Race and Redistribution in the United States: An Experimental Analysis[*]

Jesper Akesson, Robert W. Hahn,
Robert D. Metcalfe and Itzhak Rasooly[†]

September 7, 2022

Abstract

Scholars have suggested that White American support for welfare is related to beliefs about the racial composition of welfare recipients. While a host of observational studies lend credence to this view, it has not yet been tested using the tools of randomized inference. In this study, we do this by conducting two incentive-compatible experiments ($n = 9,775$) in which different participants are randomly given different signals about the share of welfare recipients who identify as Black and White. Our analysis yields four main findings. First, 86% of respondents greatly overestimate the share of welfare recipients who are Black, with the average respondent overestimating this by almost a factor of two. Second, White support for welfare is inversely related to the proportion of welfare recipients who are Black—a causal claim that we establish using treatment assignment as an instrument for beliefs about the racial composition of welfare recipients. Third, just making White participants think about the racial composition of welfare recipients reduces their support for welfare. Fourth, providing White respondents with accurate information about the racial composition of welfare recipients (relative to not receiving any information) does not significantly influence their support for welfare.

Keywords: redistribution, race, field experiment, beliefs.
JEL codes: D90, C93

1 Introduction

Scholars have suggested that White American support for welfare is related to beliefs about the racial composition of welfare recipients (Quadagno et al., 1994; Ribar and Wilhelm, 1999; Luttmer, 2001; Alesina and Glaeser, 2004; Alesina and Ferrara, 2005; Lee and Roemer, 2006; Gilens, 2009; Alesina and Giuliano, 2011). As Alesina et al. (2001) put it, "Racial animosity in the US makes redistribution to the poor, who are disproportionately black, unappealing to many voters." Moreover, several observational studies support this view. For example, Luttmer (2001) finds an association between individual support for welfare and the fraction of local welfare recipients who share their ethnicity. Similarly, Alesina et al. (2001) note that the US spends much less on welfare than other more ethnically homogeneous countries.

In this paper, we provide what we believe is the first *causal* test of the relationship between White American support for welfare and the racial composition of welfare recipients. We do this by conducting two large-scale experiments ($n = 9,775$), in which individuals are randomly allocated to a control group that receives no information, a 'low' signal group, or a 'high' signal group. The low signal group receives information suggesting that a relatively low share of welfare recipients is Black (20% in experiment 1 and 8% in experiment 2). In contrast, the high signal group receives information suggesting that a relatively high share of welfare recipients is Black (26% in experiment 1 and 52% in experiment 2). Importantly, we are able to present different participants with different signals without employing deception. To do this, we examine the racial composition of welfare recipients for different years and months of birth, and only give participants data about particular sub-samples. Since the numbers can vary dramatically for such sub-samples, this allows us to present different individuals with very different estimates.[1]

In addition, we elicit beliefs about the racial composition of welfare recipients using an incentive-compatible procedure. Our main outcome variable—support for welfare—is measured by asking participants if they would like to donate money to nonprofit organizations that work to increase or decrease welfare spending in the US. By using treatment assignment as an instrument for individual beliefs about the racial composition of welfare recipients, it is possible to obtain Local Average Treatment Effect (LATE) estimates of the effect of such beliefs on the support for welfare using two-stage least-squares regression.

The experiments contain several other features aimed at understanding the link between

---

[1]We explicitly tell participants that the estimates they view are based on a particular sub-sample. Our data come from the Survey of Income and Program Participation (USCB, 2018).

race and welfare support. For example, we embed a priming experiment within the control group of our second experiment by randomly varying whether we measure support for welfare before or after eliciting participants' beliefs about the racial composition of welfare recipients. This allows us to investigate whether simply prompting White individuals to think about the racial composition of recipients alters their attitudes towards welfare. We also collect data on a number of variables that may potentially mediate—and help explain—the effect of beliefs about the racial composition of welfare recipients on welfare support.

Our empirical analysis yields four main findings. First, participants dramatically overestimate the fraction of welfare recipients who are Black.[2] More specifically, the average White respondent in our sample estimates that 38% of welfare recipients are Black, and the average Black respondent estimates the figure at 35%. These numbers greatly exceed the true value, which is approximately 21% (USCB, 2018). The finding that people overestimate the share of welfare recipients that are Black matches up with prior surveys on this issue (*e.g.*, Arthur and Edwards-Levy (2018)). It also appears to be driven, at least in part, by widespread overestimation of the share of the population that is Black.[3]

Second, we estimate how beliefs about the share of welfare recipients that is Black influences support for welfare. To do this, we use the randomized treatment assignment as an instrument for beliefs about the racial composition of welfare recipients (while excluding the control groups from our analyses). We argue that the instrument is likely to be exogenous because the only channel through which treatment assignment can plausibly influence support for welfare is via beliefs (point estimates) about the racial composition of recipients.[4] In addition, our instruments are highly informative, with F-statistics of 65 and 11253 in the first and second experiments respectively.

The instrumental variables (IV) analysis leads to our second finding: higher beliefs about the share of welfare recipients who are Black reduces White respondents' support for

---

[2]In our study, we define welfare as the following four programs: Temporary Assistance for Needy Families (TANF), Medicaid, the Supplemental Nutrition Assistance Program (SNAP), Housing Assistance, and Supplemental Security Income (SSI).

[3]Previous research suggests that some individuals and politicians believe that it is in their interest to perpetuate false beliefs, narratives, and stereotypes against Black Americans, which may help explain why participants hold incorrect beliefs (Glaeser, 2005; Steele, 2011; Bordalo et al., 2016). It may also result from long-term systemic racism (Feagin, 2013).

[4]Omitting the control group is crucial for this argument since subjects in the control group may have been differentially primed to think about race relative to subjects in the treatment groups, leading to a violation of the exclusion restriction. Moreover, we do not find that participants' confidence in their beliefs differs between the two treatment groups, while those in the treatment groups are more confident than those in the control group.

welfare. For example, in the first experiment, we find that a 1 percentage point increase in beliefs about the share that is Black leads to a 0.7 percentage point increase in the share that donate to the anti-welfare nonprofit ($p = 0.023$). Moreover, a 10 percentage point shift in beliefs has as large an effect on the share that donate to the anti-welfare nonprofit as the difference in donation rates between conservative and liberal respondents.

We then compare the relationship between beliefs about the share of welfare recipients who are Black and respondents' support for welfare estimated using IV to the relationship estimated using ordinary least squares regressions (OLS). Strikingly, we find that the (causal) IV estimates are close to identical to the (possibly biased) associations obtained using simple OLS. Moreover, the results from the instrumental variables analyses are robust to dropping apparently less attentive subjects, re-weighting the sample to match the demographics of the US population, and several other variations on our main specification. Further, the results do not differ significantly across different White sub-groups. Taken together, these findings provide strong support for both the internal and external validity of our main results.

