the slider the amount in dollars ($) you decide to reallocate to participant A. The figure below will reflect the choice you make.

Please confirm your choice in the slider. Remember this choice **determines the final bonuses for this survey.**

I confirm the choice in the slider.

→

Figure A13: Attention check



Please indicate from **which countries were the two participants matched to you sampled from** (the order does not matter, and they could have both been sampled from the same country or different countries):

- [ ] USA
- [ ] Other

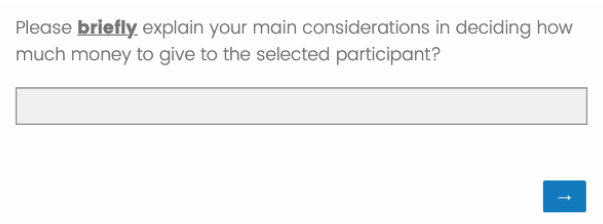
→

If other, which country?

Canada

→

Figure A14: Open-ended question



Please **briefly** explain your main considerations in deciding how much money to give to the selected participant?

→

Figure A15: Policy views

You are now going to read a number of statements. In each case, we want you to say whether you Strongly Agree, Somewhat Agree, Do not agree or disagree, Somewhat Disagree, Strongly Disagree. **Read carefully each statement**.

The government should use local tax revenues to fund health insurance only within the local communities they were raised, **rather** than to fund health insurance across all communities nationwide.

- Strongly agree
- Somewhat agree
- Do not agree or disagree
- Somewhat disagree
- Strongly disagree

The government should redistribute local tax revenues as welfare payments across all communities nationwide, **rather** than only within the local communities they were raised.

[options]

The government should send foreign aid to countries that are our international allies, **rather** than to countries that are in most need of help.

[options]

The government should take measures to ensure no child of the world is disadvantaged in access to education, the labor force, and marriage, **rather** than focusing on American children.

- Strongly agree
- Somewhat agree
- Do not agree or disagree
- Somewhat disagree
- Strongly disagree

→

Figure A16: Altruism elicitation

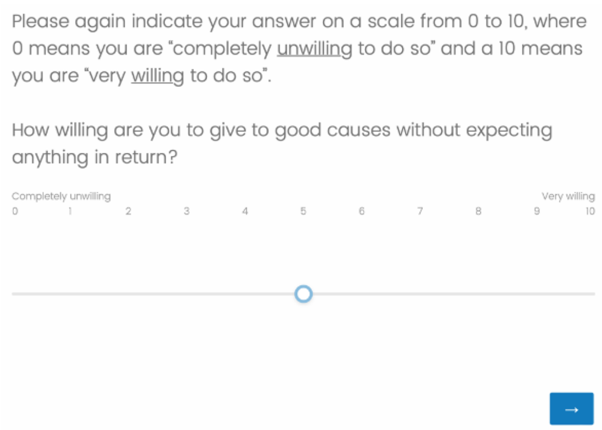


Figure A17: Background demographic information. Part 1.

Background questionnaire

Please answer the following questionnaire. You must answer these questions truthfully.

What is your age?

Which party did you support in the last U.S. elections?

- Republican
- Democrat
- Independent
- I didn't support any party

Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else?

- Democrat
- Independent
- Republican
- Other

What is the highest level of school you have completed or the highest degree you have received?

- High school graduate or less
- Bachelor's degree in college (4-year)
- Master's degree
- Doctoral degree
- Professional degree (JD, MD)

Figure A18: Background demographic information. Part 2.

In which state do you currently reside?

Which race or ethnicity best describes you?

- ☐ White or Caucasian
- ☐ Black or African American
- ☐ American Indian/Native American
- ☐ Hispanic / Latina(o)
- ☐ Asian
- ☐ Other
- ☐ Prefer not to say

Chapter 2

The role of a majority-minority status and ingroup affinity in shaping social preferences

Abstract: Globalization, conflict, and climate change are generating unprecedented demographic shifts which affect the distribution of groups in society and therefore the majority-minority status of individuals. This study investigates the impact of such status on social preferences, focusing on preferences for giving and acceptance of inequality. Using an online sample of 1600 U.S. participants, I randomly assign participants to roles as members of a majority or minority in a social context. The findings reveal that majority-minority status does not significantly affect these social preferences. However, the proportion of participants with ingroup affinity, defined as whether the participant feels closer to the ingroup relative to the outgroup, is slightly higher in a minority context than a majority context. Finally, an exploratory analysis suggests that ingroup affinity plays a role in the presence of ingroup bias in social preferences.

I am grateful to my supervisors Bertil Tungodden and Alexander Cappelen for invaluable feedback in the development and writing of this project. This paper also benefited from helpful comments by Akshay Moorthy. Funding for this project was provided by the ERC Advanced Grant 788433. This study was preregistered in the AEA RCT Registry (AEARCTR-0010179).

2.1 Introduction

Globalization, conflict, and climate change are leading to unprecedented movement of people from different groups (e.g., nationality, ethnicity, etc.) across and within countries. These significant demographic shifts are transforming individuals' immediate social environments, including the neighborhoods they live in, and the classrooms and workplaces they attend. Notably, these changes alter the relative sizes of groups in a given social setting, and therefore people's majority-minority status. An individual could form part of a majority, when there are more people that share the individual's group membership than others, or a minority, when there are more people from a different group, relative to the individual's group members. Such status could significantly influence how individuals perceive their surroundings and interact with others.

A substantial body of interdisciplinary research has explored the relationship between the size of migrant groups and the attitudes of natives toward them (Pottie-Sherman and Wilkes, 2017), towards other sets of migrants (Fouka and Tabellini, 2022), hate crimes (Cikara, Fouka and Tabellini, 2022), and homophily (Karimi et al., 2018). Much of this discussion, stemming from "Group threat theory" in sociology, suggests that an increase in the size of a minority group can lead to more negative attitudes from the majority (Blalock Jr, 1967; Schlueter and Scheepers, 2010), though the relationship is not always unambiguous (Citrin and Sides, 2008; Hjerm, 2007). While existing evidence primarily focuses on the macro perspective of diversity, using national-level migration scenarios, this study centers on a micro-level exposure to social contexts that individuals encounter in their daily lives, and how it shapes their interactions with others. This paper investigates how an individual's majority-minority status in an immediate social context influences their social preferences, finding a limited influence of such status on behavior towards others.

Causally estimating the impact of majority-minority status on behavior is challenging, as this status is often endogenous. Differences in group compositions across social contexts might stem from specific context characteristics, e.g., a rich neighborhood might be more homogeneous. Moreover, there can be selection concerns in the exposure of individuals to a social context where they are a majority or a minority. This paper addresses this challenge by experimentally varying exposure to different social contexts, wherein a decision-maker is randomly assigned a context where she has a majority or minority status.

This paper studies decisions in three incentivized allocation games, using a controlled experiment in a sample of around 1600 participants from the U.S., recruited through Prolific. Participants were randomly placed in a social context where they assumed roles as members of either a minority or a majority. Three findings emerge. First, the majority-minority status of participants does not significantly influence their allocation decisions, suggesting a limited role in social preferences. Second, while perceptions of social proximity do not vary significantly across different social contexts, a minor shift is observed in ingroup affinity, defined by whether an individual feels closer to a member of the ingroup than the outgroup when participants transitioned from a majority to a minority status. Finally, ingroup affinity emerges as a crucial factor in driving ingroup biases in social preferences.

The experimental design follows Carvajal (2024) by exogenously manipulating social context. A decision-maker from the U.S. is randomly assigned to a social context composed of five matched individuals, which includes fellow U.S. nationals (ingroup) and participants from China (outgroup), varying in group size. The decision-maker can have either a (i) majority status, where she is matched with four U.S. nationals and one participant from China, or a (ii) minority status, where she is matched with one U.S. national and four participants from China. The objective is to examine social preferences across these two contexts, measured using three allocation tasks: a dictator game, where the participant is a stakeholder, and two where the participant is a spectator.

In the first allocation decision, the dictator game, a decision-maker decides which amount to give from their endowment towards a randomly selected receiver from the matched social context. This introduces a second layer of randomization, corresponding to whether the receiver is a fellow U.S. national (ingroup) or a participant from China (outgroup). The design allows to examine the effects of social context towards both: ingroup and outgroup, as well as the presence of ingroup bias in allocations. The findings suggest that neither the allocations towards an individual from a specific country, nor the ingroup bias in allocations is affected by the social context. Also, the ingroup bias in allocations in both social contexts is positive, though weak.

The two allocation tasks as spectators are collected to assess inequality acceptance, where participants are presented with an unequal distribution of resources and have the option to equalize outcomes by destroying a portion of the advantaged individual's money. Each participant faces two types of situations: first, a *local inequality*, where resources are distributed unequally between two people from the same country. Second, a *global inequality*, where resources are distributed unequally between two individuals from different countries. The findings are threefold. On one hand, around 60% of the participants accept an inequality when it is local, regardless of the country of both participants. On the other hand, when the inequality is global, participants accept 10-15 p.p. more inequality when it benefits the ingroup (65%) relative to when it benefits the outgroup (50-55%). Notably, the findings indicate that prior exposure to majority-minority status does not significantly influence inequality acceptance in either the local or the global inequality scenarios.

Altogether, the results on the behavioral measures of social preferences indicate that the majority-minority status of an individual seems not to play a strong role in social preferences in this setting.

Findings on recent work that studies the effect of social context on prosocial behavior show that perceptions of social proximity play an important role in allocation decisions (Carvajal, 2024; Robson, 2021). Therefore, to understand the main findings on social preferences, this study also examines perceived social proximity. In both types of social contexts to which the participants are exposed, there are members of the ingroup and of the outgroup. Thus, I am able to measure the *ingroup affinity* of an individual, defined as the perception that a member of the ingroup is closer than a member of the outgroup, and study how it varies according to the size of each group, which has not been previously studied. Using the Inclusion of the Other in the Self (IOS) scale developed by Aron, Aron and Smollan (1992), participants indicate their closeness to participants from both the ingroup and the outgroup. This yields an ingroup affinity measure, where individuals present ingroup affinity when they indicate closer proximity to their ingroup than to the outgroup.

The results are twofold: while individual social proximity levels towards both ingroup and outgroup remain constant across contexts, the proportion of individuals with ingroup affinity marginally increases when participants have a minority status. However, this change is modest, around 7 p.p. compared to the 54% of individuals with ingroup affinity in the majority context and seems not to significantly influence allocations. Notably, the average ingroup bias in proximity, given by the difference between the proximity to the ingroup and to the outgroup, is already positive and high in both contexts, where the DM feels 30% closer to an ingroup than an outgroup (with a base of 2 in the 7-point scale), limiting the scope of impact of social context on perceptions and behavior, suggesting an explanation for the lack of effects in the allocation decision.

An interesting finding is that the percentage of individuals with ingroup affinity, within a context, is stable regardless of whether the receiver in the allocation task is an ingroup or an outgroup, making ingroup affinity balanced across both situations. Therefore, it enables an exploratory analysis studying heterogeneity in allocation decisions based on ingroup affinity, which can play a key role.

First, regarding the dictator game, among individuals with ingroup affinity, a notable ingroup bias in allocations exists, with decision-makers in this group allocating 1 USD more to their ingroup than to the outgroup, relative to the average giving of 6.3 USD. Conversely, individuals without ingroup affinity exhibit no such ingroup bias, highlighting the role of ingroup affinity as a key driver in dictator game allocation patterns. Second, inequality acceptance also varies by types. For local inequalities, both participants with and without ingroup affinity exhibit similar levels of acceptance, around 60%, regardless of the group membership of the two individuals (ingroup or outgroup). In contrast, acceptance of global inequalities reveals differing patterns. Individuals without ingroup affinity accept inequality consistently, at around 60% level regardless of who is favored by the inequality, the ingroup or the outgroup. In contrast, individuals with ingroup affinity accept 70% of inequalities favoring their ingroup, compared to only 50% when benefiting the outgroup. These findings suggest that relative social proximity within a context not only influences allocation decisions but also modifies inequality acceptance, favoring the ingroup.

In summary, while the majority-minority status of an individual appears to have a limited role in perceptions and prosocial behavior towards others in a society, perceptions of social proximity, particularly ingroup affinity, significantly explain ingroup biases in social preferences.

This paper contributes to our understanding in three main strands of literature. First, it expands our knowledge about the impact of outgroup size on interactions with others (Blalock Jr, 1967; Schlueter and Scheepers, 2010; Fouka and Tabellini, 2022). Unlike previous studies predominantly focused on the migrant setting, which examine how the size of migrant groups influences attitudes towards them using self-reported measures (Pottie-Sherman and Wilkes, 2017), or reports of hate crimes towards members of migrant groups (Cikara, Fouka and Tabellini, 2022), this research makes use of three incentivized behavioral measures of social preferences. It also shifts the focus from macro-level group sizes at the national level to the influence of majority-minority status in immediate, small-scale social contexts, such as those in workplaces and classrooms. Importantly, this paper expands the scope of study from solely examining the effect of outgroup size on behavior towards the outgroup to also include its impact on behavior towards the ingroup, an aspect that has received limited attention.

Additionally, this paper intersects with research on group identity and self-categorization in social preferences (Tajfel and Turner, 1979; Akerlof and Kranton, 2000; Charness and Chen, 2020). Previous studies have demonstrated the influence of perceived social proximity on prosocial behavior towards specific individuals (Carvajal, 2024; Robson, 2021; Bicchieri et al., 2022). This research extends these findings by considering ingroup affinity within a social context, by comparing perceptions towards members of the ingroup relative to the outgroup, a factor that has not been adequately explored previously. The results highlight the significant role of ingroup affinity in shaping social preferences.

Finally, the paper contributes to our understanding of the determinants of social preferences, particularly in terms of preferences for giving and inequality aversion (Fehr and Schmidt, 1999; Charness and Rabin, 2002; Cappelen et al., 2007). While most prior research has focused on how individuals accept inequality between individuals with indistinct characteristics or without explicit specification of group membership, this paper provides insights into how giving and inequality acceptance can be influenced by salient group membership, often in a direction favoring the ingroup, in accordance with results obtained by (Chen and Li, 2009). Moreover, it underscores that not all individuals are equally sensitive to group membership in their behavior, but the differences generated by group membership are more pronounced among individuals displaying ingroup affinity.

The remainder of the paper is structured as follows. In Section 2.2, I explain the social preferences measures used in this experiment and propose the design that allows me to identify the effects of minority-majority status on these measures. Section 2.3 defines the empirical strategy. In Section 2.4, the main results are presented, as well as suggestive evidence of the role of ingroup affinity in defining social preferences. Finally, in Section 2.5, I discuss the potential extensions of the paper and future directions.

2.2 Experimental Design

I first define social context in the experimental setting and its manipulation. Second, I describe the three allocation decisions of the study: one dictator game, as a stakeholder, and two inequality acceptance games, as a spectator. Then, I describe the social proximity elicitation and other collected outcomes. Finally, I summarize the sample and the procedures.

Social context: Majority-Minority Status. A decision-maker (DM) is informed that they are participating in an international study and that they will be matched with five other individuals. The DM is always from the U.S., and the matched participants are a combination of participants from the U.S. (ingroup) and from China (outgroup). The DM can be exposed to one out of two possible social contexts, which differ in the size of each of the groups, and therefore, affect the minority-majority status of the DM in this setting. First, in the *majority* context, the DM is a member of a majority, where the five matched participants consist of four fellow U.S. nationals and one participant from China. Second, in the *minority* context, the DM is a member of a minority, where the five matched participants consist of one fellow U.S. national and four participants from China. Each DM will face only one social context.

Dictator game. The DM and the five matched participants, together, share a total of 65 USD, initially distributed as follows: 40 USD for the DM, and 5 USD for each of the five matched participants. After informing the DM about the initial distribution of the 65 USD and the country of each of the matched participants, one of the matched participants is randomly selected to be a *receiver*.[1] The DM is informed about the selection process. Now, the final payoffs are determined as follows: the DM decides how much of their initial endowment of 40 USD would they allocate only towards the receiver. The participants that were not selected, the non-receivers, each will have a final payoff of 5 USD as initially determined. Note that the final pay-offs of the DM and the receiver depend on the DM's choice, and the pay-offs of the non-receivers are independent of the choice.

Importantly, the DMs are informed that, in their group, only they are able to change the initial distribution of the pay-offs, and that the decision is anonymous. This aims to avoid potential signaling and social image concerns in determining the allocation decision (Andreoni and Bernheim, 2009).

Inequality acceptance games. Following exposure to a social context, the DM makes two additional allocation decisions as a spectator. In each decision, the DM is matched with two new participants (e.g., A and B). A total of 5 USD is initially distributed in an unequal way: 3 USD for A and 2 USD for B. The DM must choose between two options: (1) reduce the bonus of A to 2 USD to equalize bonuses, or (2) keep the inequality by not changing the initial distribution. I measure inequality acceptance in two settings. First, a *local inequality*, where the DM makes the spectator decision while matched with two new participants from the same

[1]The probabilities for any of the matched participants of being selected are not equal, but all participants have a probability higher than $\frac{1}{5}$ of being selected.

country (either ingroup or outgroup). Second, a *global inequality*, where the DM makes the spectator decision while matched with two new participants, one from the USA (ingroup) and one from China (outgroup). The order of which type of inequality situation is presented first, local or global, is randomized. As described before, each situation presented to the participant involves an initial inequality.

The matched participants with the advantageous inequality in both tasks come from the same country as the country of the randomly selected, for a given DM. Therefore, for the local inequality, the participant can face either two members of the ingroup or of the outgroup, and for the global inequality, they face one member of each group, where the inequality can be advantageous towards either the ingroup or the outgroup.

The aim of this task is to eliminate self-regarding considerations, as the DM is not a stakeholder, and isolate potential inequality aversion considerations in a decision where an inequality between two other individuals exists. In particular, I am interested in whether the group membership of the two individuals with an existing inequality matters for their preferences for equality.

Social proximity elicitation and ingroup affinity. Aimed to understand the effect of social context on allocations, I obtain a proxy measure for perceived social proximity using the Inclusion of Other in the Self (IOS) scale, developed by Aron, Aron and Smollan (1992), where I ask the respondent to indicate the closeness they feel towards a participant from a specific group, with options guided by two circles with different levels of overlap. No overlap means not close at all, and the bigger the overlap, the closer the respondent feels towards the specific participant. The measure has been validated as a reliable measure for perceived social proximity in comparison to other more sophisticated survey methods (Gächter, Starmer and Tufano, 2015). Moreover, it has been widely incorporated in recent research in economics and psychology (Goette and Tripodi, 2021; Bicchieri et al., 2022; Gächter, Starmer and Tufano, 2022; Carvajal, 2024).

As the respondents will be facing social contexts with participants from both: the ingroup (U.S.) and the outgroup (China), I ask for the DMs perceived social proximity towards participants from both groups. This allows me to construct for each respondent a summary measure of proximity towards others within a social context: *ingroup affinity*, defined by whether an individual feels closer to a member of the ingroup relative to a member of the outgroup. I construct then an indicator measure, where an individual has ingroup affinity if the perceived social proximity towards a fellow U.S. national is strictly higher than the perceived social proximity towards a person from China. This measure will allow us to study the role of ingroup affinity in preferences for giving and inequality aversion.

Demographic outcomes. In the survey, key demographic information such as education, political affiliation, ethnicity, sex, country of birth, country of residence, age.

Figure 2.1 illustrates two examples of social contexts a DM can be exposed to. The DM is matched to five other participants and informed about their nationality, as well as the initial distribution of the 65 USD. In red, it is indicated which of the five participants was randomly selected to be the receiver. Each figure represents the two social contexts a DM can be exposed to. In Figure 2.1a, the DM is a member of a majority, meaning that the DM from the U.S. is matched with four other participants from the ingroup (U.S.) and one participant from the outgroup (China). On the other hand, in Figure 2.1b, the DM is a member of a minority, meaning that the DM from the U.S. is matched with one member of the ingroup and four members of the outgroup. Note that in both situations presented a participant from the U.S. was selected, thus, the DM will make an allocation decision towards an ingroup member. Subsequently, the DM will indicate their perceived social proximity towards both, an ingroup member and an outgroup member. Therefore, in both situations, the DM faces the exact same choice sets, the only difference corresponds to the context, given by the four non-receivers; which is a key feature of the design.

2.2.1 Treatment conditions

Figure 2.2 provides an overview of the experiment. For the study of the role of majority-minority status on social preferences, the experiment follows a 2x2 between-subject experimental design, where respondents are randomized across two dimensions. First, one out of two potential social contexts is randomly selected where the DM is either (i) a member of a majority, or (ii) a member of a minority. Second, whether the allocation decisions are beneficial towards an ingroup member (fellow U.S. national), or an outgroup member (from China). This means in the dictator game the randomly selected receiver was either ingroup or outgroup, and in the inequality acceptance games, the advantageous inequality favors a person with the same group membership as the receiver.

The timeline of the experiment is as follows: the DMs are informed about the social context they are in, and subsequently, they are indicated who will be the receiver in the dictator game. After completing the dictator game, I elicit the perceived social proximity towards both, participants from the ingroup and the outgroup.

The order of whether the social proximity towards an ingroup member or an outgroup member in a context is asked first depends on the identity of the receiver. If the receiver is ingroup (outgroup), then the DM is asked first about proximity towards the ingroup (outgroup). Following the exposure to a social context and the first two tasks, the DMs face the two inequality acceptance games: one in which they face a local inequality, and one in which they face the global inequality. Whether the participant faces first the local or the global inequality situation is randomized. Note that for each type of inequality acceptance situation (local or global), the DM faces one out of two variations. If the DM in the allocation decision was randomly matched to allocate to a participant from the ingroup (outgroup), then in both inequality acceptance tasks the DM faces inequalities that are advantageous towards a member of the ingroup (outgroup), i.e. in the local inequality the DM faces two individuals of the ingroup (outgroup) and in the global inequality the DM faces an initial inequality where the ingroup member has more than the outgroup member.

2.2.2 Sample and Procedures

The survey was conducted online with a sample of 1600 participants from Prolific. Participants from the U.S. were recruited between July 14th and July 16th, 2023. The average response duration was around 4 minutes. The experimental design was preregistered at the AEA RCT Registry (AEARCTR-0010179).[2] The full instructions are made available in Appendix 2.C.

Table A1 in Appendix 2.A provides a summary of the outcomes collected in the experiment. The sample consists of around 50% of

[2]Deviations from the pre-analysis plan are specified in section 2.B.

women, around 20% identifying as Republican, a mean age of 40, mostly white (80%), and a majority college educated (60%), and 40% with income before taxes higher than 75000 USD. I observe a successful randomization, where individual characteristics are balanced across treatments.

2.3 Empirical strategy

The main goal of this paper is to study the role of majority-minority status on social preferences, measured using three allocation decisions: the dictator game, the local inequality acceptance game, and the global inequality acceptance game. The paper runs the following specification for the analysis of the main outcomes, where the dependent variable y_i corresponds to the allocation decision in each of the three allocation tasks, the social proximity towards the receiver, and ingroup affinity:

$$y_i = \beta_0 + \beta_1 Minority_i + \beta_2 Outgroup_i + \beta_3 Minority_i \times Outgroup_i + X_i + \varepsilon_i \quad (2.1)$$

where $Minority_i$ is an indicator variable which takes value 1 if the DM has a minority status in the social context and 0 otherwise. $Outgroup_i$ is an indicator variable, defined as follows: when y_i is the allocation in the dictator game, $Outgroup_i$ takes value 1 if the allocation is towards a member of the outgroup and 0 otherwise. When the dependent variable is the allocation in the local or global inequality acceptance game, it takes value 1 if the advantageous inequality favors a member of the outgroup, and 0 otherwise. Therefore, variable $Outgroup_i$ will be referred to as the indicator variable that takes value 1 if the allocation decision benefits an outgroup member, and 0 otherwise. Finally, X_i is the set of collected demographic characteristics.

When the dependent variable is the allocation in USD in the dictator game, the coefficient for the variable $Outgroup_i$ (β_2) estimates the difference between the mean allocation towards the outgroup, relative to the ingroup, when the DM has a majority status. For the case in which the dependent variable is inequality acceptance, the coefficient estimates the difference in the percentage of participants that accept inequality when it benefits the outgroup relative to when it benefits the ingroup, when the DM has a majority status. Therefore, both correspond to measures of ingroup bias. A negative coefficient suggests a positive ingroup bias.

To detect the effects of a change in social context on the allocation decisions, I focus on two coefficients. First, the coefficient for variable $Minority_i$ (β_1), which provides the differences in the mean allocation towards the ingroup in the dictator game, or the difference in the percentage of participants that accept inequality when it benefits the ingroup, in the minority context relative to the majority context. Second, the coefficient of the interaction term $Minority_i \times Outgroup_i$ (β_3), which corresponds to a differences-in-differences coefficient that estimates the change in the ingroup bias in allocations or inequality acceptance, as defined above, from the majority context to the minority context. A negative coefficient suggests an increase in ingroup bias.

The experimental design allows for a heterogeneity analysis over the allocation decisions within each context by ingroup affinity of participants, as will be expanded in section 2.4. With this aim, I estimate the following equation for each game and each context with the allocation in the game as the dependent variable

y_i:

$$y_i = \beta_0 + \beta_1 Affinity_i + \beta_2 Outgroup_i + \beta_3 Affinity_i \times Outgroup_i + X_i + \varepsilon_i \quad (2.2)$$

where $Affinity_i$ is an indicator variable which takes value 1 if the DM exhibits ingroup affinity, meaning that for the DM the elicited social proximity towards the ingroup is strictly higher than the proximity towards the outgroup, and 0 otherwise. $Outgroup_i$ is the same indicator variable as defined in equation 2.1. Finally, X_i is the set of collected demographic characteristics.

As with the main analysis, I study the three allocation decisions: the dictator game, the local inequality acceptance, and the global inequality acceptance. I focus on two coefficients. First, the coefficient for variable $Outgroup_i$ (β_1), which provides the difference in allocations or inequality acceptance benefiting the outgroup relative to the ingroup, when the DM does not exhibit ingroup affinity. In other words, it captures the ingroup bias for participants without ingroup affinity, where a negative coefficient corresponds to a positive ingroup bias. Second, to detect differences in behavior by the presence of ingroup affinity, I study the coefficient of the interaction term $Affinity_i \times Outgroup_i$ (β_3), which gives the difference in the ingroup bias in allocations or inequality acceptance between DMs with and without ingroup affinity.

2.4 Results

This section presents an analysis of the experimental findings. First, I examine the effects of majority-minority status on social preferences, measured through the allocation decisions. To understand the findings on the allocations, I further explore whether majority-minority status affects perceptions of social proximity and ingroup affinity. Finally, the experimental design allows an exploratory analysis that studies heterogeneity in allocation decisions by in-group affinity of the decision-makers, to understand its role in social preferences.

2.4.1 Majority-minority status and social preferences

Participants in the study give on average 6.3 USD out of 40 USD to the selected participant, with around 60% choosing to allocate a positive amount (see Figure A1a of Appendix 2.A). The level of giving is notably lower than most studies on giving as well as similar previous experiments with three participants (Engel, 2011; Carvajal, 2024). However, the low giving in a social context of six participants is consistent with work by Andreoni and Bernheim (2009) that shows that with an increased group size giving is reduced. In terms of the allocations in the inequality acceptance game, the proportion of participants that accept inequality for both scenarios, the global and the local inequality, is 40%. The proportion of participants that equalize pay-offs is comparable to similar spectator games in previous experiments (Charness and Rabin, 2002; Cappelen et al., 2007). Moreover, this is consistent with other experiments showing that participants are eager to "burn money" (Zizzo and Oswald, 2001).

I now examine the effects of majority-minority status on the allocation decisions. Figure 2.3a shows the mean allocation in USD in the dictator game towards a member of the ingroup, and a member of the outgroup across contexts: when the DM is a member of a majority and when the DM is a member of a minority. Examining the effect of majority-minority status on allocations, no significant differences are found in mean allocations towards an ingroup member or an outgroup member across contexts ($p>0.5$), indicating that a DM's majority-minority status does not influence their allocation behavior. In both contexts, the mean allocation from the DM towards an ingroup member is higher than the allocation towards an outgroup member, suggesting a positive ingroup bias in allocations. The difference in means is not significant in either context ($p>0.1$), and there is no difference in ingroup bias across contexts ($p>0.5$). When analyzing the difference in allocations towards the ingroup and the outgroup with the pooled sample, including observations from both contexts, the difference becomes significant at the 5% level.

I further study whether the exposure to a majority-minority status affects a DM's acceptance of local and global inequalities. First, Figure 2.3b shows the percentage of participants that accept the local inequality, when the inequality is between members of the ingroup (orange) and when the inequality is between two members of the outgroup (blue), across social contexts. When the inequality is local, 60% of participants equalize. The proportion is the same regardless of majority-minority status and whether the inequality is between members of the ingroup or the outgroup. Second, Figure 2.3c shows the proportion of people accepting inequality when the inequality is global, meaning an inequality between members of different groups. The proportion is calculated for the sample where the advantageous inequality was favoring the ingroup (orange) and when the advantageous inequality was favoring the outgroup (blue), across contexts. First, note that around 65% of DMs accept the inequality when the ingroup is favored, however, this percentage is the same regardless of the majority-minority status of the DM ($p>0.1$). Interestingly, the acceptance of inequality when it favors a participant from the outgroup is significantly lower than when it favors an ingroup member in both contexts at around 50-55% ($p=0.02$), and not significantly different by majority-minority status ($p>0.1$).[3] The lack of difference in local or global inequality across contexts constitutes evidence that when a DM is exposed to a setting in which they were part of a majority or a minority, subsequent inequality acceptance is not affected.

The results on social preferences are summarized in the regression Table 2.1, that displays the estimated the coefficients following specification (2.1). Columns 1 and 2 show the estimates with allocation in USD in the dictator game as the dependent variable. Columns 3 to 6 represent the results for the allocation decisions as a spectator, by showing the estimates of specification (2.1) using as a dependent variable an indicator variable that takes value 1 if the DM accepted the inequality and 0 if not. Columns 3 and 4 correspond to the local inequality situation, and columns 5 and 6 correspond to the global inequality situation. The regression results corroborate the findings, where neither the coefficient on Minority, showing the effect of majority-minority status on decisions benefiting the ingroup, nor the interaction term Outgroup $\times$ Minority, showing the effect of majority-minority on the ingroup biases in allocation decisions, are significantly different from zero for any specification using the different dependent variables: allocation in the dictator game, local or global inequality acceptance. The results persist for each depend variable even after including control variables (Columns 2, 4 and 6). This allows me to conclude

[3]The patterns are the same even if I consider only the first inequality acceptance game presented (see Figures A2a and A2b in Appendix 2.A).

that the status of the DM as a majority or a minority in a social context has no impact in social preferences in the setting studied here.

Result 1. *The majority-minority status of individuals has no impact on their social preferences, as it does not affect allocation decisions benefiting the ingroup nor the ingroup bias in allocation decisions, regardless of whether the decision-maker is a stakeholder (dictator game) or an spectator (local and global inequality acceptance).*

2.4.2 Majority-minority status and perceived social proximity

To understand the findings in the allocation decisions, I study the role of perceived social proximity on social preferences.

The average elicited social proximity using the IOS Scale was of 2.37 units in a 7-points scale, and around 37% of decision-makers consider the selected participant as "not close at all", which is an expected outcome given the anonymity among matched participants. In both social contexts presented, there were participants from both: the ingroup and the outgroup. Around 97% of participants have either no difference or a positive difference in proximity of ingroup versus outgroup (see Figure A1c in Appendix 2.A). This attest to the validity of the measure, where it is expect that participants feel closer to their ingroup than to the outgroup.

I now examine the effect of majority-minority status on perceptions of social proximity towards others. I focus on two measures: perceived social proximity towards the receiver, and ingroup affinity. A DM has ingroup affinity when they perceive a member of the ingroup as closer than a member of the outgroup within a specific context.

Figure 2.4a shows the mean perceived social proximity towards a receiver who is a member of the ingroup (in orange) and towards a receiver who is a member of the outgroup (in blue), across both contexts. The majority-minority status of the DM does not affect significantly the perceived social proximity towards a receiver, neither when the receiver is from the ingroup, nor from the outgroup ($p>0.4$). At the same time, a DM on average feels closer to an ingroup than to an outgroup, by around 30% (under a base level of 2 in a 7 points scale), showcasing an ingroup bias in proximity towards receivers (difference with $p<0.01$).The bias does not differ by majority-minority status of the DM.

Table 2.2 confirms the findings. Column 1 and 2 show the regression estimates of equation (2.1), using perceived social proximity towards the receiver as the dependent variable. The coefficient of variable Minority shows us that the proximity towards a receiver from the ingroup is not significantly different in the minority context, relative the majority context. Moreover, the coefficient of the interaction term Minority $\times$ Outgroup shows that majority-minority status does not affect the ingroup bias in social proximity, defined as the difference between perceived social proximity towards a receiver when the receiver is an ingroup member relative to when the receiver is from the outgroup.

The presence of a strong ingroup bias in social proximity in the majority context, where the DM feels 30% closer to the ingroup than the outgroup, leaves little room for changes in social context to increase the ingroup bias in social proximity when moving towards a minority context, which is corroborated with the lack of differences in the bias across contexts. The negligible impact of majority-minority status on proximity and social preferences, in this experimental setting, is consistent with previous work suggesting that an important mechanism of the effects of social context on prosocial behavior corresponds to changes in perceived social proximity generated by contextual factors (Carvajal, 2024).

I also explore the effect of majority-minority status on ingroup affinity. Figure 2.4b shows the proportion of participants that exhibit ingroup affinity. The proportion of participants with ingroup affinity increases by 5 p.p. in the minority context from 54% in the majority context.

Result 2. *The majority-minority status of an individual does not affect the perceived social proximity towards individuals; however, the share of individuals with ingroup affinity is higher in settings where the individual is a member of a minority, relative to when she is part of a majority.*

Discussion. It is important to note that within each context, whether the DM was making allocation decisions that benefit members of the ingroup or of the outgroup has no significant effect on the proportion of participants that exhibit ingroup affinity. This is shown in columns 3 and 4 of Table 2.2 that show the regression estimates of equation (2.1), using an indicator variable for the presence of ingroup affinity as the dependent variable. Coefficient Outgroup shows the difference in the percentage of participants with ingroup affinity if the participants face allocations benefiting the outgroup relative to if they faced allocation decisions benefiting the ingroup, for the subsample in the Majority context. The sum of the coefficients Outgroup and the interaction term Outgroup $\times$ Minority shows the same difference, but for the subsample in the Minority context. Note that both coefficients are not significantly different from 0 ($p>0.15$). This condition allows for the study of heterogeneity in allocation decisions within contexts, by ingroup affinity, as the proportion of participants with ingroup affinity is balanced across groups.

2.4.3 Ingroup Affinity and Social Preferences

I explore a potential additional determinant in the effect of social context on social preferences. Allocations have been shown to be affected by the perceived social proximity towards a receiver i in a social context (Carvajal, 2024). However, proximity towards other individuals present in a social context, that are different than receiver i, might also affect a DM's behavior towards individual i, an aspect that has been overlooked. This suggests that not only the *absolute* level of proximity towards a receiver matters, but also the *relative* level of proximity towards a receiver, in comparison with others in a society. In my experiment, relative proximity is given by ingroup affinity. This section explores heterogeneity in allocation decisions by ingroup affinity.

Figure 2.5 summarizes the analysis by ingroup affinity for the three collected measures of social preferences: allocation decision, local inequality acceptance, and global inequality acceptance; for each of the social contexts: with majority and minority status.

First, I focus on how allocations vary by individuals with and without ingroup affinity. Figures 2.5a and 2.5b show, for each context, the mean allocation of the DM to an ingroup member (orange circle) and to an outgroup member (blue triangle) for each type: when the DM has no ingroup affinity versus when the DM has ingroup affinity. In both contexts, when DMs have no ingroup affinity, they do not exhibit ingroup bias; if anything, they allocate more to the outgroup, though the difference is not significantly different from zero. On the other hand, when the DM displays ingroup affinity, there is a significant ingroup bias of around 1 USD, for a mean allocation in the whole sample of around 6.3 USD, in both contexts. This suggests that the ingroup bias in allocations is entirely driven by individuals with ingroup affinity.

Second, I study heterogeneity in inequality acceptance. Figures 2.5e and 2.5f depict, for each context, the proportion of individuals that accept a local inequality (between two people from the same country), meaning, the individuals that do not equalize the outcomes in the burning money game. There are two cases, when the local inequality is within ingroup members (orange circle) and when it is within outgroup members (blue triangle) for each type: when the DM has no ingroup affinity versus when the DM has ingroup affinity. The proportion that accepts the inequality is around 60% in any situation, regardless of whether the DM has ingroup affinity, or whether it is between two ingroup members or two outgroup members.

Lastly, Figures 2.5c and 2.5d display, for each context, the proportion of individuals that accept a global inequality (between people from different countries), meaning, the individuals that do not equalize the outcomes in the burning money game. There are two cases, when the global inequality is advantageous for the ingroup member (orange circle) and when it is advantageous for the outgroup member (blue triangle) for each type: when the DM has no ingroup affinity versus when the DM has ingroup affinity. For the individuals with no ingroup affinity, the fraction that accepts inequality is the same, at around 50-55%, regardless of whether the ingroup or the outgroup has an advantage. However, for the individuals with ingroup affinity, the proportion that accepts the inequality is more than 20 p.p. higher when the inequality is advantageous for the ingroup than when it is advantageous for the outgroup, in both contexts.

The results are confirmed through a regression analysis, which follows equation (2.2). Tables 2.3 and 2.4 show the results for each of the contexts: majority and minority, respectively. Columns 1 and 2 correspond to the estimates using the allocation decision in USD as the dependent variable, whereas Columns 3-6 use the proportion of individuals that accept an inequality as the dependent variable. The regression results confirm that for the individuals without ingroup affinity, there is no ingroup bias in any of the measures of social preferences, given by the coefficient Outgroup. Moreover, for the allocation decision (columns 1 and 2) and the global inequality acceptance (columns 5 and 6), the ingroup bias increases by 3.3 USD and 19.9 p.p. respectively, resulting in a significantly positive ingroup bias in both cases ($p<0.01$). In the case of the local inequality, neither participants with ingroup affinity nor participants without ingroup affinity exhibit any ingroup bias, as local inequalities are accepted at the same rate.

Altogether, these results show that not only the proximity of a DM towards an individual affects a DM's behavior towards that individual, but also the proximity of others that surround them in a social context. In particular, I provide evidence of ingroup affinity as a key factor in the development of ingroup biases in social preferences.

Result 3. *Ingroup affinity plays a role in explaining the heterogeneity in preferences for giving and acceptance of inequality, where individuals with ingroup affinity are the ones that drive any ingroup bias in social preferences.*

2.5 Conclusion

This paper studies the effects of majority-minority status on allocation decisions, inequality acceptance, and perceived social proximity. The experimental design, through a controlled setting with online participants from the U.S. and China, causally estimates the differences between decision-makers who are randomly exposed to a social context where they have either majority or minority status.

The key findings of this study are threefold. First, the majority-minority status of participants does not significantly influence their allocation decisions or acceptance of inequality, suggesting that the majority-minority status of an individual has a limited role in shaping their preferences. Second, while perceptions of social proximity do not vary significantly across different social contexts, a minor shift is observed in ingroup affinity, defined by whether an individual feels closer to a member of the ingroup than the outgroup, when participants transition from a majority to a minority status. Finally, ingroup affinity emerges as a crucial factor in driving ingroup biases in social preferences.

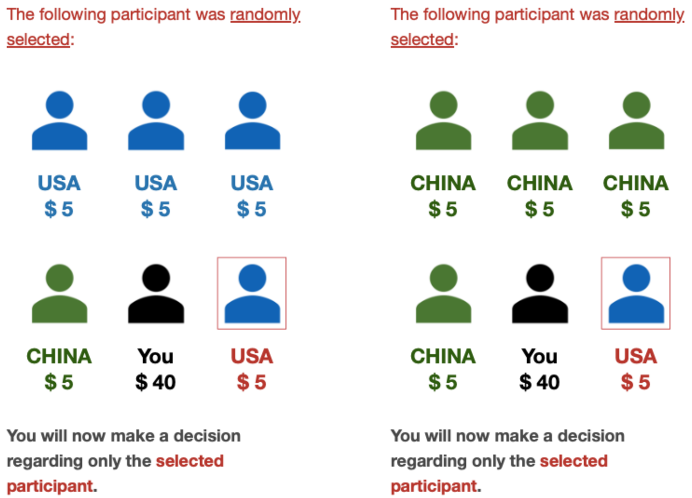
This study is not without limitations. The experimental setting, while controlled and systematic, may not fully capture the complexity and nuances of real-world social interactions. Field evidence and studies considering different sets of characteristics that people might encounter, such as ethnicity, language, and others, are needed.

This research opens several avenues for future investigation. One key area is the exploration of dynamics and transitions from a status quo of a majority to one of minority and vice versa. Additionally, further research could delve into the way ingroup affinity affects social preferences, exploring its implications in various social and organizational settings.

In conclusion, this paper adds to the existing literature by highlighting the role of social context in shaping social preferences and social identity. The findings underscore the importance of considering ingroup affinity in understanding human behavior in diverse social settings.

2.6 Main Figures

Figure 2.1: Example of a change in social context from majority status (left) to minority (right) in the allocation toward the ingroup in the dictator game.