In contrast to the result for White respondents, we do not find a significant relationship between beliefs about the racial composition of welfare recipients and support for welfare among Black respondents. This may simply be because our study is under-powered to detect such effects among the smaller sub-sample of Black respondents. However, it is also consistent with the possibility that Black Americans, unlike White Americans, do not substantially take race into account when forming attitudes towards welfare. This second possibility echoes previous findings that the impact of race-related information depends on the race of the recipient (see, *e.g.*, Washington (2006)).

We also discuss mechanisms that might explain our main result—that White Americans support welfare less when they believe that the share of welfare recipients that is Black is higher. In some sense, the explanation for this is transparent: White Americans (on average) prefer welfare spending that goes to White as opposed to Black Americans. Nonetheless, one can ask what in turn explains this underlying preference. We are able to shed some light on this issue. More specifically, we find evidence that higher beliefs about the share of welfare recipients that is Black lead to lower assessments about the perceived worthiness of welfare recipients—and the worthiness of recipients has in past studies been found to influence people's willingness to donate money (Fong and Luttmer,

2009).[5] However, one should be careful not to over-interpret this finding, not least because individuals may use claims about worthiness as an excuse for underlying racial animus (Bursztyn et al., 2020c).[6]

Next, we turn our attention to the priming treatment embedded within experiment 2, which yields our third finding: simply making White respondents think about the racial composition of welfare recipients makes them less supportive of welfare. More specifically, we find that answering the question about the share of respondents that are from different racial groups before being asked whether they would donate—as opposed to after—reduces the share that donate to the pro-welfare nonprofit by 4.9 percentage points ($p = 0.045$). This effect is fairly large relative to the effect of beliefs, and is consistent with the results in Alesina et al. (2018). For comparison, increasing participants' perception about the share of welfare recipients that are Black by 1 percentage point decreases the share that donate to the pro-welfare nonprofit by 0.1 percentage points in experiment 2 ($p = 0.037$).[7]

Finally, our fourth finding is that providing individuals with accurate information about the racial composition of welfare recipients does not seem to alter their attitudes toward welfare (relative to not receiving any information). This might seem surprising since providing such information shifts beliefs about the prevalence of Black welfare recipients downwards, which in turn might be expected to increase support for welfare (in light of our earlier result). However, one should recognize that information provision also makes the issue of race more salient, which in turn could be expected to decrease support for welfare. Given these conflicting effects, it is not unexpected that information provision does not in itself have a large effect on welfare attitudes in either direction.[8]

Our results build on a number of important literatures. Most importantly, there is a

---

[5]In contrast to our findings regarding worthiness, we do not find that our treatments alter perceptions about the efficacy of welfare spending, which has been suggested as another key determinant of welfare support (Cook and Barrett, 1992; Fong, 2001; Henry et al., 2004).

[6]In other words, people may not actually truly believe that one group is less worthy than another, but rather use claims about worthiness to justify actions that are made for other (less socially acceptable) motivations (*e.g.*, racism). Moreover, several mechanisms are consistent with our data. For example, it could be that racial animosity influences perceptions about the worthiness of Black recipients, which in turn influences support for welfare. It is also possible that racism and misperceptions about worthiness independently influence welfare support.

[7]We present a comparison with experiment 2 because the priming experiment was embedded within this experiment.

[8]It is, however, possible that we would find a different effect of providing accurate information if respondents were told that this information came from a fully representative sample. This may, however, be unlikely as many respondents believed the estimate that they were given. It is also possible that the effects of information provision depend on factors such as who is presenting the information, and the context within which the information is presented.

large body of work on whether beliefs about the ethnic distribution of welfare recipients alter support for welfare in a US context. Gilens (1995), Gilens (1996), Ribar and Wilhelm (1999), Alesina et al. (2001), Luttmer (2001) and Alesina and Glaeser (2004) all use observational data to make the case that believing that a greater proportion of welfare recipients are Black leads to less support for welfare spending. On a more theoretical level, Lee and Roemer (2006), Lee et al. (2006), Alesina and Stantcheva (2020) show how such beliefs ought to determine both welfare support and therefore equilibrium outcomes. Our study builds on this literature by providing novel experimental evidence on this important issue.[9]

Our findings are also reminiscent of those obtained by studies on the impact of immigration on preferences for redistribution. A number of studies find that an influx of immigrants can dampen support for redistribution in a variety of contexts (Alesina and Ferrara, 2005; Dahlberg et al., 2012; Kraus et al., 2019; Dustmann et al., 2019; Tabellini, 2020; Bonomi et al., 2021; Fouka et al., 2022). In addition, Alesina et al. (2018) find that just priming individuals to think about immigration reduces their support for redistribution. Both of these findings are echoed by our results.[10] It is also consistent with research suggesting that biased beliefs can cause discriminatory behavior in a range of different market settings (Sarsons, 2017; Bohren et al., 2019a,b, 2022).

Finally, our study contributes to a growing literature that examines whether information provision can alter behavior and attitudes (see Haaland et al. (2020) for a review).[11] Especially related are the studies that provide race-related information, such as estimates of the Black/White wealth gap (Onyeador et al., 2021; Alesina et al., 2021; Callaghan et al.,

---

[9]There are also a number of papers that although highly related, do not directly study the impact of beliefs about the ethnic/racial distribution of welfare recipients on welfare support. For example, O'Brien (2017) conducts a survey experiment and finds that White American support for taxation is influenced by the share of the population that is Hispanic. Wetts and Willer (2018) conduct two survey experiments and show that White Americans are less likely to support welfare if the programs are framed as primarily benefiting Black Americans, and that they are also less supportive of welfare when presented with information suggesting that Whites' status as a majority group is rapidly coming under threat. Bobo and Kluegel (1993) also find observational evidence suggesting that White Americans are opposed to race-targeted welfare policies. Finally, Gilens (1996), Harell et al. (2016) and Brown-Iannuzzi et al. (2017) conduct hypothetical vignette experiments, which suggest that providing cues about the race of particular individuals influence the amount of support that subjects deem they should receive (see also Fong and Luttmer (2009) and Gross and Wronski (2021)). Our results support the general finding in these papers that individuals have race-related preferences regarding charitable giving and welfare.

[10]Our results also support the priming race psychological research by Tesler (2015a,b).

[11]For recent examples, see Armantier et al. (2016); Delavande and Zafar (2018); Fuster et al. (2018); Andre et al. (2019); Armona et al. (2019); Roth and Wohlfart (2020); Roth et al. (2021).