(a) Treatment **Majority - Ingroup** (b) Treatment: **Minority - Ingroup**

Figure 2.2: Overview of the experiment

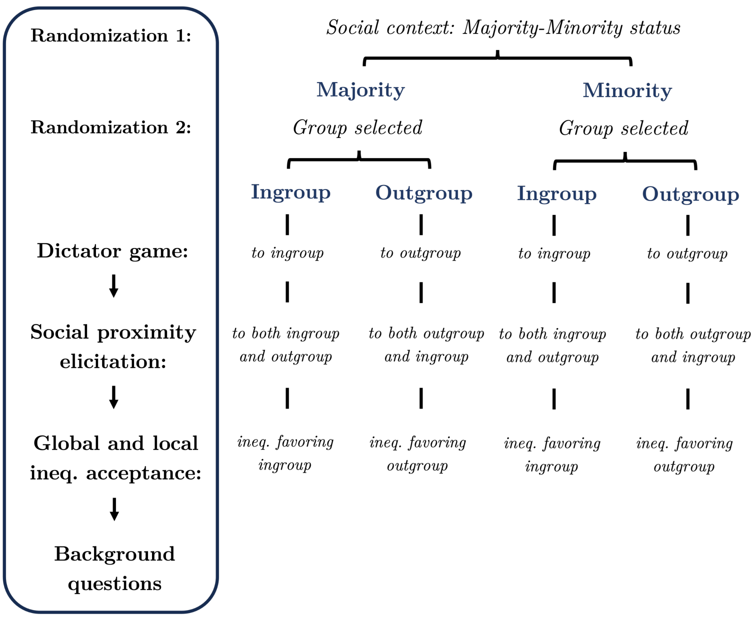
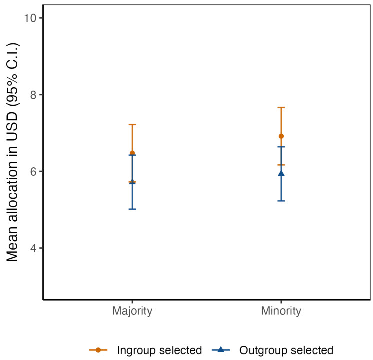
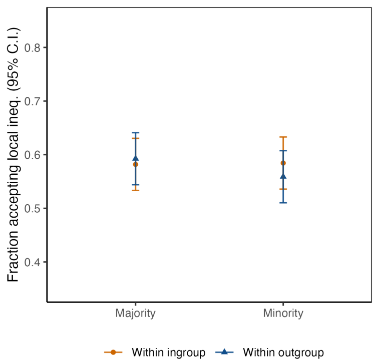


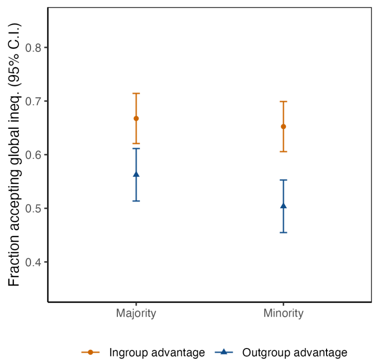
Figure 2.3: Effect of majority-minority status on social preferences



(a) Allocation in dictator game (USD)



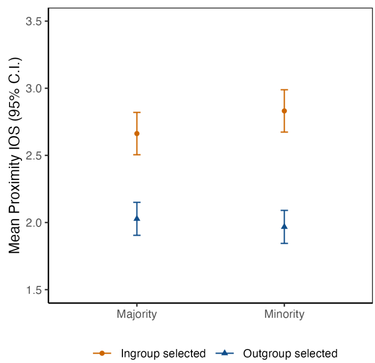
(b) Local inequality



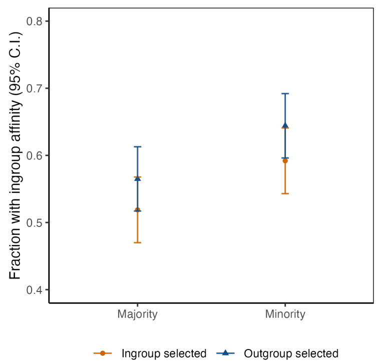
(c) Global inequality

Notes: Figure (a) shows a the mean allocation in USD (out of 40 USD) in the dictator game towards a receiver from the ingroup/outgroup and in the majority/minority context. Panel (b) plots, the percentage of decision-makers that accepted the inequality in the local inequality, when the two participants are from the ingroup/outgroup and in the majority/minority context. Panel (c) displays the fraction of decision-makers that accepted the inequality in the global inequality, when the two participants are one from the ingroup and one from the outgroup, but the inequality is advantageous for the ingroup/outgroup and in the majority/minority context.

Figure 2.4: Effects of majority-minority status on social proximity and ingroup affinity



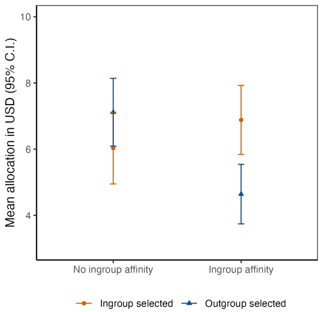
(a) Absolute proximity



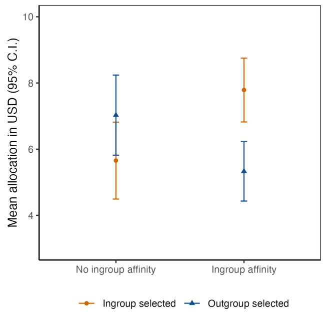
(b) Ingroup affinity

Notes: Figure (a) shows a the mean social proximity (IOS Scale from 1-7) towards a receiver from the ingroup/outgroup and in the majority/minority context, and 95% confidence intervals. Panel (b) plots, the percentage of decision-makers that exhibit ingroup affinity, meaning that they indicated that they feel strictly closer to the ingroup than to the outgroup, for the case when the receiver was from the ingroup/outgroup and in the majority/minority context.

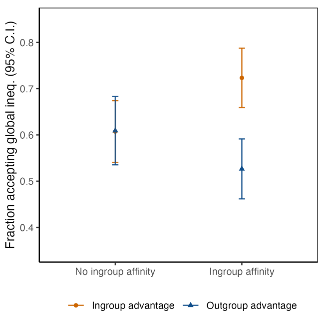
Figure 2.5: Heterogeneity in social preferences by ingroup affinity



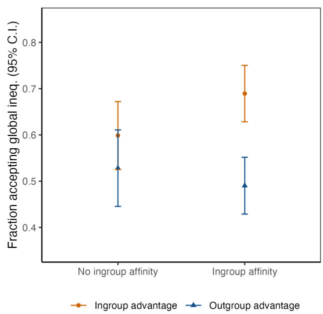
(a) Majority: Dictator



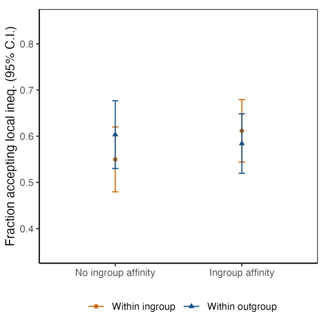
(b) Minority: Dictator



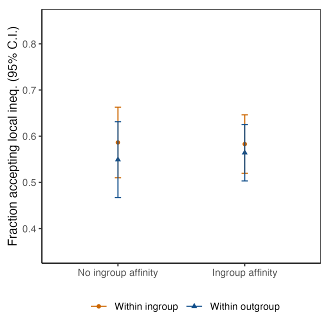
(c) Majority: Global Ineq.



(d) Minority: Global Ineq.



(e) Majority: Local Ineq.



(f) Minority: Local Ineq.

Notes: The graphs show, for each context separately, the estimates using the subsamples of participants with/without ingroup affinity, and when the allocation benefits a member of the ingroup/outgroup.

2.7 Main Tables

Table 2.1: OLS estimates of the regression on the social preferences measures.

	Dictator USD		Fraction accepts local ineq.		Fraction accepts globql ineq.	
	(1)	(2)	(3)	(4)	(5)	(6)
Minority	0.443	0.328	0.003	0.015	−0.015	0.006
	(0.539)	(0.587)	(0.035)	(0.038)	(0.035)	(0.038)
Outgroup	−0.756	−0.803	0.011	0.019	-0.105^{***}	-0.096^{**}
	(0.517)	(0.567)	(0.035)	(0.038)	(0.035)	(0.038)
Minority x Outgroup	−0.226	0.011	−0.036	−0.053	−0.044	−0.068
	(0.741)	(0.813)	(0.050)	(0.053)	(0.050)	(0.053)
Intercept	6.47	4.82	0.58	0.55	0.67	0.59
Controls	No	Yes	No	Yes	No	Yes
Observations	1,593	1,380	1,593	1,380	1,593	1,380
R^2	0.004	0.054	0.001	0.074	0.019	0.080

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table presents the results from an OLS regression with robust S.E. Columns 1 and 2 shows the estimates using the allocation decision (in USD) in the dictator game as the dependent variable. Column 3 to 6 estimate the coefficients using as the dependent variable an indicator variable which takes value 1 if the participant accepted the inequality, and 0 otherwise. Columns 3 and 4 correspond to the local inequality situation whereas columns 5 and 6 correspond to the global inequality situation. The variable Minority takes value 1 when the participant is in the Minority-status context and 0 otherwise. The variable Outgroup in columns 1 and 2 takes value 1 when the allocation is towards a member of the outgroup or 0 otherwise, whereas in columns 3 to 6 it takes value 1 when the inequality is advantageous towards a member of the outgroup, and 0 othwewise. Minority x Outgroup shows the estimate of the interaction. The even columns show specifications that control for all covariates measured, except for the elicited social proximity. The controls include: gender, age, political party, education, race and state fixed effects.

Table 2.2: OLS estimates of the regression on perceived social proximity and ingroup affinity.

	Social Proximity receveiver		Frac. ingroup affinity	
	(1)	(2)	(3)	(4)
Minority	0.169	0.092	0.073^{**}	0.036
	(0.114)	(0.121)	(0.035)	(0.038)
Outgroup	-0.635^{***}	-0.716^{***}	0.046	0.014
	(0.103)	(0.110)	(0.035)	(0.038)
Minority x Outgroup	-0.229	-0.135	0.006	0.039
	(0.144)	(0.155)	(0.049)	(0.053)
Intercept	2.66	2.42	0.52	0.4
Controls	No	Yes	No	Yes
Observations	1,593	1,380	1,593	1,380
R^2	0.066	0.128	0.008	0.066

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table presents the results from an OLS regression with robust S.E. Columns 1 and 2 shows the estimates using the perceived social proximity towards a receiver (IOS scale 1-7 points) as the dependent variable. Column 4 and 3 estimate the coefficients using as the dependent variable an indicator variable which takes value 1 if the participanthas ingroup affinity, and 0 otherwise. The variable Minority takes value 1 when the participant is in the Minority-status context and 0 otherwise. The variable Outgroup takes value 1 when the allocation is towards a member of the outgroup or 0 otherwise. The even columns show specifications that control for all covariates measured, except for the elicited social proximity. The controls include: gender, age, political party, education, race and state fixed effects.

Table 2.3: OLS estimates of the regression on social preferences, by ingroup affinity for the subsample in the context Majority.

	Dictator USD		Fraction accepts local ineq.		Fraction accepts globql ineq.	
	(1)	(2)	(3)	(4)	(5)	(6)
Affinity	0.852	1.218	0.062	0.076	0.116^{**}	0.119^{**}
	(0.771)	(0.865)	(0.050)	(0.056)	(0.050)	(0.056)
Outgroup	1.084	1.390	0.054	0.085	0.002	0.024
	(0.832)	(0.934)	(0.052)	(0.059)	(0.052)	(0.059)
Affinity x Outgroup	-3.325^{***}	-3.847^{***}	-0.081	-0.135^{*}	-0.199^{***}	-0.232^{***}
	(1.047)	(1.174)	(0.070)	(0.079)	(0.070)	(0.079)
Intercept	6.03	2.5	0.55	0.37	0.61	0.34
Controls	No	Yes	No	Yes	No	Yes
Observations	797	696	797	696	797	696
R^2	0.019	0.108	0.002	0.110	0.022	0.105

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table presents the results from an OLS regression with robust S.E. for the subsample in the Majority status context. Columns 1 and 2 shows the estimates using the allocation decision (in USD) as the dependent variable. Column 3 to 6 estimate the coefficients using as the dependent variable an indicator variable which takes value 1 if the participant accepted the inequality, and 0 otherwise. Columns 3 and 4 correspond to the local inequality situation whereas columns 5 and 6 correspond to the global inequality situation. The variable Affinity takes value 1 when the participant exhibits ingroup affinity and 0 otherwise. The variable Outgroup in columns 1 and 2 takes value 1 when the allocation is towards a member of the outgroup or 0 otherwise, whereas in columns 3 to 6 it takes value 1 when the inequality is advantageous towards a member of the outgroup, and 0 othwewise. Affinity x Outgroup shows the estimate of the interaction. The even columns show specifications that control for all covariates measured, except for the elicited social proximity. The controls include: gender, age, political party, education, race and state fixed effects.

Table 2.4: OLS estimates of the regression on social preferences, by ingroup affinity for the subsample in the context Minority.

	Dictator USD		Fraction accepts local ineq.		Fraction accepts globql ineq.	
	(1)	(2)	(3)	(4)	(5)	(6)
Affinity	2.133^{***}	2.037^{**}	-0.003	-0.004	0.091^{*}	0.086
	(0.780)	(0.915)	(0.050)	(0.055)	(0.050)	(0.055)
Outgroup	1.374	1.725^{*}	-0.037	-0.043	-0.071	-0.069
	(0.908)	(1.030)	(0.057)	(0.061)	(0.057)	(0.061)
Affinity x Outgroup	-3.830^{***}	-3.750^{***}	0.018	0.001	-0.128^{*}	-0.158^{**}
	(1.113)	(1.267)	(0.073)	(0.079)	(0.073)	(0.079)
Intercept	5.65	5.42	0.59	0.76	0.6	0.79
Controls	No	Yes	No	Yes	No	Yes
Observations	796	684	796	684	796	684
R^2	0.020	0.107	0.001	0.123	0.027	0.138

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table presents the results from an OLS regression with robust S.E. for the subsample in the Minority status context. Columns 1 and 2 shows the estimates using the allocation decision (in USD) in the dictator game as the dependent variable. Column 3 to 6 estimate the coefficients using as the dependent variable an indicator variable which takes value 1 if the participant accepted the inequality, and 0 otherwise. Columns 3 and 4 correspond to the local inequality situation whereas columns 5 and 6 correspond to the global inequality situation. The variable Affinity takes value 1 when the participant exhibits ingroup affinity and 0 otherwise. The variable Outgroup in columns 1 and 2 takes value 1 when the allocation is towards a member of the outgroup or 0 otherwise, whereas in columns 3 to 6 it takes value 1 when the inequality is advantageous towards a member of the outgroup, and 0 othwewise. Affinity x Outgroup shows the estimate of the interaction. The even columns show specifications that control for all covariates measured, except for the elicited social proximity. The controls include: gender, age, political party, education, race and state fixed effects.

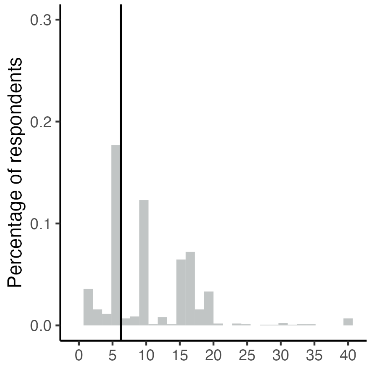
Appendices

2.A Additional Figures and Tables

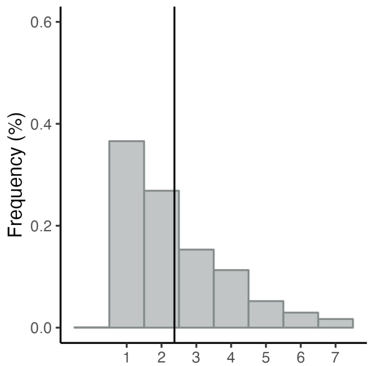
Table A1: Summary statistics of covariates by treatment.

	Majority				Minority			
	Ingroup		Outgroup		Ingroup		Outgroup	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	40.62	13.46	39.53	13.39	41.23	46.49	41.80	47.63
Female	0.50	0.50	0.46	0.50	0.51	0.50	0.52	0.50
College	0.58	0.49	0.60	0.49	0.61	0.49	0.60	0.49
White	0.80	0.40	0.81	0.39	0.77	0.42	0.82	0.39
Republican	0.21	0.41	0.18	0.38	0.20	0.40	0.19	0.39
Income > 75k	0.38	0.49	0.42	0.49	0.38	0.49	0.43	0.50

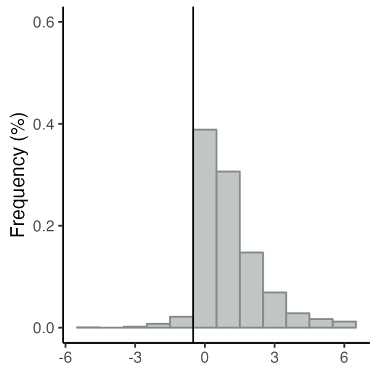
Figure A1: Histograms of the full sample for the allocations in the dictator game, the social proximity towards the receiver, and the difference between social proximity towards the ingroup and the outgroup.



(a) Allocation in dictator game (USD)



(b) Proximity towards the receiver



(c) Proximity towards Ingroup – Outgroup

Notes: In Figure (a) and (b) the vertical line corresponds to the mean value of the distribution. In Figure (c) the vertical line partition the distribution between negative and non-negative values..

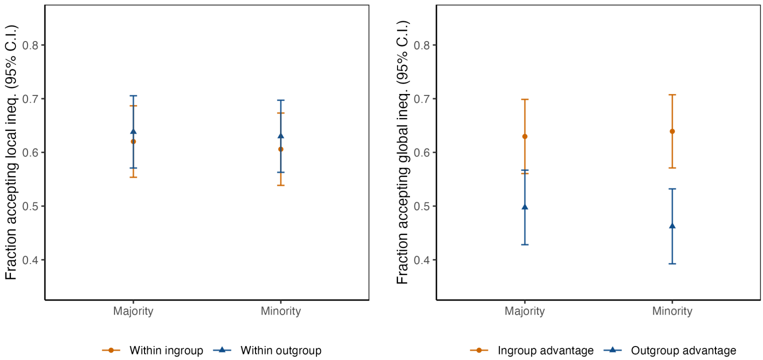
Table A2: OLS estimates of the regression on the alternative specifications for the dictator game.

	Allocation in USD		Positive giving	
	(1)	(2)	(3)	(4)
Minority	0.703	0.506	0.013	−0.014
	(0.495)	(0.543)	(0.034)	(0.037)
Outgroup	−0.414	−0.501	−0.050	−0.060
	(0.477)	(0.526)	(0.035)	(0.038)
Minority x Outgroup	−0.658	−0.398	−0.019	0.010
	(0.690)	(0.761)	(0.049)	(0.054)
Intercept	6.05	4.06	0.62	0.53
Controls	No	Yes	No	Yes
Observations	1,582	1,370	1,593	1,380
R^2	0.004	0.050	0.004	0.054

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table presents the results from an OLS regression with robust S.E. Columns 1 and 2 shows the estimates using the allocation decision (in USD) in the dictator game as the dependent variable, but restricting the sample to the participants that did not have give their full endowment (40 USD). Column 3 and 4 estimate the coefficients using as the dependent variable an indicator variable which takes value 1 if the participant gave a positive amount in the dictator game, and 0 otherwise. The variable Minority takes value 1 when the participant is in the Minority-status context and 0 otherwise. The variable Outgroup in columns 1 and 2 takes value 1 when the allocation is towards a member of the outgroup or 0 otherwise, whereas in columns 3 to 6 it takes value 1 when the inequality is advantageous towards a member of the outgroup, and 0 othwewise. Minority x Outgroup shows the estimate of the interaction. The even columns show specifications that control for all covariates measured, except for the elicited social proximity. The controls include: gender, age, political party, education, race and state fixed effects.

Figure A2: Effect of majority-minority status on inequality acceptance for only the first inequality acceptance game presented.



(a) Local inequality (b) Global inequality

Notes: Panel (a) plots, the percentage of decision-makers that accepted the inequality in the local inequality, when the two participants are from the ingroup/outgroup and in the majority/minority context. Panel (b) displays the fraction of decision-makers that accepted the inequality in the global inequality, when the two participants are one from the ingroup and one from the outgroup, but the inequality is advantageous for the ingroup/outgroup and in the majority/minority context. The data in this graphs correspond to only the decisions in the first inequality acceptance game presented.

2.B Deviations from Pre-Analysis Plan

This paper was preregistered in the AEA RCT Registry (AEARCTR-0010179). In this section, I detail the deviations from the stated pre-analysis plan (PAP).

The main deviation from the PAP corresponds to the comparison groups. This data collection was preregistered as comparing three different types of social contexts:

- **Non-contrasting** context, where the DM allocates to a receiver in a society where all five matched individuals share the same group membership.
- **Minimum-contrast** context, where the DM allocates to a receiver that belongs to the group of four out of five matched individuals that share the same group membership.
- **Maximum-contrast** context, where the DM allocates to a receiver that does not belong to the group of four out of five matched individuals that share the same group membership.

However, the comparison of decisions towards the ingroup versus towards the outgroup in the contexts as defined above corresponds to analyzing differences in allocations towards individuals with different characteristics in different contexts, meaning two elements in the treatment are changing. Thus, it is difficult to cleanly identify effects of a single change in the setting. To avoid confounds and make decisions towards members of the ingroup and the outgroup comparable across contexts, I rather focus on the following comparison groups:

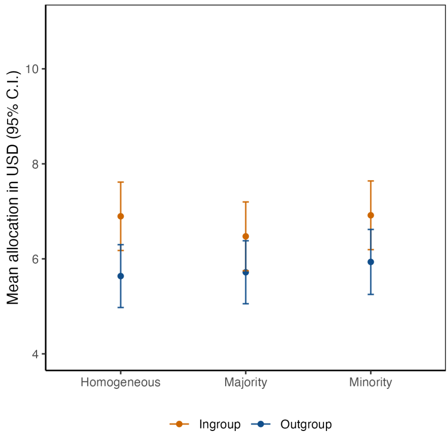
- **Majority** context, where the DM allocates to a receiver in a context where four out of five matched individuals share the same group membership as the DM, i.e., they are part of a majority.

- **Minority** context, where the DM allocates to a receiver in a context where one out of five matched individuals share the same group membership as the DM, i.e., they are part of a minority.

Given the change in the comparison groups, the main focus of the paper shifted from studying effects of changes in the level of contrast in allocations, to instead studying the effect of majority-minority status of decision-makers over their allocation decisions. Given this change in focus, I do not include the "Non-contrasting" treatment in the analysis, as I consider them independent for the assessed research question.

Figure A3 summarizes the allocation decisions: the mean allocation in USD in the dictator game, and the percentage of participants that accepted the local and the global inequality. The values are computed by whether the ingroup or the outgroup is benefited in the allocation decision, across the three social contexts collected: "Non-contrasting", "Majority" and "Minority". It is visible that there are no differences in allocations towards the ingroup nor the outgroup, regardless of which context the participant is in.

Figure A3: Effect of majority-minority status on social preferences



Notes: Figure (a) shows a the mean allocation in USD (out of 40 USD) towards a receiver from the ingroup/outgroup and in the non-contrasting, majority and minority context.

2.C Survey Questionnaire

Figure A4: Consent.

Welcome to this research project! We very much appreciate your participation.

Procedures

The study consists of a series of questions and decisions that might affect your final payoffs and the payoffs of others. Please make sure to always read the instructions carefully and answer truthfully.

Payment

For completing this survey, you will receive a participation fee of 1.75 USD / 1.4 GBP and be automatically enrolled in a lottery by the research team where you can win up to 40 USD, where 1 in every 100 participants will be selected.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to future participation in other studies conducted by us.

Confidentiality

All data obtained from you will be kept confidential. Your ID will only be used for making payments, and will not be used for any further purposes. Only anonymized data from the survey might be shared in open science repositories.

Verification

At the end of this survey, you will be automatically redirected to Prolific.

Questions about the Research

If you have questions regarding this study, you may contact:

thechoicelab@nhh.no

Please write "ACCEPT" in the box below if you wish to participate.

Figure A5: Page 1.

This is an international study that will be run on multiple samples of respondents coming from different places.

Before starting the survey, we will match you with **five other respondents** from the study. The six of you, together, will participate in a lottery to **win a total of $65 as a bonus** that will be distributed between the six (1 in every 100 groups of six will be selected).

If you win the lottery, **the amount each one receives depends on your choice only.** Therefore, please **read the instructions carefully**.

According to the previous paragraph, how will the final $65 bonus be distributed between the six of you? You must answer this question correctly in order to proceed.

- It depends on the choice of the other five respondents.
- It depends on your choice only.
- It will always be 1/6 for each.
- None of the above

→

Figure A6: Page 2.

You have been matched to the following **five other participants**. Below, for each individual, you have information about:

1. The **sample** they were drawn from.
2. The **initial distribution** of the **$65 bonus**.

One of the other five participants will be randomly selected**.** You will make a decision regarding only the selected participant.

→

Figure A7: Page 3.

Figure A8: Allocation Decision

You have the possibility to redistribute your share of the bonus **between yourself** and the **selected participant**. You can now:

- Give none or up to $40, of the bonus initially allocated to you, to the selected participant.

Your final bonus and the bonus of the selected participant is determined by your choice. The final bonus of the other (not selected) participants is independent of what you choose. None of the participants besides you is able to change their bonuses, and will not know by whom or how the bonuses were determined. The stated amounts will be paid out using the local currency equivalent to a $ in the US.

Indicate, using the slider, the amount in dollars ($) you decide to reallocate to the selected participant. The figure below will reflect the choice you make.

0 4 8 12 16 20 24 28 32 36 40

Give to selected participant (press on slider):

Please confirm your choice in the slider. Remember this choice **determines the final bonuses for you and the matched participants.**

I confirm the choice in the slider

Figure A9: First closeness elicitation

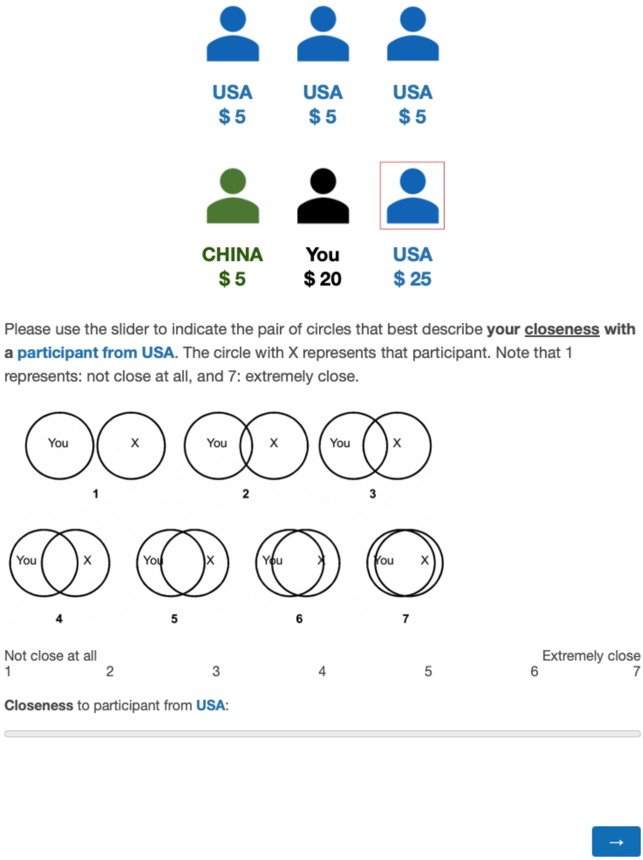


Figure A10: Second closeness elicitation

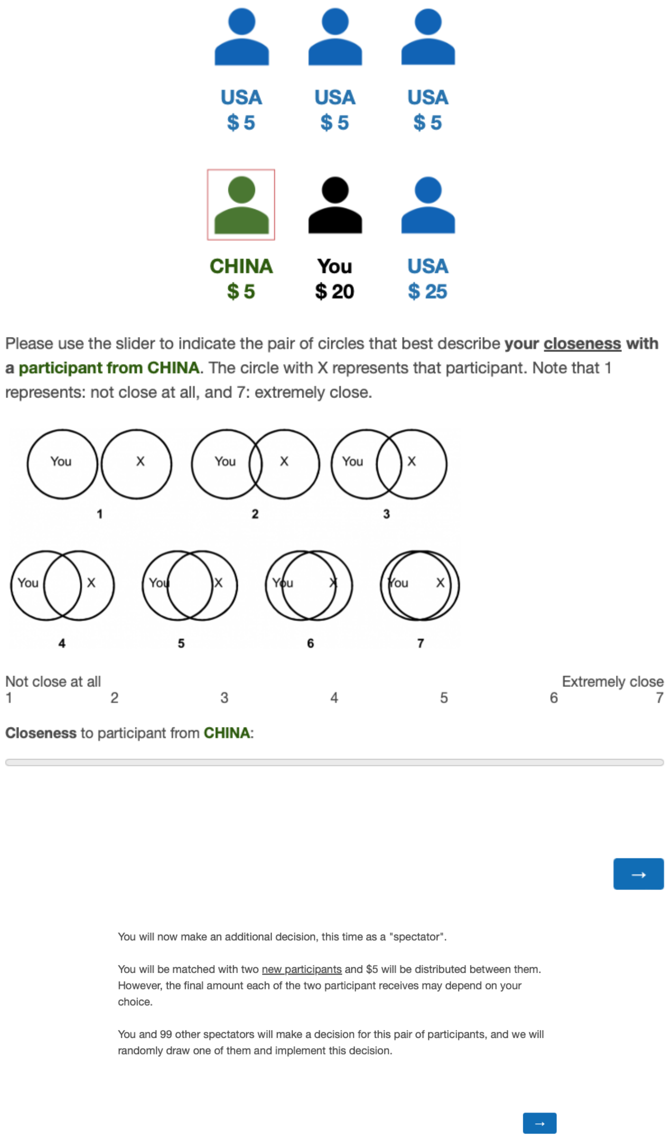


Figure A11: First inequality aversion elicitation

You were matched with new participants **A and B**. Below you have information about the country they were sampled from.

A total of $5 has been arbitrarily distributed as follows:

A B

USA $3 CHINA $2

You have the possibility to equalize the bonuses of A and B, by setting the bonus of A to $2.

You can now:

1. **Keep bonus unchanged**: A gets $3 and B gets $2.
2. **Equalize bonus**: A gets $2 and B gets $2.

Neither A or B will know how the bonus was determined. The stated amounts will be paid out using the local currency equivalent to a $ in the US.

Please indicate the alternative your prefer. The figure below will reflect the choice you make.

A B

USA $3 CHINA $2

Equalize bonus: A gets $2 and B gets $2

Keep unchanged: A gets $3 and B gets $2

Figure A12: Background demographic information. Part 1.

Background questionnaire

Please answer the following questionnaire. You must answer these questions truthfully.

What is the highest level of school you have completed or the highest degree you have received?

- High school graduate or less
- Bachelor's degree
- Master's degree
- Doctoral degree
- Professional degree (JD, MD)

What was your total household income before taxes during the past 12 months?

- Less than $25,000
- $25,000-$49,999
- $50,000-$74,999
- $75,000-$99,999
- $100,000-$149,999
- $150,000 or more
- Prefer not to say

In which state do you currently reside?

You will now make an final decision, this time as a "spectator".

You will be matched with two new participants and $5 will be distributed between them. However, the final amount each of the two participant receives may depend on your choice.

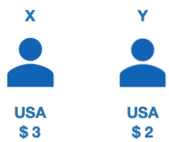
You and 99 other spectators will make a decision for this pair of participants, and we will randomly draw one of them and implement this decision.

→

Figure A13: Second inequality aversion elicitation

You were matched with new participants **X and Y**. Below you have information about the country they were sampled from.

A total of $5 has been arbitrarily distributed as follows:



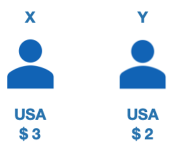
You have the possibility to equalize the bonuses of X and Y, by setting the bonus of Y to $2.

You can now:

1. **Keep bonus unchanged**: X gets $3 and Y gets $2.
2. **Equalize bonus**: X gets $2 and Y gets $2.

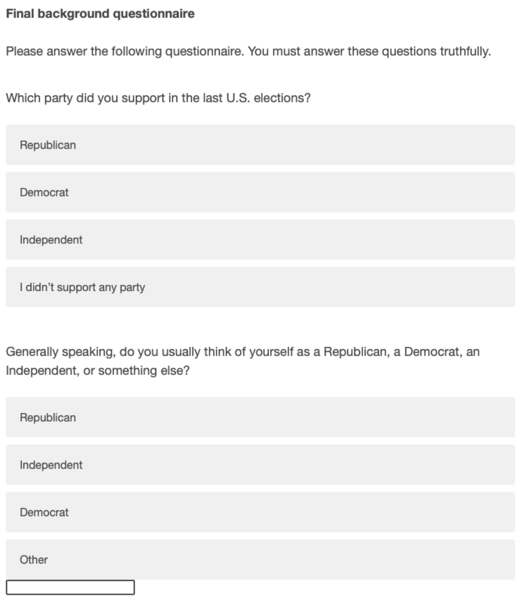
Neither X or Y will know how the bonus was determined. The stated amounts will be paid out using the local currency equivalent to a $ in the US.

Please indicate the alternative your prefer. The figure below will reflect the choice you make.



Equalize bonus: X gets $2 and Y gets $2

Keep unchanged: X gets $3 and Y gets $2



Final background questionnaire

Please answer the following questionnaire. You must answer these questions truthfully.

Which party did you support in the last U.S. elections?

Republican

Democrat

Independent

I didn't support any party

Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else?

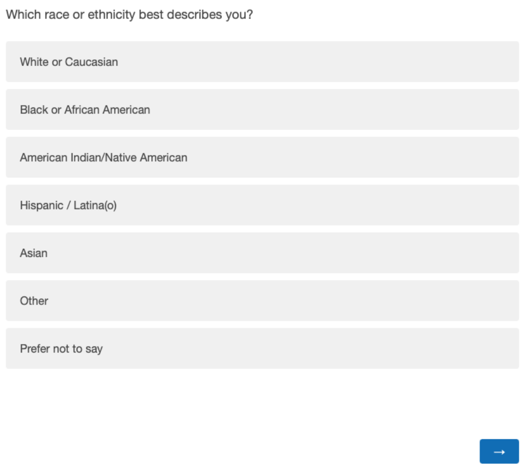
Republican

Independent

Democrat

Other

Figure A14: Background demographic information. Part 2.



Which race or ethnicity best describes you?

White or Caucasian

Black or African American

American Indian/Native American

Hispanic / Latina(o)

Asian

Other

Prefer not to say

Chapter 3

Will Artificial Intelligence get in the way of achieving gender equality?

Abstract: The promise of generative AI to increase human productivity relies on developing skills to become proficient at it. There is reason to suspect that women and men use AI tools differently, which could result in productivity and payoff gaps in a labor market increasingly demanding knowledge in AI. Thus, it is important to understand if there are gender differences in AI usage among current students. We conduct a survey at the Norwegian School of Economics collecting use and attitudes towards ChatGPT, a measure of AI proficiency, and responses to policies allowing or forbidding ChatGPT use. Three key findings emerge: first, female students report a significantly lower use of ChatGPT compared to their male counterparts. Second, male students are more skilled at writing successful prompts, even after accounting for higher ChatGPT usage. Third, imposing university bans on ChatGPT use widens the gender gap in intended use substantially. We provide insights into potential factors influencing the AI adoption gender gap and highlight the role of appropriate encouragement and policies in allowing female students to benefit from AI usage, thereby mitigating potential impacts on later labor market outcomes.

This chapter is written together with Catalina Franco and Siri Isaksson. The authors would like to thank Andrea Bocchino, Finn Casey, Iver Finne, Erika Povea, María Recalde, Stig Tenold, Heidi Thysen, and Bertil Tungodden. Audiences at the ASSA 2024 meeting, FAIR, and NHH provided useful comments and feedback. This study was preregistered in the AEA RCT Registry (AEARCTR-0012452).

3.1 Introduction

Within a year of its release, ChatGPT has already left a mark. Companies have expressed interest in candidates with knowledge of how to use the tool (CNBC, 2023), and new well-paid jobs as "prompt engineer" are quickly emerging (WSJ, 2023). Recent studies indicate that the use of artificial intelligence (AI) tools such as ChatGPT provides substantial productivity gains across domains. For instance, allowing access to AI tools improved output quality in professional writing tasks among online workers by 18% (Noy and Zhang, 2023), increased solutions to issues in real-life customer support tasks per hour by 14% (Brynjolfsson, Li and Raymond, 2023), and reduced the time developers used in coding tasks by 56% (Peng et al., 2023). Although exact economic impacts are hard to predict and depend on policies adopted (Brynjolfsson and Unger, 2023), AI proficiency is likely to shape labor market paths and success in the near future. Therefore, it is crucial to assess the adoption and use of these new technologies by students facing this fast-paced labor market, particularly amidst the current heated debate on whether to allow or ban the use of ChatGPT.

This paper focuses on a potential disparity in adoption and proficiency in ChatGPT use based on gender, a side overlooked so far in the debate. Gender likely plays a role in AI adoption based on previously documented gender disparities in internet usage (the so-called "gender digital divide") (Bimber, 2000; OECD, 2018), in technology-related career choices (Buser, Niederle and Oosterbeek, 2014; Buser, Peter and Wolter, 2017; Cimpian, Kim and McDermott, 2020), and in confidence regarding skills in male-dominated tasks and the prevalence of stereotypes (Bordalo et al., 2016*b*). Using a survey experiment on university students in Norway, we find substantial gender differences in both adoption and proficiency of ChatGPT usage. We also identify potential explanatory factors influencing this gender gap.

The rapid growth and unprecedented capabilities of ChatGPT and other generative AI technologies have raised concerns among educational institutions, prompting calls for regulatory measures regarding its use. Varying policies have been proposed ranging from outright bans to embracing and incorporating AI tools in the learning process. Those arguing that it should be banned cite fears about students submitting inauthentic and potentially plagiarized work, substituting the development of critical and problem-solving skills as students get easy and quick answers, and information privacy concerns. Supporters of embracing ChatGPT think that the technology is here to stay and should be incorporated in the classroom to guide students in making a productive use of it.[1] We believe gender should be a crucial aspect in the debate on whether to ban or allow AI use by students, as differential responses may unintentionally create gender-biased policies.[2] With this aim, our survey also provides evidence of large differences in how female and male students would respond to university bans of the tool, and shows that imposing university bans on the use of ChatGPT widens the gender gap in intended use substantially.

We conducted a preregistered anonymous survey experiment

[1]See Lo (2023) for a review of the advantages and disadvantages of ChatGPT use in education.

[2]Since there is no evidence yet on whether ChatGPT helps or hurts students, we are agnostic on who will be harmed by these policies. For example, ChatGPT use could hinder critical and problem-solving skills or help develop them further. In this case, even though the demand for AI skills will most likely increase in the near future, acquiring these skills at the expense of critical and problem-solving skills will probably not help students in the labor market. On the other hand, if acquiring AI skills does not hinder other labor-market relevant skills, differential adoption will likely make slow adopters fall behind by missing learning and career opportunities in a rapidly-developing labor market.

with 514 students at the Norwegian School of Economics (NHH) in November 2023.[3] Participants in the study were recruited in class from the first and third-year bachelor's program cohorts, as well as from the master's program. We collected student self-reports on current ChatGPT usage and measured prompting skills. We preregistered hypotheses that perceptions about the technology, preferences regarding its use, and their experience and exposure could constitute potential factors influencing adoption and skills, and collected measures accordingly for each factor.[4] Crucial for understanding potential differential responses to policies allowing or forbidding ChatGPT, we included a vignette experiment describing a course that students would hypothetically be enrolled in. Keeping all other information constant, the description randomly displayed whether the professor explicitly allowed or forbade the use of ChatGPT in the course, and students were asked to report their intended use of ChatGPT throughout the course.

We report three main findings. First, female students are much less likely to currently use ChatGPT than male students. A larger fraction of women than men report having heard of ChatGPT but do not use it (11.2% vs. 2.5%), and to having used it only a few times (31.9% vs. 23%). Further, the proportion of women who report using it all the time is almost half that of men (25.4% vs. 44.3%). Overall, the raw gap in high ChatGPT use (occasionally or all the time) is 17.2 percentage points (pp) or 30% over a base of 57% among women. This estimate does not reflect gender dif-

[3]NHH is the most selective higher education institution in Norway, so our effects should be interpreted as a lower bound. We consider anonymity to be important in this context because we want to obtain truthful responses and not responses reflecting what they think should be correct if they knew we were matching survey answers to their academic record.

[4]Besides reporting raw gender gaps, we add these factors as well as other baseline variables to examine the extent to which they explain the gaps.

ferences in course selection since the curriculum is mostly fixed within cohorts at NHH. Moreover, including year in college and admission grade as well as measures of risk and time preferences reduces the gender gap to 9.8 pp. That is, these baseline characteristics explain about 43% of the raw gap. Adding the full set of potential factors influencing adoption, capturing perceptions, preferences, and experience/exposure related to ChatGPT further reduces the gap to 1.2 pp, which is not statistically different from zero.

Second, men are more skilled at ChatGPT prompting than women even after controlling for baseline use. We measure proficiency by asking students to write a prompt that we then feed into ChatGPT to assess whether it gets the correct answer.[5] While there is no significant gender gap in time spent writing the prompt (132 seconds on average), we find gender gaps of about 0.3 standard deviations (SDs) in the variables measuring the number of characters written (179 on average for men), and success rate of the prompt (getting the correct answer 36% of the time on average for men).[6] The gender gap in success rates disappears when adding the full set of controls, including the potential factors influencing use. Specifically, gender differences in perceptions, such as the confidence that the prompt will give the correct answer, have the most explanatory power, showing the strength of the relationship between confidence and ability.