2021)[12], and the amount of discrimination faced by Black Americans in the labor market (Haaland and Roth, 2021).[13] Moreover, our fourth finding—that providing accurate information fails to change behavior—echoes the null result obtained by Alesina et al. (2018) and Hopkins et al. (2019), albeit in a very different context.

Finally, at the highest level, our study contributes to the wider literature on racial disparities within the US. Studies on this topic have documented a wide array of inequalities in social, legal and economic outcomes between White and Black Americans (Chiteji and Stafford, 1999; Barsky et al., 2002; Charles and Hurst, 2002; Gittleman and Wolff, 2004; Altonji and Doraszelski, 2005; Charles and Guryan, 2008; Boustan, 2010; Rothstein, 2017; Bayer and Charles, 2018; Derenoncourt et al., 2021). Insofar as our findings map into welfare policy (discussed later), they may help go some way to explaining a portion of these disparities.

The remainder of our paper is structured as follows. Section 2 outlines the design of our two experiments, and Section 3 provides our central results. Finally, Section 4 concludes with a discussion of the policy implications of our results.

2 Experimental design

We conducted two experiments: the first ($n = 5,793$) in January 2021 and the second ($n = 3,982$) in October 2021. Participants were recruited via Prolific Academic and were paid a flat fee of \$1.60 in exchange for their participation (they were also entered into various lotteries, as outlined below).[14] The experiments took place within two Qualtrics surveys, both of which took an average of 13 minutes to complete. We also recorded the time taken by respondents on all important questions. While the experiments shared many similarities, there were some key differences that we describe below.

---

[12]The fact that individuals have inaccurate beliefs about race-related issues is also picked up by studies on perceptions of racial equality in the US—see Brodish et al. (2008); Eibach and Keegan (2006); Kraus et al. (2017, 2019); Kuo et al. (2020).

[13]Our result is also related to the work of Bursztyn et al. (2020b) and Bursztyn and Yang (2021) who find that people have incorrect beliefs about certain norms, but such beliefs can be corrected, which then impacts on subsequent relevant behavior.

[14]More information about Prolific Academic can be found at https://www.prolific.co/. Peer et al. (2017) show that participants recruited via Prolific Academic are less dishonest, are less likely to fail attention checks, and produce higher quality data than participants recruited via other comparable online research platforms. Similar findings on the data quality from Prolific are more recently demonstrated by Gupta et al. (2021).

2.1 Experiment 1

The first experiment[15] began by asking respondents some standard demographic questions, including their age, state of residence, educational attainment, gender, race/ethnicity, household income, political affiliation, and whether they had ever been on welfare.[16] In the course of asking respondents whether they had been 'on welfare', we explicitly defined what we meant by the term 'welfare' for the purposes of the survey. More specifically, we told respondents that 'welfare' refers to any of the following programs: (i) Temporary Assistance for Needy Families (TANF); (ii) Medicaid; (iii) Supplemental Nutrition Assistance Program (SNAP); (iv) Housing Assistance; and (v) Supplementary Security Income (SSI).[17]

Before moving to the main treatment, we exposed respondents to a 'strategic attention check' (following Alesina et al. (2018)). To do this, we asked respondents whether we should use their responses, or instead whether they should discard their responses since they had not devoted their full attention to the questions so far. The main aim of this question was to prompt respondents to pay attention in the next (and more important) section of the survey.

We then elicited respondents' beliefs about the share of welfare recipients who identify as White, Black, or as belonging to other races and ethnicities. We also asked what proportion of welfare recipients they thought identify as Spanish/Hispanic/Latino. The ethnic/racial classifications that we used throughout were those used by the US Census Bureau. We incentivized accurate answers to these questions by telling respondents that answers within 2 percentage points of the truth would be rewarded with entry into a lottery for $100. We closed the belief elicitation section by asking respondents about their degree of confidence in their estimates.

Having elicited respondents' beliefs, we then randomly assigned respondents into either the 'high' treatment group, the 'low treatment' group, or a control group. The high treatment group was told:

---

[15]All questions asked in this experiment can be viewed here: https://brown.co1.qualtrics.com/jfe/form/SV_br468nd3DknmvWu

[16]In addition, we also asked participants about their news consumption, how much they believe that the US government spends on welfare, and whether they voted in the 2020 Presidential Election.

[17]As an empirical matter, it is unclear whether the extent to which welfare attitudes depend on race is affected by the precise programs included in the definition. Nonetheless, we thought that an explicit definition could be helpful, not least because it allows us to calculate whether individual perceptions about welfare are accurate. Our definition of welfare includes the five largest means-tested redistributive programs (excluding the Earned Income Tax Credit (EITC), as this program has a work requirement).

Estimates* from 2017 suggest that, out of every 100 American adults on welfare,

- 63 were White
- 26 were Black
- 11 belonged to other ethnic groups

*Please note that these estimates were obtained using the Survey of Income and Program Participation (2018). The statistics were computed for individuals who were born in the month of October (in any year) and may thus not be fully representative of the overall population. See https://www.census.gov/programs-surveys/sipp.html for more information.

Those in the low treatment group were also given estimates about the racial/ethnic distribution of welfare recipients. However, those estimates were obtained from those born in September (in any year), and yielded the following distribution: 68% White, 20% Black, and 12% belonging to other groups.[18] Finally, the control group received no such information. We stress that, although we provided different information to members of the different treatment groups, all information provided was fully accurate and the nature of the estimates was disclosed (*i.e.*, we did not employ any deception).

In an effort to ensure that participants processed the information provided, we next asked them to recall the estimates that they had just been shown. Respondents were then told if their answer had been correct, and were then shown the treatment estimates for a second time. While the main motivation of this quiz was to encourage respondents to further internalize the estimates, the answers to the quiz may also have provided further information about the attentiveness of particular subjects (as discussed below).[19]

Immediately after randomly exposing participants to the treatments, we once again elicited (treated) respondents' beliefs about the ethnic/racial distribution of welfare recipients. To do this, we told respondents that, while the estimates which they had been presented were from 2017, they should instead "think about now". They were then asked, out of every hundred American adults on welfare, how many are Black, White, or neither. As before, we incentivized correct answers (within 2 percentage points of the truth) and asked respondents how confident they were in their answers (on a 5-point Likert scale).

After re-eliciting beliefs, we recorded our main outcome of interest, namely whether re-

---

[18]The corresponding White/Black shares for those born in January, those born in February, and so on are: 68/21, 65/21, 66/23, 64/24, 67/23, 68/21, 65/25, 64/25, 68/20, 63/26, 64/24, and 66/22.

[19]Note that participants in the control group were not given this quiz as they did not receive any treatment information. For this same reason, we also did not re-elicit their beliefs about the racial/ethnic distribution of welfare recipients.

spondents supported welfare. To measure this in an incentive-compatible way, we first told respondents that they would automatically be enrolled into a lottery for $100. We then asked them whether they would like to donate their potential winnings to either of two non-profit organizations, one chosen to be 'pro-welfare' and the other chosen to be 'anti-welfare'. Respondents were given the following information about the non-profits:

1. **The Center on Budget and Policy Priorities** is a progressive American organization (think tank) that works to ensure that policymakers consider the needs of low-income people. Many of these people receive food stamps, housing assistance or other forms of welfare assistance.

2. **The Foundation for Government Accountability** is an organization (think tank) that focuses on welfare and health care reform. Many of the policies proposed by this think tank would have the impact of reducing federal welfare spending.

We used an incentive-compatible outcome (donations) in order to minimize any possible experimenter demand effects and ensure that subjects were attentive when providing answers. This approach has also been used in other experimental studies on political attitudes. For example, Bursztyn et al. (2020a) use donations to a xenophobic organization as their main outcome of interest when studying how social norms influence racist expression. Somewhat closer to our setting, Alesina et al. (2018) ask respondents if they would like to donate prospective lottery winnings to charities supporting low-income adults or children in an attempt to capture attitudes toward redistribution.

Having measured respondents' attitudes toward welfare, we then investigated the mechanisms through which beliefs about the racial composition of welfare recipients might influence support for redistribution. One explanation for any observed effects would simply be that respondents weight the welfare of White and Black welfare recipients differently, for example because of racial animosity or in-group preferences (Alesina and Stantcheva, 2020). Other possibilities are that participants believe that the efficacy of welfare spending is different for Black and White recipients, or that Black and White recipients are not equally worthy of welfare (see, *e.g.*, Cook and Barrett (1992), Fong (2001), Henry et al. (2004)). Given the difficulty of measuring racial animosity and in-group preferences (in light of experimenter demand effects, and also because the act of measuring such factors is likely to prime participants), we did not study these factors. We did, however, explore the potential importance of beliefs about efficacy and worthiness in explaining any observed effects.

There are, in principle, two ways of doing this. First, one could employ our general

experimental design while, for example, also holding constant the perceived worthiness of Black and White recipients (e.g., by informing recipients that Black and White welfare respondents are equally likely to be poor through no fault of their own). The goal would be to check if any observed effects decrease once mediating factors have been held fixed. However, one problem with this design is that it is very challenging to adequately change participants' beliefs about worthiness and efficacy, partly because such matters cannot be easily settled using objective data. Another problem is that providing such information might inadvertently reveal the purpose of the study, and consequently alter the treatment effects. Finally, it would be difficult to interpret the results of this experiment because information about worthiness or efficacy might alter the observed treatment effects by simply changing the baseline level of welfare support.

The alternate design, which we pursue, is to ask respondents about their views on the efficacy of welfare and worthiness of welfare recipients, and to measure whether participants provide different responses in the two treatment groups. To do this, we asked respondents whether they thought that "welfare programs help lift Americans out of poverty" to capture an "efficacy" channel; and whether they thought that "people who receive welfare are poor through no fault of their own" to capture a "worthiness" channel. With this design, we avoid the issues raised above, and we are able to determine whether shifting beliefs about the racial composition of welfare recipients influences perceptions about the efficacy of welfare and worthiness of welfare recipients. We are, however, unable to determine the causal impact of these factors on welfare support in a fully rigorous way.

We concluded the experiment by asking respondents about their likelihood of winning the lottery to ensure the validity of our incentivization. We also measured their 'implied beliefs' about the share of welfare recipients that are from different racial/ethnic groups. To do this, we asked what share of different groups they believed were on welfare and what share of the population belong to different groups (allowing for a calculation of 'implied beliefs' using Bayes rule).[20] Finally, we asked an open-ended question on what they thought the study was about. This question allowed us to drop those who understood that the survey was about racial attitudes (as a robustness check) as we wanted to avoid experimenter demand effects.

---

[20]For example, if a participant believes that the 10% of the population are Black, that 30% of the Black population are on welfare, and that 20% of the population are on welfare, then their 'implied belief' for the share of welfare recipients who are Black is $0.3 \times 0.1/0.2 = 0.15$.

2.2 Experiment 2

The structure of the second experiment[21] broadly mirrored that of the first (see Figure 1 for an outline of the experimental structure). However, there were several important differences. First, we dramatically increased the difference in the signals that the two treatment groups were given in order to ensure that our results are not sensitive to the exact parameter values used in the first experiment. Rather than providing information on welfare recipients born in different months (experiment 1), we provided information on welfare recipients born in different months of a particular year (December 1986 for the 'low' group and April 1987 for the 'high' group). This generated an 8% Black, 75% White, and 17% 'other' distribution in the 'low' treatment group; and a 52% Black, 31% White, and 17% 'other' distribution in the 'high' treatment group.[22] As in the first experiment, we informed participants how the estimates had been generated, and that they might be unrepresentative of the population as a whole.

---

[21] All questions asked in this experiment can be viewed here: https://brown.co1.qualtrics.com/jfe/form/SV_2363JGhin9dKoQK

[22] The White/Black shares for those born in December 1986, January 1987, February 1987, and so on are: 75/8, 83/17, 82/18, 68/27, 31/52, 72/23, 65/24, 45/43, 53/43, 51/26, 51/49, and 73/21. This illustrates the substantial degree of variation in the SIPP data.