Third, the gender gap in intended use widens if ChatGPT is banned. About 80% of both female and male students random-

[5]Large Language Models (LLMs) provide different answers each time a prompt is submitted, which in some cases might be correct or not. Therefore, we run each prompt over 100 times and collect how many times the prompt gives the correct answer.

[6]The fractions of women and men with at least a 50% success rate are 25% and 37%, respectively.

ized into the professor “allows” ChatGPT treatment state that they would use it in the course. However, women in the professor “forbids” ChatGPT treatment are 38 pp less likely to use it than women in the “allows” treatment. A gender gap equal to 20 pp opens up since men’s intended use is much less likely to be affected by the bans. The gender gap in intended use and the within-gender reaction to the policy is virtually unaffected by adding the full range of control variables and potential factors influencing adoption mentioned above. Hence, we conclude that other aspects that might vary by gender, such as rule-following behavior, obeying the authority or having trust in the professor’s recommendation since they know what is best for students, must be behind the differences in intended use after the policy. Most importantly, this shows that banning ChatGPT in the classroom might have large unintended consequences by putting female students behind their male peers in AI adoption.

Finally, we discuss additional descriptive findings in light of the existing literature on gender differences in choices. Using self-reported admission grades,[7] we look into differences in use and reactions to the hypothetical “allow/forbid” policy by admission grade quintile. While men across all grade quintiles have similar usage rates (gravitating around 80%), women in the top quintiles are much less likely to be currently using ChatGPT (around 40%) relative to women in the bottom two quintiles (over 80%) and than men in any quintile. The responses to the policy are quite similar across admission grade quintiles for men, while only similar across

[7] Admission grades are based on high school GPA and retakes of courses for students who do not get in on their first attempt. 273 of 514 students provided valid admission grade responses, which prevents us from doing a full heterogeneity analysis. There are no gender differences in the likelihood of reporting the admission grade nor in the reported grade.

quintiles for women in the "allow" treatment. In the "forbid" treatment, women's intended adoption rates are much smaller in the top quintiles.[8]

We note the resemblance between our findings by admission grade quintile and previously documented patterns of top women, in particular, exhibiting behaviors most dissimilar from men. For example, men are less sensitive to the grade they obtain in a principles class when deciding whether to major in the same field as that class. Women are much more sensitive, with only the women earning the highest grades in the principles class declaring a major in the same field (Rask and Tiefenthaler, 2008; Ost, 2010; Goldin, 2015; Avilova and Goldin, 2018; Kugler, Tinsley and Ukhaneva, 2021; Ugalde, 2022). Niederle and Vesterlund (2007) find that women in the top performance quartile are willing to compete to a similar extent as men in lower quartiles, and Coffman (2014) finds that expert women are less likely to speak up. We document similar patterns in a completely new domain: using ChatGPT and responses to policies on ChatGPT use, a skill that is becoming increasingly relevant for labor market success. Crucially, we contribute to the previous literature by showing that top women may be willing to adopt behaviors in which there are ex-ante gender differences through a change in policies or recommendations that alter how the behavior is portrayed. If women are disproportionately affected by negative portrayals of what it means to choose a major when one's grade was not among the best or what it means to enter a competition when one's chances of winning are not the highest, they may simply opt out.

Our findings suggest that more positive portrayals (it is okay/allowed

[8] We also see that women who do not use ChatGPT at baseline have much larger reactions to the "forbid" policy than women who already use it occasionally or all the time and than men in all usage categories.

to apply/compete even if you fail) by an authority figure (e.g., a professor) may go a long way in closing gender gaps in choices. Moreover, we believe that gender differences in rule-following or trusting advice from authority opens up a new agenda of research in the topic of gender and behavior.

An additional implication of our results is that recent findings on how using AI recruitment tools increases gender diversity in the workplace (Avery, Leibbrandt and Vecci, 2023) may be attenuated by women not having the requirements to apply for the increasing number of jobs that require AI skills. If women develop AI skills to a lesser extent than men while in college, as we document, the prospect of increasing gender diversity with debiased recruitment (Pisanelli, 2022; Awad et al., 2023) may be harder to attain.

3.2 Setting and Research Design

3.2.1 Participants and recruitment

Participants in our survey are recruited from the first and third year of the bachelor's and master's programs at NHH. The school offers a five-year program consisting of three years of a bachelor's program in economics and business administration followed by two years of a master's program in either economics and business administration or international management. Education is free and students who are admitted into the bachelor's program automatically get a slot for the master's programs and typically continue with the master's, but can leave after completing the three years of the bachelor's program only.[9]

[9] The bachelor's program is taught in Norwegian, while the master's programs are taught in English.

The bachelor's program at NHH is the most popular program in Norway listed as the first choice of most applicants to higher education. In 2023, it was listed as a first choice by 2,170 applicants who competed for 500 slots.[10] 50% of admitted students come straight from high school (first-time admission) and the other 50% usually retake some subjects or do some activities after graduating from high school that grants them higher admission points to be more competitive in the admission process. The 2023 admission cutoffs for the first-time admission and regular admission were 55.6 and 59.5, respectively. For reference, grades in Norway go from 1 to 6, and GPAs are calculated from high school grades and the score in five to six exams taken throughout high school (Landaud et al., (accepted). The cutoffs, calculated by multiplying the GPA by 10, illustrate that successful applicants in both admission categories typically achieve scores close to a perfect 6 in every school and exam subject.

In the bachelor's program students take 4 subjects every semester, for a total of 24 of which only 6 are elective.[11] Subjects in the master's programs involve 6 subjects and a master's thesis, where at least 3 of the 6 subjects must be selected from a list of mandatory subjects. We believe that the small role of elective courses, particularly in the bachelor's program, make a strong case that our results are not simply driven by gender differences in the choice of subjects that are more or less amenable to the use of ChatGPT.

Students participating in the survey were recruited during lecture hours of two of the mandatory courses of the bachelor's pro-

[10] Almost 5,000 applicants listed the NHH program in any rank of their list. There were 62,757 higher education slots in Norway in 2023 (Direktoratet for høyere utdanning og kompetanse, 2023).

[11] There are no electives in the Autumn semester of the first year (where half of our sample is recruited from), and one elective thereafter except in the last semester of the program in which students can choose two electives.

gram: a first-year and a third-year course, as well as one of the core courses in the master's program. The survey experiment was pre-registered in the AEA RCT Registry (AEARCTR-0012452) and the data was collected subsequently in November 2023. The survey was anonymous and implemented in the classroom using a QR code. Students lasted on average 7 and a half minutes responding the survey.[12]

3.2.2 Anonymity and Participant Incentives

In considering the best format to administer the survey, we weighed the prospect of linking student responses to their future academic performance with the potential for misrepresentation of ChatGPT use and experimenter-demand effects if students knew that the survey was not anonymous. Since this is the first study documenting patterns in student use of ChatGPT, we opted for anonymity as we put the highest value on truthful responses.

Related to anonymity, incentivizing the prompting exercise and second-order beliefs questions would have required collecting some personal information to provide incentives. We also opted for conducting the survey in the classroom to avoid students getting external help (from someone else or from ChatGPT) to get the prompt correct.

Validation exercises have found strong similarities in the use of hypothetical and unincentivized measures relative to incentivized elicitations and real-world behavior across different domains (Hainmueller, Hangartner and Yamamoto, 2015; Brañas-Garza et al., 2021, 2023; Enke, Rodríguez-Padilla and Zimmermann, 2022; Falk et al., 2023). At the same time, there has been an increase in the use

[12] On average, women spend 7.7 minutes and men 7.3 minutes. The difference is not statistically significant.

of unincentivized measures in economics research (Ameriks et al., 2020; Bernheim et al., 2022; Stango and Zinman, 2023; Almås, Attanasio and Jervis, 2023; Andre et al., 2022). Given the restrictions in our scenario and the concerns over potential effects of incentives on reporting actual capabilities, we opted for the use of unincentivized questions.

3.2.3 Survey design

The survey consists of four sections: background characteristics, a hypothetical vignette experiment, a prompting skills task, and a questionnaire on the use and attitudes about ChatGPT, presented to the respondent in that order. The questionnaire is in Appendix 3.B.

Background characteristics. First, participants were asked questions on demographic and academic background, including gender, whether the student is from Norway, risk and time preferences measured through survey questions following Falk et al. (2018). Students were given the possibility of reporting or not their admission grade, with 273 students reporting valid responses out of the 514 respondents (53% of the sample).

Use and attitudes about ChatGPT. Participants indicated self-reported use of ChatGPT. Our baseline use outcome is obtained from the question *"How familiar are you with ChatGPT?"*, with choices corresponding to *low use* if the participant indicated: "not heard about it", "heard about it but not using it myself", or "used it a few times", which indicates none or limited use; and *high use* if the participant indicated "use it occasionally", or "use it regularly", which indicates continuous use. Participants also selected the types of tasks they "typically ask ChatGPT to help with".

Prompting skills measure. To measure a participant's skill in the use of ChatGPT, we presented participants with an image of the "Ebbinghaus illusion",[13] and asked them to write in a text box the query/prompt they would provide to ChatGPT to arrive at the correct official name of the visual phenomenon represented by the image. We use three outcome measures based on the prompting exercise: time spent writing the prompt, the number of characters written, and the success rate of the prompt, given by the proportion of ChatGPT answers that mention the official name out of over 100 queries made, for each prompt.

Potential factors influencing usage. We elicited attitudes of respondents regarding ChatGPT, which we preregistered and classified into three categories of primary factors affecting ChatGPT usage: (i) preferences, (ii) perceptions, and (iii) exposure/experience. In terms of preferences, we aim to measure potential utilitarian costs or benefits associated with ChatGPT usage and examine the role of persistence in the use of technology. Concerning perceptions, we focus on four key areas: perceived usefulness of ChatGPT, whether ChatGPT usage is considered cheating, trust in the accuracy of information provided by ChatGPT, and confidence in one's abilities to use the technology. Lastly, we explore the exposure/experience factor, analyzing how prior exposure to ChatGPT might influence its adoption.

Hypothetical vignette experiment. Participants were presented with a hypothetical scenario, describing a course the participants are hypothetically enrolled in. The course description indicates

[13]The Ebbinghaus-Titchener illusion (Titchener, 1901) is represented by two circles of the same size that are surrounded by a different context each: the first circle is surrounded by small circles and the second circle is surrounded by big circles. When most observers view these figures, the context affects perceptions of size. The image used is presented in Appendix 3.B.

how it is evaluated and we experimentally vary a statement of whether the professor explicitly allows or forbids the use of ChatGPT in the course as follows:

> *"Imagine you are enrolled in a course on Environmental Policy and Economic Impact. This course explores the intersection of environmental regulations, economic incentives, and their effects on industry practices and sustainability. The professor explicitly allows/forbids the use of ChatGPT during coursework. It is an 8-week course with final evaluation given by a final in-person written exam."*

Subsequently, the respondent was asked: *"Given this scenario, how likely are you to use ChatGPT throughout the course?"*, where the choice consists of indicating intended use in a 5-point scale from "Very unlikely" to "Very likely".

Participants stratified by gender were randomly allocated into one of two treatment conditions: (i) when the professor explicitly *allows* the use of ChatGPT, and (ii) when the professor explicitly *forbids* the use of ChatGPT. This allows us to causally study the effects of the allow/forbid policy on intended use. A second layer of randomization was the type of evaluation of the course, where the evaluation could be either an in-person exam or a home exam.[14]

3.2.4 Sample characteristics

Almost 55% of our sample is male, which is close to the historical male student representation at NHH of about 60% (Hirshman

[14] Respondents that were presented with the home exam scenario were asked a second question: *"Given this scenario, how likely are you to use ChatGPT during the final exam?"* This way, respondents would differentiate the use of ChatGPT throughout the course and during the exam in order to make the measures comparable across different evaluation scenarios. We are not using this layer of randomization in this draft.

and Willén, 2022). 54% and 40% of the sample are in the first and third year of the bachelor's program, respectively. Men are statistically more willing to take risks and give up something beneficial today in order to benefit more from that in the future (Falk et al., 2018) than women in the sample. While only 53% of the sample provided a valid answer for their admission grade, there are no gender differences in the likelihood of reporting the grade or in the grade itself. On average, the admission grade is 5.6 (median equal to 5.7) for both men and women, and the distributions look quite similar.

3.2.5 Empirical Strategy

We use two main econometric specifications. For the outcomes related to baseline use and prompt success rate described in Section 3.2.3, we focus on estimating the gender gap using an indicator for whether the participant is a male student:

$$y_i = \alpha_0 + \alpha_1 \text{Male}_i + X_i \gamma + \varepsilon_i \tag{3.1}$$

We measure the gender gap through the coefficient α_1. In our main results table, we present the raw gap along with a series of controls X_i including baseline use (for the success rate outcome only), background characteristics, and preferences, perceptions, and experience as described above.

Our second econometric specification involves estimating the gender gap for the policy reaction to allowing/forbidding ChatGPT in the hypothetical course presented in the vignette experiment:

$$y_i = \beta_0 + \beta_1 \text{Male}_i + \beta_2 \text{ChatGPT forbidden}_i + \beta_3 \text{Male}_i \times \text{ChatGPT forbidden}_i + X_i\gamma + \epsilon_i \quad (3.2)$$

The outcome y_i is equal to 1 for students who state that they are likely or very likely to use ChatGPT during the course. The coefficient β_1 provides the gender gap when ChatGPT is allowed, β_2 represents the policy response (from allowed to forbidden) among women, and β_3 measures the differential change in the policy response for men relative to women. Similarly as in specification 3.1, we add different types of controls that help us understand the influence of the preregistered factors on our results.

3.3 Results

3.3.1 Female Students Are Less Likely to Use ChatGPT Than Male Students

We begin by analyzing the responses to the survey question *"How familiar are you with ChatGPT?,"* which contains 5 answer options: not heard, heard but not use, used few times, use occasionally, and use all the time. Figure 3.1 shows the proportion of responses in each category split by gender, with the height of the bars adding up to 100% within gender. Women are much more likely to be represented in low use categories. 11.2% of women while 2.5% of men state that they have heard about ChatGPT but do not use it. 31.9% of women and 23% of men have used it a few times. Only 1 out of 514 students answered not to have heard about it. In contrast, men are overrepresented in the use all the time category with 44.3% of men relative to 25.4% of women. The proportions in the use occasionally category are similar with 31.5% of women and 29.8% of men.

We statistically estimate the gender gap in use through specification 3.1, where the outcome is a binary measure indicating a high use if the participant responds use occasionally or use all the time. Overall, the raw gender gap in high use at baseline is estimated at 17.2 pp or 30% over a base of 56.9% of women using ChatGPT occasionally or all the time (see Column 1 in Panel A of Table 3.1).

To understand the overall gender gap in use, it may be insightful to plot this variable by a measure of relative academic skill.[15] For example, given a level of skill, students may use ChatGPT less or more depending on how they think it complements or substitutes their own skills. Figures 3.4a and 3.4b show, for women and men separately, the fraction of students reporting a high use by quintile of the admission score distribution.[16] The fraction of men with high ChatGPT use (Figure 3.4b) is between 73% in the highest quintile up to 84% in the middle quintile, so it is quite homogeneous across quintiles. In contrast, the fraction of women with high ChatGPT is strongly and negatively correlated with admission grade quintile. In the bottom two quintiles, the fraction of women with high baseline use is similar to the fraction of all men, while for the three top quintiles, the fraction of women with high baseline use is below 45% (Figure 3.4a). A regression estimating the correlation between the raw admission grade and the high baseline use indicator yields a negative and significant coefficient for both men and women, but it is almost four times larger for

[15] Admission grades tend to be correlated with college GPA, which in turn increases hiring interest by employers (Kessler, Low and Sullivan, 2019).

[16] Quintiles are calculated pooling men's and women's admission grades.

women (-0.378) than for men (-0.097).

Our results are in line with previous findings suggesting a correlation between women's choices according to their position in the skill distribution in choices based on laboratory tasks (Niederle and Vesterlund, 2007; Coffman, 2014) and on the grade in a principle's class determining what college major students enroll in (Rask and Tiefenthaler, 2008; Ost, 2010; Avilova and Goldin, 2018; Kugler, Tinsley and Ukhaneva, 2021; Ugalde, 2022). In our setting, we find that top women engage less in ChatGPT use, a behavior that may be perceived as showing that they are not as qualified as they are.[17] If female students interpret ChatGPT as such, they may care more about how they will be perceived by employers down the line given the evidence that student beliefs about hiring decisions affect important decisions such as which college major to pursue (Ugalde, 2022). Women may opt for not using ChatGPT to avoid giving the wrong signal to fellow students, professors, or employers. However, we show in Section 3.3.4 that institutional policies on ChatGPT use can affect intended use and that women are as likely to intend using it as men under certain scenarios, which suggests that needing to prove themselves in college and to employers must not be playing a first-order role in the gender gap in use.

3.3.2 Male Students Are More Skilled at Prompting than Female Students

As mentioned before, proficiency in AI tools like ChatGPT is becoming an increasingly important skill for labor market success. We documented in the previous result that female students, in par-

[17] For example, Williams (2014) find that a majority of female scientists report in a survey that they feel the need to provide more evidence of their competence than others to prove themselves to their colleagues.

ticular top women, are using ChatGPT to a lesser extent than male students. Lower use can directly impact proficiency since acquiring it probably results from continued use with a tool. We show that male students are more skilled at writing successful ChatGPT prompts than female students even after accounting for baseline use.

Figure 3.2 shows standardized versions of three outcomes measuring prompt quality: Time spent, success rate out of 108 runs of the same prompt on ChatGPT, and number of characters written. These variables are standardized relative to the mean and standard deviation of men in the sample. The stars next to each gap visualization correspond to the statistical difference in the raw gap.

On average, men spend 132 seconds writing their prompt, and women spend less time, but not statistically significantly so. The average success rate recording the fraction of times that the prompt obtains the desired answer (Ebbinghaus illusion) is 36% for male students and lower by 11 pp or 0.26 SD for women (see also Table 3.1, Column 1 in Panel B). Lastly, female students write about 0.3 SD fewer characters in their prompt relative to a mean of 179 characters among male students. The success rate and number of characters seem to be strongly and positively correlated, as shown in Figure A1.

Table 3.1, Panel B, Column 1 quantifies the raw gap in prompt success rates. On average, women have a success rate of 24.9%, meaning their prompt gives the correct answer about 27 times out of 108 ChatGPT runs. The gender difference is estimated at 11.1 pp, which means that men get the correct answer about a third of the time on average.

As expected, in Figures 3.5a and 3.5b, students at the top of the admission grade distribution have higher success rates with their prompts regardless of gender. In the top quintile of the distribution, students have success rates of about 41-42%, which is in stark contrast to the overall average of 31%. As with the high baseline use outcome, men have more homogeneous success rates across quintiles than women. Even though women in quintile 1 have the highest ChatGPT baseline use, their success rate (17%) is the lowest among all and almost half of the success rate for men in the same quintile (30%), who have similar levels of baseline use.

3.3.3 Potential factors influencing adoption and skills

As discussed in section 3.2.3, we preregistered and measured three main categories of potential factors influencing adoption and use of ChatGPT: (i) preferences, (ii) perceptions, and (iii) experience or exposure. The results are summarized in Figure 3.7.

Gender Differences in Preferences

As preferences, we consider three factors. For the first two, we ask students to indicate their agreement or disagreement with the following two statements: *"I think ChatGPT is enjoyable to use"*, and *"I think ChatGPT is difficult to use"*, representing a utilitarian benefit and cost from using ChatGPT, respectively. The choices range on a 5-point scale from strongly agree to strongly disagree. Figure 3.7b shows the percentage of students that indicated disagreement with the claim that ChatGPT is difficult to use, and agreement with the claim that using ChatGPT is enjoyable. While 63% of women find ChatGPT not difficult to use, and 68% find it enjoyable to use; in both cases, the percentage of men is higher than women, by 7 pp for disagreement that it is difficult to use, and by 12 pp for agreement that it is enjoyable to use, the latter being significant at the 1% level. This suggests that men have stronger preferences for the use of ChatGPT, as they find it more enjoyable (higher utilitarian benefit), and less difficult (lower utilitarian cost) to use than women.

We also measured what we refer to as "persistence", where the participants were asked *"If ChatGPT does not provide the desired answer on your first attempt, how many additional attempts do you typically make?"* with four options ranging from "One more try" to "I keep until satisfied".[18] We find that 58% of female students indicate that they attempt two more tries or more, compared to 73% of male students, a difference significant at a 1% level. This indicates that men are more persistent in attempting to obtain desired results than women, which can lead to gender differences in the use of ChatGPT, as men would be more eager to maintain longer "conversations" with ChatGPT for a specific query. Moreover, the gap in persistence could generate differences in skill, as men can learn more from the increased prompting experience.