Figure 1: Experimental design (experiment 2)
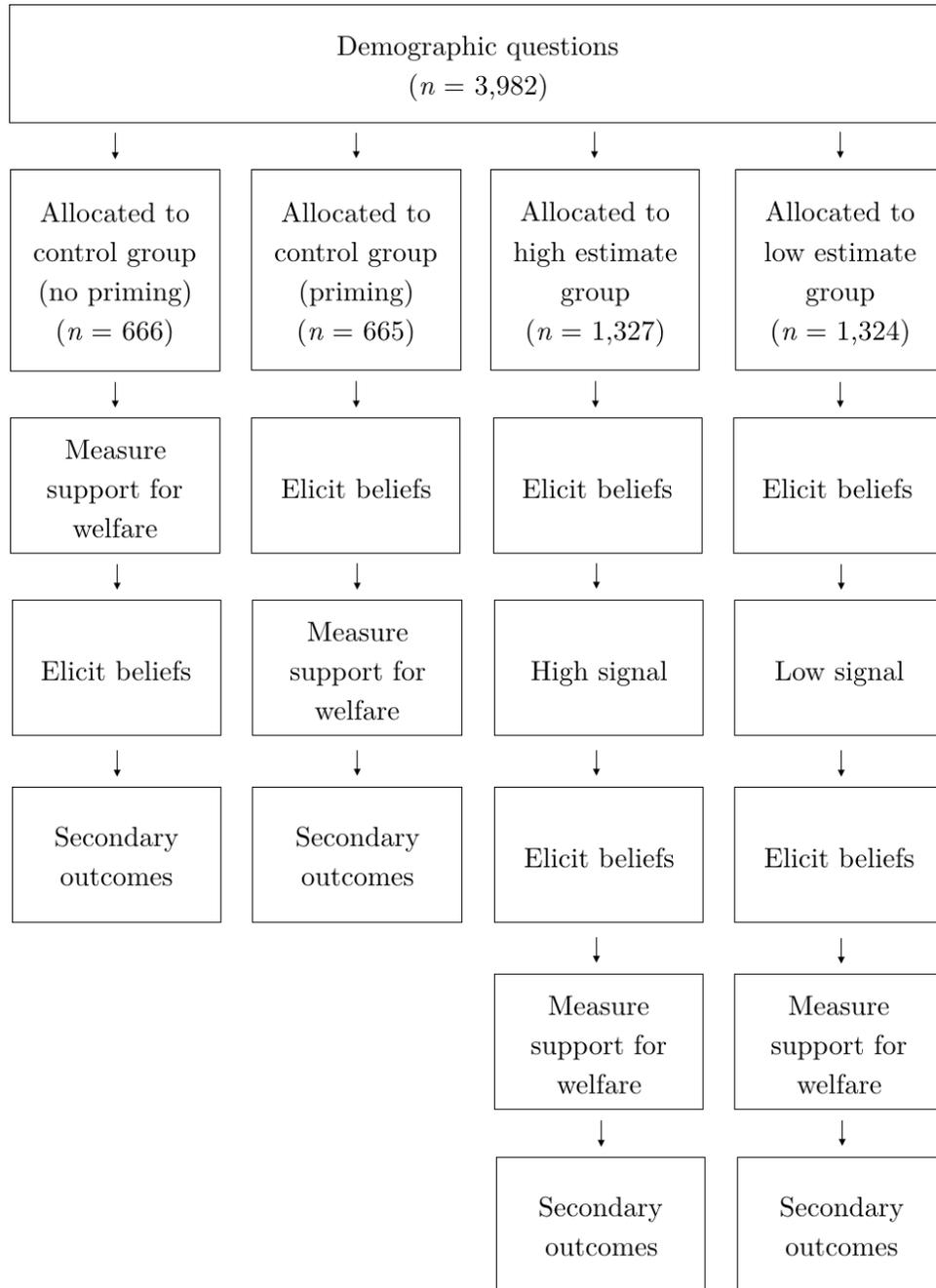
Notes. In this figure, we present the experimental procedure for experiment 2. The structure of experiment 1 is very similar, with the main difference being that we do not randomize those in the control condition into two groups.

Second, we recorded the extent to which participants displayed 'social desirability bias' (*i.e.*, a tendency to say and do things in order to conform with the relevant group) prior to presenting the treatment information. To do this, we asked participants questions from the Marlowe-Crowne Social Desirability Scale (Crowne and Marlowe, 1960). We elicited this information to allow us to conduct a robustness check where we drop those who exhibit high degrees of social desirability bias as these individuals may be less likely to reveal their true beliefs and attitudes—particularly if they realized what the survey was about. This approach has been used in a number of other experimental studies, such as Dhar et al. (2018).

Third, similar to Alesina et al. (2018), we embedded a priming experiment into our survey. More specifically, those who were randomly assigned to the control group were further randomized into one of two groups: 1) those asked whether they would like to donate money to the non-profits before being asked about the share of welfare recipients that belong to different ethnic/racial groups, and 2) those asked if they would like to donate money after being asked about the share of welfare recipients that belong to different ethnic/racial groups. In other words, we randomized whether the question about ethnicity came before or after the questions about welfare support. This allowed us to investigate whether being primed to think about race might influence support for welfare.

Finally, we included two additional questions at the end of the survey that provide us with descriptive evidence about the relationship between people's beliefs about the share of welfare recipients that are Black and their support for welfare. The first question asked whether respondents took race into account when deciding whether to donate money to one of the two non-profits. The second question asked whether participants had thought about the share of welfare recipients that belong to different ethnic/racial groups prior to taking the survey. We asked these questions to further understand whether these beliefs about race matter for expressed welfare support both within the survey and beyond it.

2.3 Data and sample

In total, we sampled 9,775 participants, 5,793 of whom took part in the first experiment and 3,982 of whom took part in the second experiment. Most participants were White, but we also had 1,846 Black participants take part in the first experiment.[23] While we made some effort to sample representatively, the sample was not altogether balanced on common demographic characteristics. Notably, the sample was disproportionately

---

[23]We primarily recruited White participants since our main hypothesis (*i.e.*, that White individuals are less inclined to support welfare if a greater proportion of recipients are Black) concerns this group.

female (59% female and 66% female in experiments 1 and 2 respectively) and dispro-
portionately liberal (58%) in experiment 1 (but not in experiment 2)—see Tables A1 and
A2 for descriptive statistics.[24] As a result, we checked that our results were robust to
re-weighting the sample to match US population demographics (see discussion below).

3 Results

3.1 Pre-treatment beliefs about race and welfare

We begin by examining the accuracy of respondent beliefs about the racial distribution of
welfare recipients. Figure 2 plots respondent beliefs about the share of welfare recipients
who are Black, pooled across the two experiments. On average, respondents estimate
that 37% of welfare recipients are Black (and that 39% are White, 25% of welfare recip-
ients are neither, and 27% of welfare recipients are Hispanic). These average estimates,
however, are quite far from the truth: for example, data from the Survey of Income and
Program Participation suggest that the share of welfare recipients who are Black is around
21%.[25] As these numbers might indicate, overestimating the share of welfare recipients
who are Black is common in the sample, with 86% of respondents overestimating this fig-
ure. This finding is consistent with prior survey evidence on this issue (see, *e.g.*, Delaney
and Edwards-Levy (2018)).

Next, we examine which groups are most likely to overestimate the share of welfare re-
cipients who are Black. To do this, we calculate the fraction of respondents in various
subgroups who estimate the share to be at least 26 percentage points (*i.e.*, an overesti-
mate of at least 5 percentage points).[26] Table A4 displays the results (see also Table A5
for a linear probability model that reports similar findings). As can be seen, conservatives
and moderates are significantly more likely to overestimate the share of welfare recipi-
ents who are Black than liberals. Similarly, women are more likely to overestimate the
figure than men, and White respondents are more likely to overestimate the share than
Black respondents. Strikingly, however, each of these subgroups still overestimates the
share substantially, including Black respondents who on average estimate the figure at
35%.

---

[24]See also Table ?? for a broader array of descriptive statistics.

[25]See https://www.census.gov/programs-surveys/sipp.html for more information about the SIPP
methodology.

[26]As one might expect, we obtain similar results using alternative cut-offs.

Figure 2: Pre-treatment beliefs
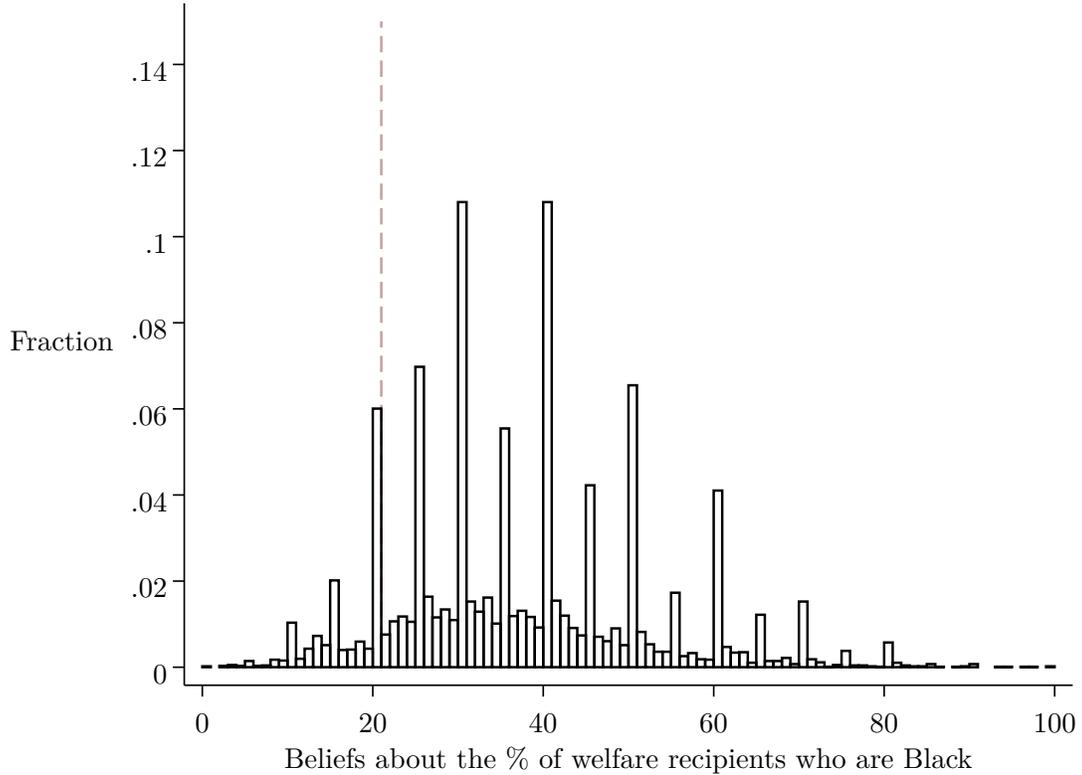
Notes. In this figure, we present the distribution of participants' pre-treatment beliefs (pooled across both experiments) about the share of welfare recipients who are Black. The *y*-axis shows the share of participants who held a particular belief, and the *x*-axis shows the particular point estimate. The dashed line at 21% shows the true share of welfare recipients who are Black. We do not exclude any participants when constructing this figure.

In principle, there are a number of reasons why we might observe these systematic over-estimates. While we are unable to examine all factors that contribute to these beliefs, three potentially important factors—suggested by Bayes rule—are beliefs about the share of Black people who are welfare recipients, beliefs about the share of the US population that is Black, and beliefs about the share of the US population that is on welfare. While we find that respondents do not exactly use Bayes rule, a regression of our outcome on these three variables suggests that they do influence our outcome in the expected direction: for example, increasing beliefs about the share of the US population that is Black by one percentage point is associated with a 0.242 percentage point increase in the perceived share of welfare recipients who are Black (see Table A6).[27] Moreover, we find that respondents

---

[27]The fact that the effect is less than one percentage point suggests that participants are insufficiently sensitive to this component of Bayes Rule. See Augenblick and Rabin (2021) and associated references.

greatly overestimate this statistic, putting the share of the US population that is Black at 27%.[28] This seems likely to be a contributor to the widespread overestimation, although the causal relationship could of course also go in the opposite direction.

3.2 Updating beliefs about the racial composition of welfare recipients

We now examine how our treatments influence beliefs about the share of welfare recipients who are Black, only focusing on White respondents as our main hypothesis (*i.e.*, that White Americans are less supportive of redistribution to Black Americans) concerns this subgroup.[29] Figure 3 displays the distribution of beliefs in the high and low treatment groups of both experiments (the upper panel for experiment 1 and the lower panel for experiment 2). As can be seen, both experiments succeeded in shifting the distribution of beliefs in the expected way. In experiment 1, the distributions are centered around 20% and 26% (*i.e.*, the exact estimates with which subjects were presented in their respective treatments). In experiment 2, the distributions are centered around 8% and 52%, which are again the estimates from the two treatments. While the differential assignment to the treatments shifted beliefs in the expected way, there is some clustering around the provided estimates (which does not pose any problems for our analyses). In part, the clustering may be because subjects were presented with information from a sub-sample from 2017, and therefore chose to make some revisions when estimating the figure for the population in 2020.

---

[28]Similar overestimates have been observed in previous studies, e.g., Nadeau et al. (1993). Relatedly, there is also evidence that people overestimate how many immigrants (both legal and illegal) are in the US (Abramitzky and Boustan, 2017; Alesina et al., 2018).

[29]We only focus on White respondents in the remainder of the empirical section unless explicitly stated otherwise.