We now aim to understand the relationship between the gender differences in responses to survey questions related to preferences in the use of ChatGPT, and the self-reported adoption and skills of the technology. In our regression analysis, incorporating preferences factors helps explain a significant part of the gender gap in ChatGPT use, with the gender gap coefficient, previously at 17.2 pp, now being 5.1 pp and not statistically significant (see Table 3.1, Panel A, Column 3). However, the same exercise on the success rate of the prompt (Panel B of Table 3.1) keeps the gap at around 10 pp, a similar level to the initial gap. This suggests that our measures of preferences seem to capture part of the gender gap in use,

[18] We also allow participants to indicate that they do not use it, and these participants are excluded from the analysis of these covariates.

but not in their ChatGPT skills.

Gender Differences in Perceptions

We also consider belief-based motives that can affect behavior in our setting, which we categorize as perceptions. We consider four relevant sets of perception over the use of ChatGPT: (i) perceptions over its use as cheating, (ii) perceptions of its usefulness, (iii) trust in its accuracy in providing information, and (iv) confidence in one's own skills using it. The perceptions are illustrated in Figure 3.7a, which shows the percentage of participants who align with a series of statements, representing the different sets of perceptions.

First, students might not adopt the technology if they perceive its use is unethical/cheating. To measure this, participants were asked to indicate agreement or disagreement with the following two statements: *"Using ChatGPT as an aid to solve assignments in a course is equivalent to cheating"*, and *"Using ChatGPT as a learning aid in a course is equivalent to cheating"*, with options ranging on a 5-point scale from strongly disagree to strongly agree. Figure 3.7a shows the percentage indicating either strongly disagree or somewhat disagree with the statement. While the majority of participants disagree with considering the use of ChatGPT as equivalent to cheating, there are important gender differences, with around 13 pp more men disagreeing relative to women, a significant difference at the 1% level. However, it is important to highlight that when the use of ChatGPT is as a learning aid, 83% of participants disagree with the use being equivalent to cheating, relative to a 58% disagreeing when the use is as an aid to solve assignments. Moreover, a related statement to those on cheating in our survey is *"It is easy for professors to identify if a student has used ChatGPT"*, which measures the perceived risk of getting caught using Chat-

GPT. Figure 3.7a shows that while 43% of women disagree with the statement, the proportion of men that disagree is weakly higher (51%).

A highlighted factor in previous work on the "gender digital divide" in driving gender differences in the use of the internet corresponded to the perceptions of the usefulness of the technology in different tasks (OECD, 2018). We capture perceptions of usefulness of ChatGPT by asking students to indicate *"What do you believe are the main advantages of using ChatGPT in coursework?"*. Figure 3.7a shows the percentage of students that indicated each statement as an advantage of using ChatGPT. While almost no one sees no advantages of using ChatGPT, there are strong gender differences in perceptions of usefulness. Only 17% of women believe it can improve their grades in a course, whereas 32% of men believe it can, almost double the proportion. There are also strong differences in the perception that it increases accuracy or work quality, with 29% of women and 42% of men holding this belief. Additionally, slightly fewer than half of female students (48%) believe that ChatGPT improves the learning of course methods, whereas the majority of men (63%) hold this belief. The differences mentioned are significant at the 1% significance level. However, in terms of saving time, there are no strong gender differences in perceptions, with most men (86%) and women (80%) believing it is a main advantage of ChatGPT. Altogether, these results show that men perceive ChatGPT as more useful than women, consistent with previous findings in other technology-related settings.

There could also be potential differences in trust in the accuracy of the information provided by ChatGPT, affecting the perceived benefits of using the technology. To capture this, we presented participants with a screen capture of a real prompt and answer submitted to and by ChatGPT, respectively, where the participant was asked whether they trust that the information provided by ChatGPT was accurate, using a 4-point scale from "Completely trust" to "Completely distrust".[19] Figure 3.7a shows that there are no differences in trust in the accuracy of information provided by ChatGPT, where a majority of men and women (63%) indicated either "Somewhat trust" or "Completely trust".

Finally, confidence in their skills in using the technology might affect women's willingness to engage with ChatGPT, as it has been shown in previous research using male-dominated settings (Coffman, Collis and Kulkarni, 2023). To measure confidence, we take advantage of the prompting task the students performed, and asked them, *"How confident do you feel that the query you just provided will make ChatGPT get the information you need?"*, with choices within a 4-point scale ranging from "Not confident at all" to "Extremely confident". We observe important differences in confidence by gender. 60% of women and 80% of men indicate some level of confidence in their prompt. Moreover, as represented in Figure A2a, around 40% of men indicate feeling very or extremely confident in their own prompt being correct, relative to only 17% of women.

Overall, men have more positive perceptions towards ChatGPT than women, and these seem to play a key role in explaining the gender gap in ChatGPT use and prompting skills. The gap in both outcomes vanishes once we use the measures of perceptions as controls (see Table 3.1, Column 4). In particular, the perceptions of usefulness and confidence seem to be particularly important

[19] The query asked to ChatGPT in the example provided was the following: "What is the poverty rate in Denmark". The participants were later asked, *"Based on this response from ChatGPT, how much do you trust that the poverty rate reported is accurate?"* (see Appendix 3.B).

in explaining both gaps. This is represented in Table A1, where we observe how the gap changes after controlling for the different sets of perceptions. Column 1 in Panel A shows that controlling for usefulness reduces by half the gender gap in baseline use. Column 5 in Panel B shows that the variable that has the most explanatory power for the gender gap in success rate is how confident they feel that their prompt will provide the correct answer. Adding the level of confidence by itself reduces the gender gap in success rates to 4.1 pp, which is no longer statistically significant. Figure A2a shows that there are indeed large gender differences in the levels of confidence that the prompt provided will give the correct answer. Panel (b) further shows that success rates are positively correlated with confidence levels, and there are no gender differences in success rates within a stated level of confidence.

Confidence explaining away the gender gap in success rates can have two different interpretations. One is that men are better at prompting and they know it. The other is that men are more overconfident about their prompting skills, suggesting that their high level of confidence is at least partially unfounded.[20] In the latter interpretation, people with overconfident beliefs would exert more effort since exaggerated beliefs have a motivational value (Chen and Schildberg-Hörisch, 2019). We assess levels of under- and overconfidence in our sample by constructing two indicator variables. Underconfidence is defined as having a success rate in the prompt of at least 50%, but stating being only slightly confident or not confident at all in the prompt. Overconfidence is the

[20] Male overconfidence has long been documented in academically-related domains such as relative performance in adding tasks (Niederle and Vesterlund, 2007, 2011) and cognitive tests (Buser, Gerhards and Van Der Weele, 2018; Möbius et al., 2022), as well as in domains related to the job search of recent graduates (Cortés et al., 2022). For other references see Croson and Gneezy (2009).

opposite, that is, having a success rate below 50% and stating to be very or extremely confident in the prompt. About 16% of both men and women are underconfident, while 9.5% of women but 19.5% of men are overconfident. Even though the gender difference in overconfidence is large, almost 65% of the men have an accurate idea of their skill, which suggests that overconfidence is not the full story behind men having more successful prompts. Actual proficiency and being aware of it is a main part of the story.

Gender Differences in Experience or Exposure

Finally, a gender gap in the use and skills of ChatGPT might be driven by male and female students having different levels of experience or exposure to the technology, through peers or previous experience. To measure exposure through peers, we asked participants to *"indicate the percentage of people you believe use ChatGPT"* for three different groups: their group of friends, students in their course, and professors at NHH.[21] Figure 3.7c shows the average percentage indicated by the students for each of the groups. Note that there is a significant difference in the percentage of friends that they believe use ChatGPT, with women stating that it is around 63%, and men indicating that it is 70% of their friends. However, their beliefs about other students in the course and professors at NHH are not different, where both believe around 72% of other students and only around 40% of professors use ChatGPT.[22] In a related proxy measure of experience, we asked stu-

[21] To avoid concerns of men and women having different anchors when estimating this percentage, we provided the following statement before the question: "A survey conducted among university students in the US in the Spring of 2023 reports that 30% of students use ChatGPT for their schoolwork.".

[22] The differences in the percentage of friends may be driven by women having more female friends than men, and not necessarily from inaccurate assessments of

dents whether they have *"ever received inaccurate or misleading information from ChatGPT?"*, with possible answers being "No, never", "Yes, few times" and "Yes, many times", as well as an option for those who have not used it. In Figure 3.7c, the percentage of students who have experienced inaccurate information is 16 pp higher for men than for women, the latter being only 27%. Altogether, this evidence shows that not only do men have higher exposure to ChatGPT from their surroundings, but they also have more previous experience.

However, when relating the differences in exposure to the differences in use and skills regarding ChatGPT, Table 3.1 shows that controlling for our exposure measures does not explain a significant part of the gap generated in either of our main outcome variables (Column 5), where the gender gaps are still significant at the 5% level. This evidence suggests a limited role of exposure in explaining the gender gap.

3.3.4 Hypothetical Policy Experiment: Forbidding ChatGPT Would Widen the Gender Gap in Use

Given the current policy discussion around the world, we included in the survey a policy experiment to assess student responses to policies allowing or banning the use of ChatGPT. We rely on a hypothetical vignette experiment in which we randomize, at the student level, whether the professor in the hypothetical course they are taking allows or forbids ChatGPT use during the course, as described in Section 3.2.3.[23]

the fractions of friends using ChatGPT by either gender.

[23]Randomizing this type of policy in real institutions would not be attainable due to the importance that the issue of ChatGPT has for educators that will make the policy difficult to randomize, and the required sample sizes using randomization at the institution level.

Figure 3.3 plots the raw gender gaps in intended use when ChatGPT is allowed or forbidden. Intended use equals one if students state that they are likely or very likely to use ChatGPT during the course described in their randomly assigned scenario. When ChatGPT is allowed, over 80% of both men and women intend to use it. However, forbidding ChatGPT opens a large and statistically significant gap in intended use. While men respond to the ban with a decrease of 17.6 pp, from 87% intending to use when allowed to 70% when forbidden, the response of women is much larger at 37.9 pp, from 81% when allowed to 43% when forbidden.

The point estimate for the gender gap in intended use following specification 3.2 is in Table 3.1, Panel C, Column 1. When ChatGPT is allowed, the gap is 6.4 pp and not statistically significant. A gender gap in intended use equal to 20.3 pp opens up as a result of the forbidding policy (see interaction coefficient).

We note that intended use is higher for both men and women under the hypothetical scenario when the professor explicitly mentions that ChatGPT is allowed in the course than the baseline use that we documented in Section 3.3.1. Our take on this difference is that, to the best of our knowledge, the professors in the courses we recruited participants from did not make any explicit statements on whether or not ChatGPT should be used in the course.[24] When not explicitly stated, the default behavior is up to students' interpretation, and some of them may interpret no rule as not encouraged.

As with our previous results, we add different sets of control variables to identify whether the hypothesized factors influencing

[24]The professors in the master's course encouraged the use of ChatGPT but the sample coming from that course is very small.

use and skills can also be behind the differential policy responses. Unlike the gender gaps in high baseline use and prompt success rates, the responses to policies forbidding ChatGPT are not explained by any of the hypothesized factors influencing adoption. Columns 2-6 in Table 3.1, Panel C show that the coefficients remain similar in magnitude and statistically significant when adding different sets of controls independently or all controls at once.

Given the wide set of controls that we collected, our interpretation of the prevalence of the gender gap after adding the controls is that inclinations towards rule-following, obedience to authority, and trust in the professor's recommendations, play crucial roles in shaping the divergence in intended use. For instance, if female students are more predisposed to follow established rules and trust the guidance of authority figures, they may be more cautious or reserved in adopting new technologies, even if those technologies are intended to enhance learning experiences.

The crucial implication of these findings is the potential unintended consequences of banning ChatGPT in the classroom. Such a prohibition, intended to maintain a level playing field or address concerns by educators, might inadvertently contribute to a gender gap in AI adoption. By restricting access to this technology, female students could be placed at a disadvantage compared to their male peers, hindering their exposure to and familiarity with AI tools, as well as their prospects of success in a rapidly evolving labor market.

3.4 Discussion and Conclusion

We conducted a student survey at the Norwegian School of Economics to understand the current use of and proficiency in AI tools such as ChatGPT. We find large gender disparities in both dimensions, with male students being more likely to have already adopted and being more proficient at ChatGPT one year after its initial release. Importantly, policies banning ChatGPT in educational institutions would further widen the gender gap in use.

The implications arising from these findings could have profound significance for the career trajectories of female students. The observed gender disparity in ChatGPT usage raises concerns about potential barriers for women in accessing opportunities in a rapidly evolving job market that increasingly values AI proficiency. One potential constraint is that women who do not become proficient in AI tools refrain from applying to jobs that ask specifically for AI skills since their job decisions have been found to depend on features of the job or the workplace where women differ from men, i.e., competitiveness (Flory, Leibbrandt and List, 2015; Samek, 2019). Another constraint is that, even if they apply, women who do not acquire AI skills may find themselves at a disadvantage in the selection process for a growing number of positions that demand competence in this technology. In addition, once on the job, AI will likely drive differences in productivity and efficiency, leaving those that do not know how to use it properly behind. This could mean that women will miss out on promotions and career advancement if they lack AI skills. This discrepancy not only affects individual career prospects but also contributes to perpetuating gender imbalances in industries where AI proficiency is becoming a prerequisite, hindering diversity and inclusion efforts.

Our results also have wider implications regarding whether AI will reduce or exaggerate existing inequalities between high- and low-skill workers. The main idea is that labor demand is prone to decrease in tasks closely substitutable with the new technology, while it is inclined to rise in tasks that complement it (Brynjolfsson and Mitchell, 2017).[25] The results from early work suggest that AI can reduce inequalities between workers. An experiment with customer support agents shows that the low-skill agents using an AI tool that provides conversational guidance are able to increase their number of issues resolved per hour to the level of the high-skill agents (Brynjolfsson, Li and Raymond, 2023). Similarly, software developers with less developing experience benefit most from having access to the AI tool GitHub Copilot Peng et al. (2023). Our results suggest that, at least in the case of students developing general competencies at the undergraduate level, who can all be considered high-skill, those with higher admission grades have more to gain from AI tools because they are more successful at writing prompts. This implies that the potential benefit of AI tools hinges on the ability to interact with the AI, and that the top women in our sample, who have the lowest ChatGPT adoption rates, are those who may have more to lose from not becoming proficient at AI tools. In the future, the significance of prompting skills may diminish, as recent research has found that the newer version ChatGPT-4 can solve complex tasks in multiple domains with performance close to human level and without the need of special prompting (Bubeck et al., 2023), but for now it is a key skill.

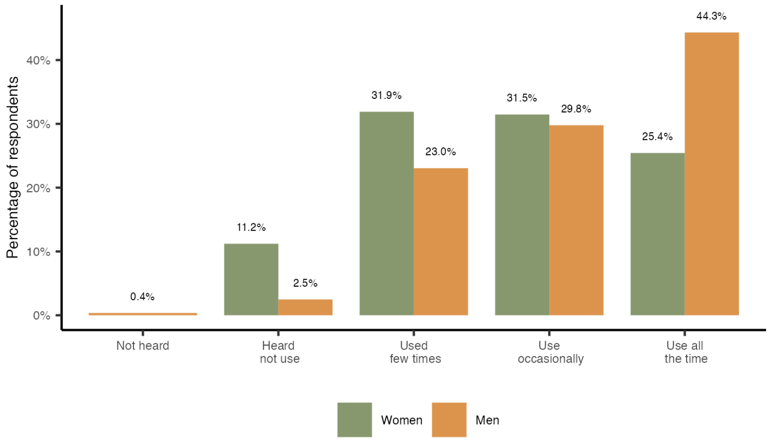
In considering our results' implications for student learning and non-AI skill development, it is crucial to address potential interference between ChatGPT use and other essential skills in edu-

[25] A study on the early look at the potential impact of large language models such as GPTs finds that around 80% of the US workforce will see that at least 10% of their tasks will be affected by LLMs. In addition, the early predictions suggest that 15% of the tasks can be completed faster while keeping the quality level (Eloundou et al., 2023).

cation and the labor market, such as critical thinking and problem-solving. We still lack evidence on whether AI adoption affects students' learning or grades, but if more traditional skills are easy to assess during exams or recruitment, students relying heavily on AI tools might find themselves at a disadvantage. Interestingly, the gender gaps we have identified could, in this context, offer advantages to women over men. As AI tools become more integral to work and daily life, influencing, for example, career choices (Reeder and Lee, 2022), the balance between traditional and AI skills in education and the labor market remains uncertain. Nevertheless, given the most likely scenario in which AI becomes increasingly important, it is in the hands of institutions to foster the development of both skill sets in a mutually beneficial manner. This is particularly crucial for female students, who, tending to adhere to rules, should be empowered with the confidence that they can adeptly develop and apply both types of skills, ensuring success in their chosen educational paths and careers.

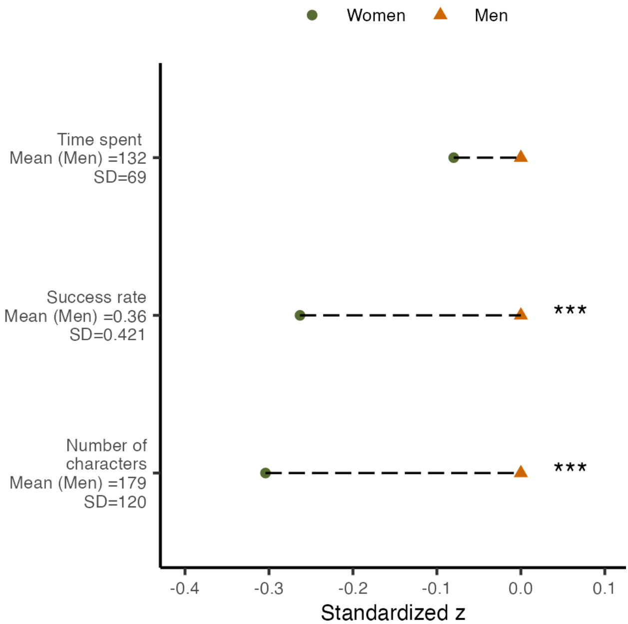
3.5 Main Figures

Figure 3.1: Baseline use



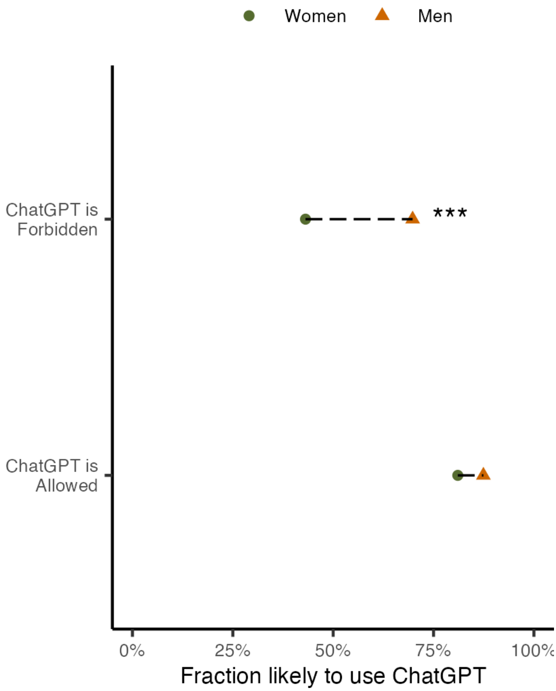
Notes: The figure shows a bar plot with the percentage of women and men indicating each answer to the question "How familiar are you with ChatGPT or similar tools?". Within gender the percentages across categories add up to 100%.

Figure 3.2: Prompt quality



Notes: The figure plots, by gender, the mean standardized values of three variables: time spent in the prompting task in seconds, success rate, and the number of characters of each prompt. All variables were standardized using the mean of men for each variable.

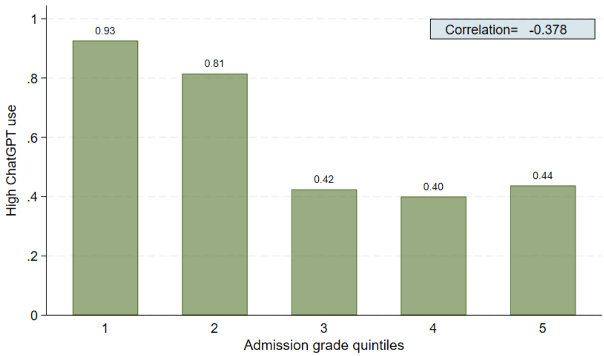
Figure 3.3: Policy responses



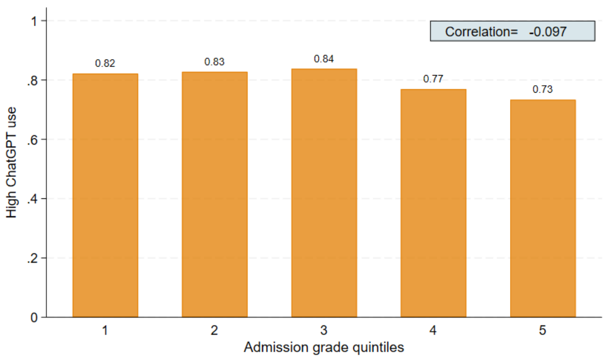
Notes: The figure shows, by gender, the fraction of participants that indicated "Somewhat likely" or "Very likely" to the question of how likely would they use ChatGPT in the hypothetical course presented in the vignette experiment. We show the estimates for the two randomly assigned scenarios: professor "forbids" and "allows" treatment.