Figure 3: Belief updating in experiments 1 and 2
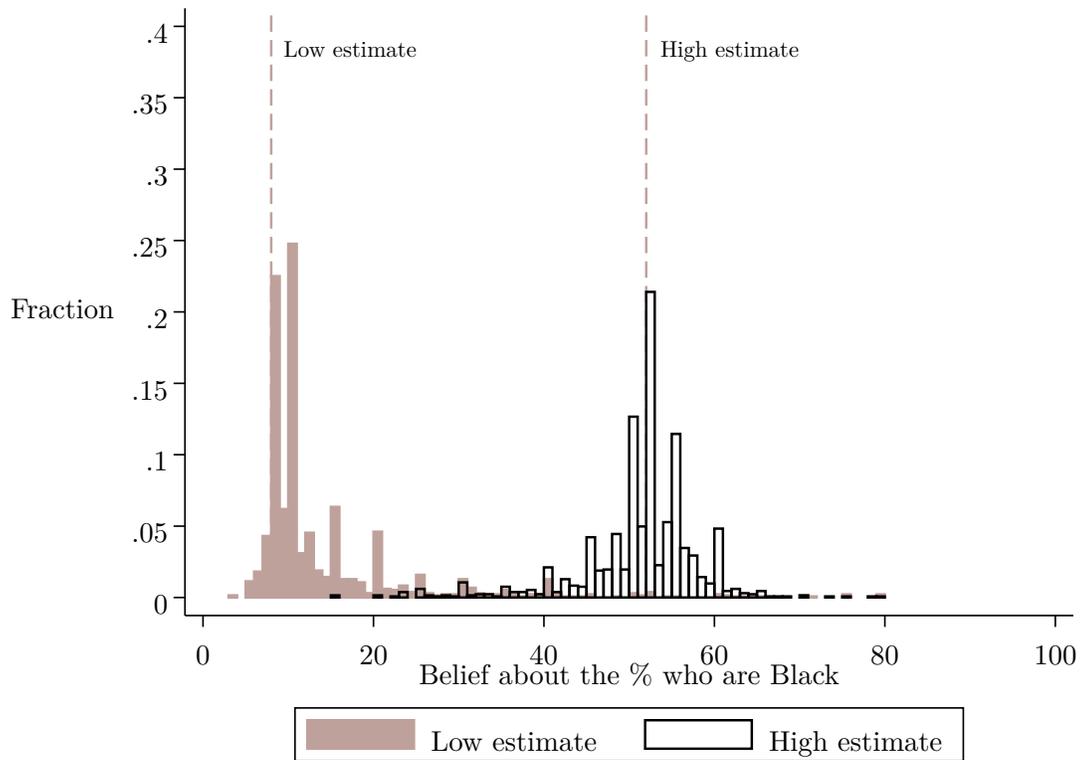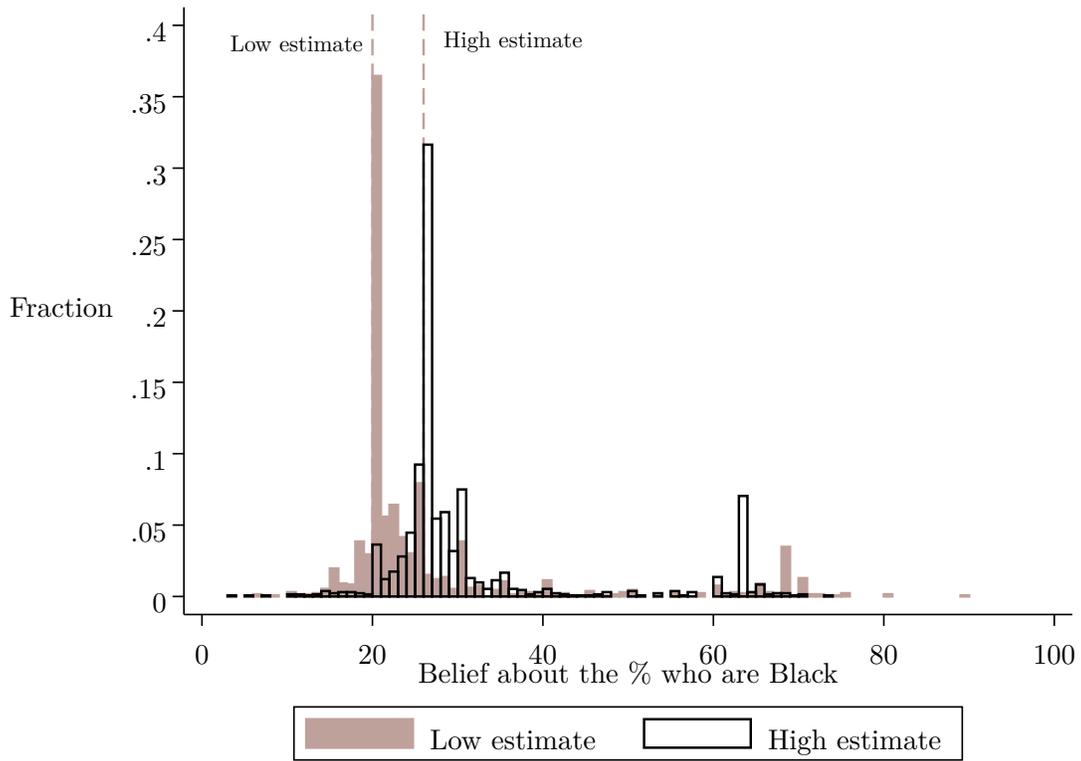
Notes. The top panel presents the distribution of beliefs about the share of welfare recipients who are Black in the high- and low-signal groups, respectively, in the first experiment. The bottom panel does the same for the second experiment. We restrict our attention to White respondents when constructing these figures, and we do not include participants in the control groups.

As discussed earlier, we manipulate beliefs in this way so that we can use treatment assignment as an instrument when estimating the effect of beliefs on attitudes towards welfare. To confirm that our instruments are informative, we now regress beliefs about the share of welfare recipients who are Black on treatment assignment in both experiments (omitting the control groups from the analysis, as these are not used when conducting the instrumental variable analyses). As can be seen from Table A7, allocating individuals to the high as opposed to low treatment group increases beliefs by 4.16 percentage points on average in experiment 1 ($p < 0.001$). In experiment 2, allocating individuals to the high as opposed to low treatment group increases their beliefs on average by 36.9 percentage points ($p < 0.001$). Since the F-statistics are large in both cases (65 and 11253 in experiments 1 and 2 respectively), we can confidently conclude that our instruments are informative.[30]

Next, we examine whether our instruments are exogenous. In other words, we investigate whether our instrument (treatment assignment) can influence our outcome (donations) through a variable other than beliefs (by which we specifically mean 'point estimates'). We can only think of two ways in which our instruments might have influenced relevant factors other than beliefs: 1) they might influence the salience of beliefs; and 2) they might influence the confidence with which beliefs are held. Since both treatment groups are provided with similarly formatted information, we can reject the first possibility: both groups should be equally primed to consider the information as salient. To investigate the second possibility, we estimate the effect of treatment assignment on stated confidence. As Table A8 shows, we do not find a significant difference in the confidence with which beliefs are held between those allocated to the high and low groups in both experiments (*i.e.*, the coefficients in the first two columns are not significant at a 5%-level).[31] We conclude that differences in confidence across groups are unlikely to be an issue for biasing our key coefficient estimates and thus our instruments are exogenous.[32]

---

[30]The F-statistic is most likely larger in the second experiment since that experiment presents the individuals in the treatment groups with more contrasting estimates.

[31]As a robustness check, we control for confidence when conducting the instrumental variable analyses and we do not find that this changes our estimates appreciably. We also examine the effects of the treatment using different ways of coding the confidence outcome (*e.g.*, treating it as a binary or continuous variable). See Section 3.3 for more information.