Figure 3.4: Gender differences in baseline use by admission grade quintiles

(a) Women



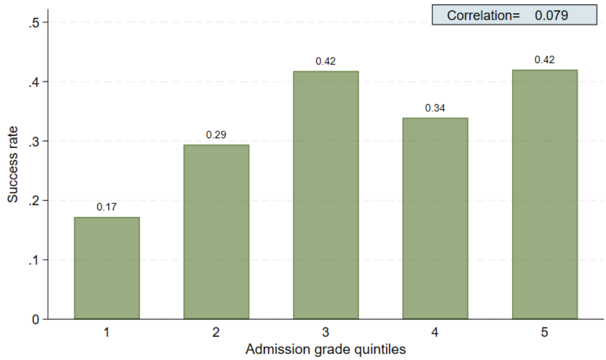
(b) Men



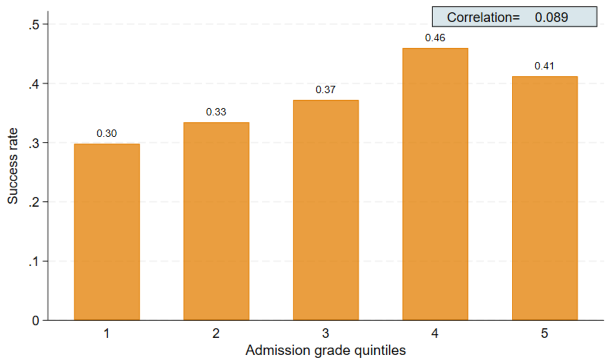
Notes: Panels (a) and (b) show the proportion of women and men, respectively, with high baseline use of ChatGPT across the self-reported admission grade quintiles (273 respondents).

Figure 3.5: Gender differences in prompt succcess by admission grade quintiles

(a) Women



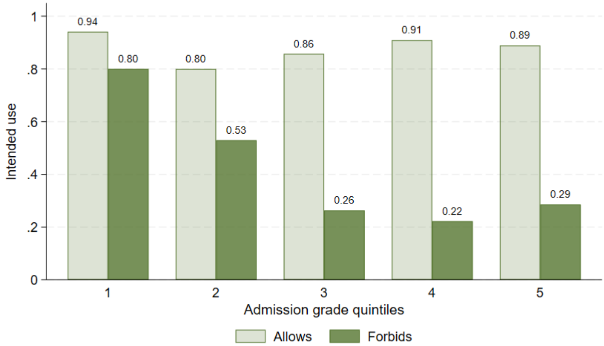
(b) Men



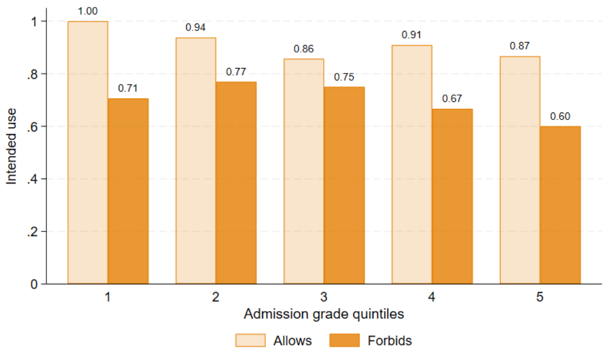
Notes: Panels (a) and (b) show the average success rate in the prompting task for women and men, respectively, across the self-reported admission grade quintiles (273 respondents).

Figure 3.6: Gender differences in policy response by admission grade quintiles

(a) Women

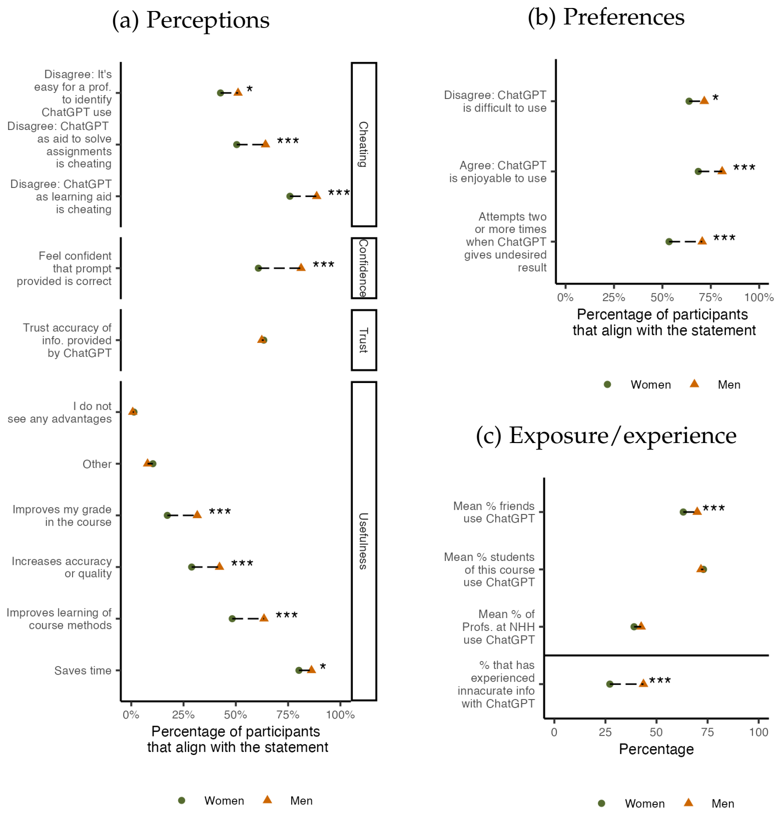


(b) Men



Notes: Panels (a) and (b) show the proportion of individuals who indicated likely intended use of ChatGPT in the vignette experiment for women and men, respectively, across the self-reported admission grade quintiles (273 respondents). In brighter colors is the intended use in the professor "allows" treatment, whereas in darker colors is the intended use in the "forbids" treatment.

Figure 3.7: Potential factors influencing use and skill: gender differences in attitudes



Notes: Panels (a) and (b) show, by gender, the percentage of participants whose answer aligns with each statement on the left of the corresponding graph. Panel (a) shows the results for the statements related to perceptions, while Panel (b) for the statements related to preferences. Panel (c) shows the variables capturing the exposure/experience channel, where the first three rows indicate, by gender, the mean estimate of the percentage of individuals that the participant believes use ChatGPT within the three indicated groups. The last row shows the percentage of participants that indicated to have experienced inaccurate information from ChatGPT. All gender gaps are raw estimates, without any controls. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

3.6 Main Tables

Table 3.1: Main results

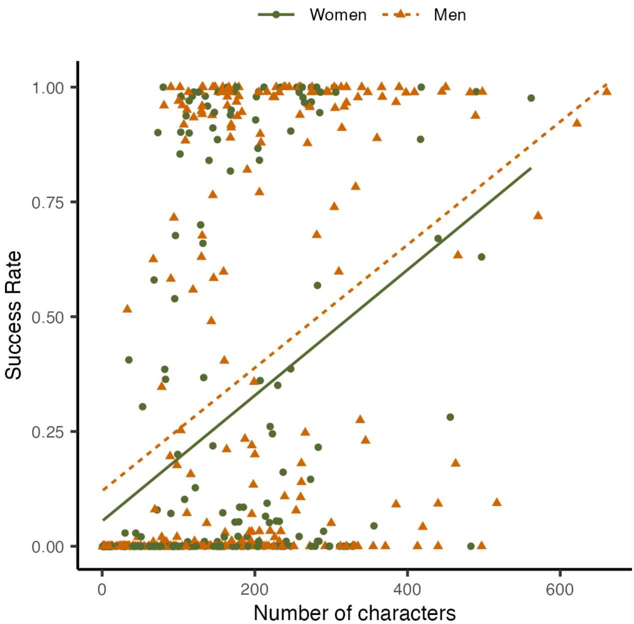
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Use ChatGPT occasionally or all the time (baseline use)						
Male	0.172***	0.098**	0.051	0.023	0.079**	0.012
	(0.042)	(0.043)	(0.037)	(0.043)	(0.037)	(0.040)
Constant	0.569***	1.362***	0.587***	0.127	0.193**	0.537
	(0.033)	(0.227)	(0.145)	(0.210)	(0.079)	(0.352)
Controls	None	Academic, risk & time	Preferences	Perceptions	Exposure/ experience	All
Observations	514	514	514	514	514	514
Panel B: Prompt success rate						
Male	0.111***	0.116***	0.101***	0.021	0.103***	0.026
	(0.037)	(0.043)	(0.038)	(0.038)	(0.039)	(0.044)
Constant	0.249***	-0.663	0.462**	0.061	0.405***	-0.772
	(0.026)	(0.403)	(0.200)	(0.249)	(0.084)	(0.558)
Controls	None	Baseline use, academic, risk & time	Preferences	Perceptions	Exposure/ experience	All
Observations	514	514	514	514	514	514
Panel C: Policy response (likely or very likely to use ChatGPT)						
Male	0.064	-0.047	-0.010	-0.037	0.010	-0.056
	(0.046)	(0.045)	(0.044)	(0.047)	(0.043)	(0.052)
ChatGPT forbidden	-0.379***	-0.384***	-0.396***	-0.397***	-0.355***	-0.391***
	(0.059)	(0.053)	(0.056)	(0.054)	(0.054)	(0.054)
Male × ChatGPT forbidden	0.203***	0.211***	0.221***	0.228***	0.201***	0.218***
	(0.076)	(0.072)	(0.072)	(0.070)	(0.070)	(0.072)
Constant	0.810***	1.160***	0.647***	0.547**	0.521***	1.040**
	(0.037)	(0.332)	(0.205)	(0.257)	(0.086)	(0.469)
Controls	None	Baseline use, academic, risk & time	Preferences	Perceptions	Exposure/ experience	All
Observations	514	514	514	514	514	514

Notes: Panels A and B show point estimates on gender differences in baseline use and success rate of the prompts written by students, respectively. Panel C shows point estimates on intended use from random variation on whether the professor allows or forbids the use of ChatGPT in a hypothetical course presented to the students. Each column title indicates what control variables are included in the regression. Column 1 presents raw estimates and Column 6 includes all controls added one by one in Columns 2-5. Academic controls include year in college, admission grade and an indicator for whether the admission grade is missing. Risk and time preferences are collected using the survey questions from the World Preferences Survey. Preferences include questions on whether students enjoy or find it difficult to use ChatGPT, as well as a measure of persistence in using ChatGPT. Perceptions include views on whether ChatGPT is equivalent to cheating, how useful it is, trust and confidence in own ChatGPT skills. Exposure/experience refers to what fraction of their friends, other students in their class and NHH professors use ChatGPT. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

Appendices

3.A Additional Figures and Tables

Figure A1: Relationship between success rate and number of characters



Notes: The scatterplot displays the relationship between the number of characters that students write in their prompt (x-axis) the success rate of the prompt (y-axis), for the full sample. The plot also provides the linear fits for both men (dashed) and women (solid), where the slope is of 0.13 for both.

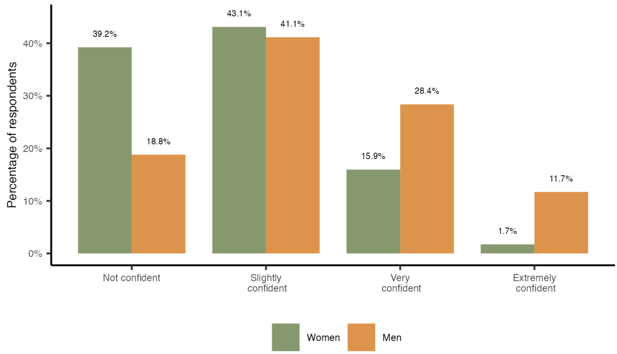
Table A1: Role of different perceptions in explaining the main results

	(1)	(2)	(3)	(4)	(5)
Panel A: Use ChatGPT occasionally or all the time (baseline use)					
Male	0.172***	0.100**	0.085**	0.171***	0.106**
	(0.042)	(0.043)	(0.041)	(0.042)	(0.043)
Constant	0.569***	0.527***	0.379***	0.572***	0.440***
	(0.033)	(0.193)	(0.069)	(0.098)	(0.044)
Controls	None	Cheating	Usefulness	Trust	Confidence
Observations	514	514	514	514	514
Panel B: Prompt success rate					
Male	0.111***	0.081**	0.099***	0.106***	0.041
	(0.037)	(0.040)	(0.037)	(0.037)	(0.037)
Constant	0.249***	0.415*	0.217***	0.133*	0.130***
	(0.026)	(0.217)	(0.063)	(0.069)	(0.028)
Controls	None	Cheating	Usefulness	Trust	Confidence
Observations	514	514	514	514	514
Panel C: Policy response (likely or very likely to use ChatGPT)					
Male	0.064	0.007	-0.003	0.063	0.022
	(0.046)	(0.045)	(0.045)	(0.047)	(0.047)
ChatGPT forbidden	-0.379***	-0.409***	-0.373***	-0.378***	-0.391***
	(0.059)	(0.055)	(0.055)	(0.059)	(0.058)
Male × ChatGPT forbidden	0.203***	0.225***	0.211***	0.206***	0.216***
	(0.076)	(0.072)	(0.072)	(0.076)	(0.075)
Constant	0.810***	0.781***	0.647***	0.850***	0.765***
	(0.037)	(0.210)	(0.072)	(0.101)	(0.048)
Controls	None	Cheating	Usefulness	Trust	Confidence
Observations	514	514	514	514	514

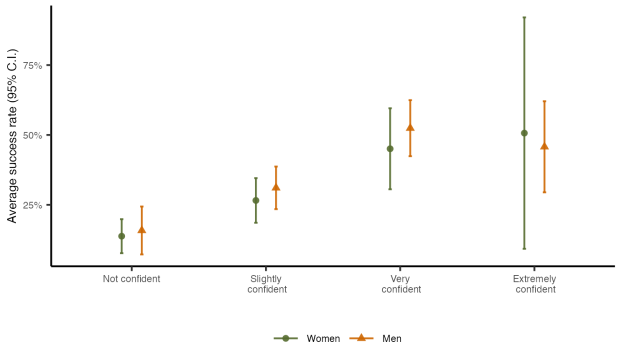
Notes: Panels A and B show point estimates on gender differences in baseline use and success rate of the prompts written by students, respectively. Panel C shows point estimates on intended use from random variation on whether the professor allows or forbids the use of ChatGPT in a hypothetical course presented to the students. Each column title indicates what control variables are included in the regression. Column 1 presents raw estimates and Columns 2-5 add a different set of perceptions variables as indicated at the bottom of the respective column. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Figure A2: Confidence in own prompt and success rates by level of confidence

(a) Confidence in own prompt



(b) Success rate by level of confidence



Notes: Panel (a) shows a bar plot with the percentage of women and men indicating each answer to the question "How confident do you feel that the query you just provided will make ChatGPT get the information you need?", which they answered after the prompting skills task. Panel (b) shows the average success rate for each answer option in the confidence question.

3.B Survey Questionnaire

Figure A3: Page 1. Consent

NHH

Welcome to this research project!

We very much appreciate your participation in this 5-minute survey. All data obtained is anonymous. Please make sure to always read the instructions carefully, **answer truthfully**, and do not leave the survey until reaching the end. Participation in this research study is completely voluntary. If you have questions regarding this study, you may contact: thechoicelab@nhh.no

Please click **Accept** below if you have understood the above and wish to participate in this study.

Accept

Figure A4: Page 2. Background characteristics

Are you from Norway?

Yes

No

To which gender identity do you most identify:

Male

Female

Non-binary / third gender

Prefer not to say

How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?

Completely unwilling to do so | Very willing to do so

0 1 2 3 4 5 6 7 8 9 10

In general, how willing are you to take risks?

Completely unwilling to take risks | Very willing to take risks

0 1 2 3 4 5 6 7 8 9 10

Figure A5: Page 3. "Allows" treatment

Imagine you are enrolled in a course on Environmental Policy and Economic Impact. This course explores the intersection of environmental regulations, economic incentives, and their effects on industry practices and sustainability. The professor explicitly allows the use of ChatGPT during coursework. It is an 8-week course with final evaluation given by a final home exam.

Given this scenario, how likely are you to use ChatGPT throughout the course?

- Very unlikely
- Somewhat unlikely
- Neither likely nor unlikely
- Somewhat likely
- Very likely

Given the scenario, how likely are you to use ChatGPT during the final exam?

- Very unlikely
- Somewhat unlikely
- Neither likely nor unlikely
- Somewhat likely
- Very likely

Figure A6: Page 4. “Forbids” treatment

Imagine you are enrolled in a course on Climate Change Economics. This course delves into the economic analysis of climate change, including the evaluation of mitigation strategies, adaptation costs, and international climate policy agreements. The professor explicitly forbids the use of ChatGPT during coursework. It is an 9-week course with final evaluation given by a final home exam.

Given this scenario, how likely are you to use ChatGPT throughout the course?

- Very unlikely
- Somewhat unlikely
- Neither likely nor unlikely
- Somewhat likely
- Very likely

Given the scenario, how likely are you to use ChatGPT during the final exam?

- Very unlikely
- Somewhat unlikely
- Neither likely nor unlikely
- Somewhat likely
- Very likely

Figure A7: Page 5. Prompting skills task

Do you know how to use ChatGPT?

Please take a moment to carefully check the image presented below.

Using the space provided, please write down the question that **you would ask to ChatGPT** to learn about the official name of this visual phenomenon. Remember ChatGPT cannot observe the image.

Figure A8: Page 6. Confidence question

How confident do you feel that the query you just provided will make ChatGPT get the information you need?

- Not confident at all
- Slightly confident
- Very confident
- Extremely confident

Figure A9: Page 7. ChatGPT use

How familiar are you with ChatGPT?

- I have not heard of it.
- I have heard of it but have not used it myself.
- I used it a few times.
- I use it occasionally.
- I use it regularly.

Figure A10: Page 8. Exposure and typical tasks

A survey conducted among university students in the US in the Spring of 2023 reports that 30% of students use ChatGPT for their schoolwork.

Now, for each of the groups below, please indicate the percentage of people you believe use ChatGPT:

0 20 40 60 80 100

Your group of friends

Students in this course

Professors at NHH

What type of tasks do you typically ask ChatGPT to help with? (Please select up to the most common three)

- Coding tasks
- Writing tasks
- Retrieving information
- Solving Math questions
- Other (Please specify)
- I don't use it

Figure A11: Page 9. Frequency by task

How frequently do you use ChatGPT for the following purposes:

Preparing for exams in a course:

- Never
- Occasionally
- Regularly

Solving home assignments for a course:

- Never
- Occasionally
- Regularly

Tasks unrelated to coursework:

- Never
- Occasionally
- Regularly

Tasks related to coursework:

- Never
- Occasionally
- Regularly

Figure A12: Page 10. Advantages (Usefulness)

What do you believe are the main advantages of using ChatGPT in coursework? (Please select all that apply.)

- Saves time.
- Increases accuracy or work quality.
- I do not see any advantages.
- Improves learning of course methods.
- Improves my grade in the course.
- Other (Please Specify)

Figure A13: Page 11.1 Agree/Disagree

How much do you agree with the following statements?

I think ChatGPT is enjoyable to use:

- Completely agree
- Somewhat agree
- Neither agree not disagree
- Somewhat disagree
- Completely disagree

Using ChatGPT as an aid to solve assignments in a course is equivalent to cheating:

- Completely agree
- Somewhat agree
- Neither agree not disagree
- Somewhat disagree
- Completely disagree

Figure A14: Page 11.2 Agree/Disagree

Using ChatGPT as a learning aid in a course is equivalent to cheating:

Completely agree

Somewhat agree

Neither agree not disagree

Somewhat disagree

Completely disagree

I think ChatGPT is difficult to use:

Completely agree

Somewhat agree

Neither agree not disagree

Somewhat disagree

Completely disagree

Figure A15: Page 11.3 Agree/Disagree

It is easy for professors to identify if a student has used ChatGPT:

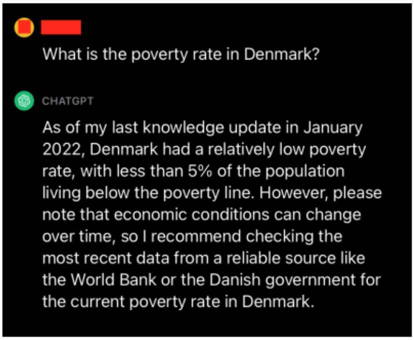
- Completely agree
- Somewhat agree
- Neither agree not disagree
- Somewhat disagree
- Completely disagree

ChatGPT is mostly a tool complementing skills rather than substituting effort:

- Completely agree
- Somewhat agree
- Neither agree not disagree
- Somewhat disagree
- Completely disagree

Figure A16: Page 12. Trust accuracy

Below is a screen capture of a query made to ChatGPT, along with the response it provided.


[redacted]

What is the poverty rate in Denmark?

CHATGPT

As of my last knowledge update in January 2022, Denmark had a relatively low poverty rate, with less than 5% of the population living below the poverty line. However, please note that economic conditions can change over time, so I recommend checking the most recent data from a reliable source like the World Bank or the Danish government for the current poverty rate in Denmark.

Based on this response from ChatGPT, how much do you trust that the **poverty rate reported** is accurate?

- Completely trust
- Somewhat trust
- Somewhat distrust
- Completely distrust

Figure A17: Page 13. Persistence and inaccuracy

If ChatGPT does not provide the desired answer on your first attempt, how many additional attempts do you typically make?

- None, I move on.
- One more try.
- Two more tries.
- I keep trying until satisfied.
- I don’t use it.

Have you ever received inaccurate or misleading information from ChatGPT?

- Yes, many times.
- Yes, few times.
- No, never.
- I don't use it.

Figure A18: Page 14. Subscription and admission grade

Do you have a subscription for using ChatGPT or other similar AI platforms?

No.

Yes, I have the free subscription.

Yes, I have the paid subscription.

What was your admission grade at NHH? Please provide an estimate if you don't remember the exact grade (or NA if you don't have):

Bibliography

Achard, Pascal, and Sigrid Suetens. 2023. "The causal effect of ethnic diversity on support for redistribution and the role of discrimination." *Journal of Economic Surveys*.

Akerlof, George A, and Rachel E Kranton. 2000. "Economics and identity." *The Quarterly Journal of Economics*, 115(3): 715–753.

Alan, Sule, Ceren Baysan, Mert Gumren, and Elif Kubilay. 2021. "Building social cohesion in ethnically mixed schools: An intervention on perspective taking." *The Quarterly Journal of Economics*, 136(4): 2147–2194.

Alesina, Alberto, and Marco Tabellini. 2022. "The political effects of immigration: Culture or economics?" National Bureau of Economic Research.

Allport, Gordon. 1954. "The nature of prejudice."

Almås, Ingvild, Orazio Attanasio, and Pamela Jervis. 2023. "Economics and Measurement: New measures to model decision making." National Bureau of Economic Research.

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew D Shapiro, and Christopher Tonetti. 2020. "Long-term-care utility and late-in-life saving." *Journal of Political Economy*, 128(6): 2375–2451.

Anderberg, Dan, Christina Felfe, Gordon Dahl, Helmut Rainer, and Thomas Siedler. 2023. "Diversity and Discrimination in the Classroom." National Bureau of Economic Research.

Andreoni, James, and B Douglas Bernheim. 2009. "Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects." *Econometrica*, 77(5): 1607–1636.

Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart. 2022. "Subjective models of the macroeconomy: Evidence from experts and representative samples." *The Review of Economic Studies*, 89(6): 2958–2991.

Aron, Arthur, Elaine N Aron, and Danny Smollan. 1992. "Inclusion of other in the self scale and the structure of interpersonal closeness." *Journal of personality and social psychology*, 63(4): 596.

Atkin, David, Eve Colson-Sihra, and Moses Shayo. 2021. "How do we choose our identity? a revealed preference approach using food consumption." *Journal of Political Economy*, 129(4): 1193–1251.

Avery, Mallory, Andreas Leibbrandt, and Joseph Vecci. 2023. "Does Artificial Intelligence Help or Hurt Gender Diversity? Evidence from Two Field Experiments on Recruitment in Tech." *Evidence from Two Field Experiments on Recruitment in Tech (February 14, 2023)*.

Avilova, Tatyana, and Claudia Goldin. 2018. "What Can UWE Do For Economics?" Vol. 108, 186–90.

Awad, Edmond, Loukas Balafoutas, Li Chen, Edwin Ip, and Joe Vecci. 2023. "Artificial Intelligence and Debiasing in Hiring: Impact on Applicant Quality and Gender Diversity." *Available at SSRN*.

Benjamin, Daniel J, James J Choi, and A Joshua Strickland. 2010. "Social identity and preferences." *American Economic Review*, 100(4): 1913–1928.

Bernheim, B Douglas, Daniel Björkegren, Jeffrey Naecker, and Michael Pollmann. 2022. "Causal inference from hypothetical evaluations." National Bureau of Economic Research.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." *American Economic Review*, 94(4): 991–1013.

Bicchieri, Cristina, Eugen Dimant, Simon Gächter, and Daniele Nosenzo. 2022. "Social proximity and the erosion of norm compliance." *Games and Economic Behavior*, 132: 59–72.

Bimber, Bruce. 2000. "Measuring the gender gap on the Internet." *Social science quarterly*, 868–876.

Blalock Jr, Hubert M. 1967. "Status inconsistency, social mobility, status integration and structural effects." *American Sociological Review*, 790–801.

Bonomi, Giampaolo, Nicola Gennaioli, and Guido Tabellini. 2021. "Identity, beliefs, and political conflict." *The Quarterly Journal of Economics*, 136(4): 2371–2411.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016*a*. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016*b*. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.

Boustan, Leah Platt. 2013. "Racial residential segregation in American cities." National Bureau of Economic Research.

Brañas-Garza, Pablo, Diego Jorrat, Antonio M Espín, and Angel Sánchez. 2023. "Paid and hypothetical time preferences are the same: Lab, field and online evidence." *Experimental Economics*, 26(2): 412–434.

Brañas-Garza, Pablo, Lorenzo Estepa-Mohedano, Diego Jorrat, Victor Orozco, and Ericka Rascón-Ramírez. 2021. "To pay or not to pay: Measuring risk preferences in lab and field." *Judgment and Decision Making*, 16(5): 1290–1313.

Brookings. 2020. "Even as metropolitan areas diversify, white Americans still live in mostly white neighborhoods."

Brynjolfsson, Erik, and Gabriel Unger. 2023. "The Macroeconomics of Artificial Intelligence." *https://www.imf.org/en/Publications/fandd/issues/2023/12/Macroeconomics-of-artificial-intelligence-Brynjolfsson-Unger*, [Online; accessed 05-December-2023].

Brynjolfsson, Erik, and Tom Mitchell. 2017. "What can machine learning do? Workforce implications." *Science*, 358(6370): 1530–1534.

Brynjolfsson, Erik, Danielle Li, and Lindsey R Raymond. 2023. "Generative AI at work." National Bureau of Economic Research.

Bubeck, Sébastien, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. 2023. "Sparks of artificial general intelligence: Early experiments with gpt-4." *arXiv preprint arXiv:2303.12712*.

Buser, Thomas, Leonie Gerhards, and Joël Van Der Weele. 2018. "Responsiveness to feedback as a personal trait." *Journal of Risk and Uncertainty*, 56: 165–192.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, competitiveness, and career choices." *The Quarterly Journal of Economics*, 129(3): 1409–1447.

Buser, Thomas, Noemi Peter, and Stefan C Wolter. 2017. "Gender, competitiveness, and study choices in high school: Evidence from Switzerland." *American Economic Review*, 107(5): 125–130.

Cappelen, Alexander W, Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden. 2007. "The pluralism of fairness ideals: An experimental approach." *American Economic Review*, 97(3): 818–827.

Cappelen, Alexander W, Benjamin Enke, and Bertil Tungodden. 2023. "Moral universalism: Global evidence." National Bureau of Economic Research.

Carvajal, Daniel. 2024. "Exposure to diversity, social proximity and ingroup bias."

Charness, Gary, and Matthew Rabin. 2002. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics*, 117(3): 817–869.

Charness, Gary, and Yan Chen. 2020. "Social identity, group behavior, and teams." *Annual Review of Economics*, 12: 691–713.

Chen, Si, and Hannah Schildberg-Hörisch. 2019. "Looking at the bright side: The motivational value of confidence." *European Economic Review*, 120: 103302.

Chen, Yan, and Sherry Xin Li. 2009. "Group identity and social preferences." *American Economic Review*, 99(1): 431–457.

Cikara, Mina, Vasiliki Fouka, and Marco Tabellini. 2022. "Hate crime towards minoritized groups increases as they increase in sized-based rank." *Nature Human Behaviour*, 6(11): 1537–1544.

Cimpian, Joseph R, Taek H Kim, and Zachary T McDermott. 2020. "Understanding persistent gender gaps in STEM." *Science*, 368(6497): 1317–1319.

Citrin, Jack, and John Sides. 2008. "Immigration and the imagined community in Europe and the United States." *Political Studies*, 56(1): 33–56.

CNBC. 2023. "Nine in 10 companies want employees with ChatGPT skills." *https://www.cnbc.com/2023/09/01/more-companies-see-chatgpt-training-as-a-hot-job-perk-for-office-wo html*, de Visé, Daniel and Klar, Rebecca. [Online; accessed 12-December-2023].

Coffman, Katherine Baldiga. 2014. "Evidence on self-stereotyping and the contribution of ideas." *The Quarterly Journal of Economics*, 129(4): 1625–1660.

Coffman, Katherine B, Manuela R Collis, and Leena Kulkarni. 2023. "Whether to Apply." *Management Science*.

Condra, Luke N, and Sera Linardi. 2019. "Casual contact and ethnic bias: Experimental evidence from Afghanistan." *The Journal of Politics*, 81(3): 1028–1042.

Cortés, Patricia, Jessica Pan, Ernesto Reuben, Laura Pilossoph, and Basit Zafar. 2022. "Gender differences in job search and the earnings gap: Evidence from the field and lab." National Bureau of Economic Research.

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

Direktoratet for høyere utdanning og kompetanse. 2023. "Hovedopptak til høyere utdanning ved universiteter og høgskoler." `https://hkdir.no/rapporter-undersokelser-og-statistikk/hovedopptak-til-hoyere-utdanning-ved-universiteter-og-hogskoler` [Online; accessed 12-December-2023].

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. "GPTs are GPTs: An early look at the labor market impact potential of large language models." *arXiv preprint arXiv:2303.10130*.

Engel, Christoph. 2011. "Dictator games: A meta study." *Experimental economics*, 14: 583–610.

Enke, Benjamin, Ricardo Rodríguez-Padilla, and Florian Zimmermann. 2022. "Moral universalism and the structure of ideology." *The Review of Economic Studies*.

Esponda, Ignacio, Ryan Oprea, and Sevgi Yuksel. 2023. "Seeing What is Representative*." *The Quarterly Journal of Economics*, qjad020.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global evidence on economic preferences." *The Quarterly Journal of Economics*, 133(4): 1645–1692.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde. 2023. "The preference survey module: A validated instrument for measuring risk, time, and social preferences." *Management Science*, 69(4): 1935–1950.

Fehr, Ernst, and Klaus M Schmidt. 1999. "A theory of fairness, competition, and cooperation." *The quarterly journal of economics*, 114(3): 817–868.

Fehr, Ernst, and Urs Fischbacher. 2003. "The nature of human altruism." *Nature*, 425(6960): 785–791.

Flory, Jeffrey A, Andreas Leibbrandt, and John A List. 2015. "Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions." *The Review of Economic Studies*, 82(1): 122–155.

Fouka, Vasiliki, and Marco Tabellini. 2022. "Changing In-Group Boundaries: The Effect of Immigration on Race Relations in the United States." *American Political Science Review*, 116(3): 968–984.

Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini. 2022. "From immigrants to Americans: Race and assimilation during the Great Migration." *The Review of Economic Studies*, 89(2): 811–842.

Fryer Jr, Roland G, and Steven D Levitt. 2004. "Understanding the black-white test score gap in the first two years of school." *Review of Economics and Statistics*, 86(2): 447–464.

Gächter, Simon, Chris Starmer, and Fabio Tufano. 2015. "Measuring the closeness of relationships: a comprehensive evaluation of the'Inclusion of the Other in the Self'scale." *PloS one*, 10(6): e0129478.

Gächter, Simon, Chris Starmer, and Fabio Tufano. 2022. "Measuring "group cohesion" to reveal the power of social relationships in team production." *The Review of Economics and Statistics*, 1–45.

Goette, Lorenz, and Egon Tripodi. 2021. "Social influence in prosocial behavior: evidence from a large-scale experiment." *Journal of the European Economic Association*, 19(4): 2373–2398.

Goldin, Claudia. 2015. "Gender and the undergraduate economics major: Notes on the undergraduate economics major at a highly selective liberal arts college." *manuscript, April*, 12.

Gold, Jason M. 2014. "Information processing correlates of a size-contrast illusion." *Frontiers in Psychology*, 5: 142.

Grossman, Gene M, and Elhanan Helpman. 2018. "Identity politics and trade policy." *The Review of Economic Studies*.

Grossman, Gene M, and Elhanan Helpman. 2021. "Identity politics and trade policy." *The Review of Economic Studies*, 88(3): 1101–1126.

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. "Validating vignette and conjoint survey experiments against real-world behavior." *Proceedings of the National Academy of Sciences*, 112(8): 2395–2400.

Hangartner, Dominik, Elias Dinas, Moritz Marbach, Konstantinos Matakos, and Dimitrios Xefteris. 2019. "Does exposure to the refugee crisis make natives more hostile?" *American Political Science Review*, 113(2): 442–455.

Heckman, James J, and Rodrigo Pinto. 2015. "Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs." *Econometric reviews*, 34(1-2): 6–31.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack. 2020. "Overconfidence and prejudice."

Hellerstein, Judith K, and David Neumark. 2008. "Workplace segregation in the United States: Race, ethnicity, and skill." *The Review of Economics and Statistics*, 90(3): 459–477.

Hett, Florian, Mario Mechtel, and Markus Kröll. 2020. "The Structure and Behavioral Effects of Revealed Social Identity Preferences*." *The Economic Journal*. ueaa055.

Hirshman, Samuel D, and Alexander Willén. 2022. "Does Increasing Risk Widen Gender Gaps?" *NHH Dept. of Economics Discussion Paper*, , (20).

Hjerm, Mikael. 2007. "Do numbers really count? Group threat theory revisited." *Journal of Ethnic and Migration Studies*, 33(8): 1253–1275.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. "Identification, inference and sensitivity analysis for causal mediation effects." *Statistical science*, 51–71.

Kahneman, Daniel, and Amos Tversky. 1972. "Subjective probability: A judgment of representativeness." *Cognitive psychology*, 3(3): 430–454.

Kahneman, Daniel, and Amos Tversky's. 1972. "Subjective probability: A judgment of representativeness." *Cognitive psychology*, 3(3): 430–454.

Karimi, Fariba, Mathieu Génois, Claudia Wagner, Philipp Singer, and Markus Strohmaier. 2018. "Homophily influences ranking of minorities in social networks." *Scientific reports*, 8(1): 11077.

Kessler, Judd B, Corinne Low, and Colin D Sullivan. 2019. "Incentivized resume rating: Eliciting employer preferences without deception." *American Economic Review*, 109(11): 3713–3744.

Klor, Esteban F, and Moses Shayo. 2010. "Social identity and preferences over redistribution." *Journal of Public Economics*, 94(3-4): 269–278.

Kranton, Rachel, Matthew Pease, Seth Sanders, and Scott Huettel. 2020. "Deconstructing bias in social preferences reveals groupy and not-groupy behavior." *Proceedings of the National Academy of Sciences*, 117(35): 21185–21193.

Kugler, Adriana D, Catherine H Tinsley, and Olga Ukhaneva. 2021. "Choice of majors: Are women really different from men?" *Economics of Education Review*, 81: 102079.

Landaud, Fanny, Éric Maurin, Barton Willage, and Alexander Willén. (accepted) 2023. "The Value of a High School GPA." *Review of Economics and Statistics*.

Lane, Tom. 2016. "Discrimination in the laboratory: A meta-analysis of economics experiments." *European Economic Review*, 90: 375–402.

Leider, Stephen, Markus M Möbius, Tanya Rosenblat, and Quoc-Anh Do. 2009. "Directed altruism and enforced reciprocity in social networks." *The Quarterly Journal of Economics*, 124(4): 1815–1851.

Li, Sherry Xin. 2020. "Group identity, ingroup favoritism, and discrimination." *Handbook of Labor, Human Resources and Population Economics*, 1–28.

Lo, Chung Kwan. 2023. "What is the impact of ChatGPT on education? A rapid review of the literature." *Education Sciences*, 13(4): 410.

Lowe, Matt. 2021. "Types of contact: A field experiment on collaborative and adversarial caste integration." *American Economic Review*, 111(6): 1807–1844.

Luttmer, Erzo FP. 2001. "Group loyalty and the taste for redistribution." *Journal of Political Economy*, 109(3): 500–528.

Möbius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat. 2022. "Managing self-confidence: Theory and experimental evidence." *Management Science*, 68(11): 7793–7817.

Mousa, Salma. 2020. "Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq." *Science*, 369(6505): 866–870.

Nathan, Noah L, and Melissa L Sands. 2023. "Context and contact: Unifying the study of environmental effects on politics." *Annual Review of Political Science*, 26.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do women shy away from competition? Do men compete too much?" *The Quarterly Journal of Economics*, 122(3): 1067–1101.

Niederle, Muriel, and Lise Vesterlund. 2011. "Gender and competition." *Annu. Rev. Econ.*, 3(1): 601–630.

Noy, Shakked, and Whitney Zhang. 2023. "Experimental evidence on the productivity effects of generative artificial intelligence." *Science*.

OECD. 2018. "Bridging the digital gender divide: Include, upskill, innovate." *OECD*.

Oh, Suanna. 2023. "Does identity affect labor supply?" *American Economic Review*, 113(8): 2055–2083.

Ost, Ben. 2010. "The role of peers and grades in determining major persistence in the sciences." *Economics of Education Review*, 29(6): 923–934.

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. 2023. "The impact of ai on developer productivity: Evidence from github copilot." *arXiv preprint arXiv:2302.06590*.

Pisanelli, Elena. 2022. "A new turning point for women: artificial intelligence as a tool for reducing gender discrimination in hiring." *Available at SSRN 4254965*.

Pottie-Sherman, Yolande, and Rima Wilkes. 2017. "Does size really matter? On the relationship between immigrant group size and anti-immigrant prejudice." *International Migration Review*, 51(1): 218–250.

Rao, Gautam. 2019. "Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools." *American Economic Review*, 109(3): 774–809.

Rask, Kevin, and Jill Tiefenthaler. 2008. "The role of grade sensitivity in explaining the gender imbalance in undergraduate economics." *Economics of Education Review*, 27(6): 676–687.

Reeder, Kristen, and Hwan Lee. 2022. "Impact of artificial intelligence on US medical students' choice of radiology." *Clinical imaging*, 81: 67–71.

Robson, Matthew. 2021. "Inequality aversion, self-interest and social connectedness." *Journal of Economic Behavior & Organization*, 183: 744–772.

Samek, Anya. 2019. "Gender differences in job entry decisions: A university-wide field experiment." *Management Science*, 65(7): 3272–3281.

Schlueter, Elmar, and Peer Scheepers. 2010. "The relationship between outgroup size and anti-outgroup attitudes: A theoretical synthesis and empirical test of group threat-and intergroup contact theory." *Social Science Research*, 39(2): 285–295.

Shayo, Moses. 2020. "Social identity and economic policy." *Annual Review of Economics*, 12: 355–389.

Stango, Victor, and Jonathan Zinman. 2023. "We Are All Behavioural, More, or Less: A Taxonomy of Consumer Decision-Making." *The Review of Economic Studies*, 90(3): 1470–1498.

Steinmayr, Andreas. 2020. "Contact versus exposure: Refugee presence and voting for the far-right." *Review of Economics and Statistics*, 1–47.

Tajfel, Henri, and John C Turner. 1979. "An Integrative Theory of Intergroup Conflict." *The Social Psychology of Intergroup Relations*, 33(47): 74.

Titchener, Edward Bradford. 1901. *Experimental psychology: A manual of laboratory practice, vol. I: Qualitative experiments.* New York: The Macmillan Company.

Turner, John C, Michael A Hogg, Penelope J Oakes, Stephen D Reicher, and Margaret S Wetherell. 1987. *Rediscovering the social group: A self-categorization theory.* Basil Blackwell.

Ugalde, Maria Paola. 2022. "Gender, Grade Sensitivity, and Major Choice."

Williams, Joan C. 2014. "Double jeopardy? An empirical study with implications for the debates over implicit bias and intersectionality." *Harvard Journal of Law & Gender*, 37: 185.

WSJ. 2023. "More companies see ChatGPT training as a hot job perk for office workers." `https://www.wsj.com/tech/ai/talking-to-chatbots-is-now-a-200k-job-so-i-applied-258bd5f0`, Stern, Joanna. [Online; accessed 12-December-2023].

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.

Zizzo, Daniel John, and Andrew J Oswald. 2001. "Are people willing to pay to reduce others' incomes?" *Annales d'Economie et de Statistique*, 39–65.