[32]Note that these points would *not* apply to a comparison of either of our treatment groups with the control group. First, members of the treatment groups turn out to be substantially more confident in their estimates than members of the control group, which is exactly what one would expect given that they (unlike the control group) have been given information. Second, and for this reason, members of the treatment groups may have been more primed to think about race than members of the control group. For this reason, we omit the control group from our main analysis.

3.3 Beliefs about the racial composition of welfare recipients and support for welfare

3.3.1 Main analysis

Having argued that our instruments are informative and exogenous, we now turn to the instrumental variables analysis. Specifically, we examine if beliefs about the share of welfare recipients who are Black influence White respondents' support for welfare, using treatment assignment as an instrument for beliefs (and excluding the control group).

For each experiment and outcome, we estimate the models

$$P_i = \theta_{0P} + \theta_{1P} b_i + u_i \tag{1}$$

and

$$A_i = \theta_{0A} + \theta_{1A} b_i + v_i \tag{2}$$

where P_i is a binary variable denoting whether individual i donates to the pro-welfare nonprofit, and A_i is a binary variable denoting whether individual i donates to the anti-welfare nonprofit. Furthermore, b_i is individual i's belief about the share of welfare recipients who are Black; u_i and v_i are error terms; and θ_{0P} and θ_{0A} represent the constant terms. Moreover, θ_{1P} is a regression coefficient which captures the local average treatment effect of beliefs on participants' propensity to donate to the pro-welfare organization, and θ_{1A} does the same for their propensity to donate to the anti-welfare organization. As discussed, we instrument for the explanatory variable in the two regressions using treatment assignment. The 'first stage' equations both take the form

$$b_i = \delta_0 + \delta_k T_i + e_i \tag{3}$$

where the T_i is a dummy indicating treatment assignment (*i.e.*, whether respondents were placed in the high or low signal group); δ_0 is a constant; and e_i is the error term. That is, we instrument for beliefs using a dummy indicating assignment to the high or low signal group. We drop the control group from these analyses to ensure that the exclusion restriction is satisfied. The regressions are conducted using two-stage least squares (2SLS).

In addition, we also conduct Intention to Treat (ITT) estimates, which gives us the direct effects of treatment assignment on participants' willingness to donate to the two organizations. As before, we drop the control group when conducting these analyses. The ITT

regressions take the form

$$P_i = \beta_{0P} + \beta_{1P} T_i + w_i \tag{4}$$

and

$$A_i = \beta_{0A} + \beta_{1A} T_i + z_i \tag{5}$$

where we use the same variable definitions as before. More specifically, β_{0P} and β_{0A} are constant terms, and β_{1P} and β_{1A} represent the effect of treatment assignment on participants' propensity to donate to the two organizations. Finally w_i and z_i are error terms. Table 1 displays results for the ITT estimates and the LATE estimate obtained from 2SLS.

Our main objective with this empirical exercise is to test whether one can reject the hypothesis that White support for welfare is unaffected by beliefs about the racial composition of welfare recipients. Within each of the experiments, we are able to reject this hypothesis if either θ_{1P} or θ_{1A} are significantly different from zero.[33]

In experiment 1, we find that beliefs have a significant causal effect on the share who donate to the anti-welfare organization (see Column 3). More specifically, for every one percentage point increase in beliefs about the share that is Black, participants become 0.7 percentage points more likely to donate to the anti-welfare organization ($p = 0.02$) (column 4). We also find evidence suggesting that higher beliefs about the share that is Black make participants less likely to donate to the pro-welfare organization, although the estimated coefficient is not statistically different from zero (Column 2).

In experiment 2, the findings are similar (see Columns 6 and 8). We find a significant effect of beliefs on the share who donate to the pro-welfare organization: for every percentage point increase about the share that is Black, participants become 0.1 percentage points less likely to donate to the pro-welfare organization (column 6). Moreover, while we do not find a significant relationship between beliefs and the share who donate to the anti-welfare organization, we do find that the estimated coefficient is positive (column 8). In summary, then, all four coefficient estimates (from both experiments) suggest that higher beliefs about the share that is Black reduce support for welfare. Note, however, that we do not obtain the exact same quantitative estimates in both experiments, which

---

[33]One could argue that we are unable to reject this hypothesis if, for example, we find that both coefficients differ from zero and that both coefficients have opposing signs. In other words, one might then be able to say that beliefs about the racial composition of welfare recipients influences behavior, but it is unclear whether it increases or decreases support, as there are offsetting effects. However, in our setting, we do not obtain such results, meaning that we do not have to deal with this type of ambiguity.

Table 1: Main results

Experiment 1	(1) ITT Pro-welfare donation	(2) LATE Pro-welfare donation	(3) ITT Anti-welfare donation	(4) LATE Anti-welfare donation
High signal	-0.0262 (0.0185)		0.0298** (0.0127)	
Beliefs about % Black		-0.0063 (0.0045)		0.0072** (0.0032)
Constant	0.3580*** (0.0132)	0.5260*** (0.1300)	0.1060*** (0.0085)	-0.0852 (0.0904)
Observations	2,646	2,646	2,646	2,646

Experiment 2	(5) ITT Pro-welfare donation	(6) LATE Pro-welfare donation	(7) ITT Anti-welfare donation	(8) LATE Anti-welfare donation
High signal	-0.0361** (0.0173)		0.0006 (0.0109)	
Beliefs about % Black		-0.001** (0.0005)		0.0000 (0.0003)
Constant	0.2920*** (0.0125)	0.3050*** (0.0178)	0.0861*** (0.0077)	0.0859*** (0.0110)
Observations	2,651	2,651	2,651	2,651

Notes. In this table, we present ITT and LATE estimates. The ITT estimates reveal the effect of assignment to the high signal group (relative to the low signal group) on participants' propensity to donate to one of the two charities. The LATE estimates reveal the effect of beliefs about the share of welfare recipients who are Black on their propensity to donate to the charities (using treatment assignment as an instrument). We exclude the control group and Black respondents from the analysis. Standard errors are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

may be due to factors such as sample composition and the precise signals used to exogenously change beliefs.[34]

---

[34]For example, the signals provided in experiment 1 were both below most participants' pre-existing beliefs, while the high estimate presented in experiment 2 was far above most people's pre-existing beliefs. Moreover, previous studies have found that updating beliefs upward or downward may have slightly different effects on behavior (see, e.g., Eil and Rao (2011)). Furthermore, one might have expected that our results would be stronger for conservative than for liberal participants. Surprisingly, however, we do not find evidence for this—see Table A9.

We can also examine the relationship between beliefs and support after combining the data from both experiments, as this affords us greater statistical power. To do this, we now instrument for beliefs using a dummy variable indicating treatment assignment (high or low signal), a dummy variable indicating which experiment participants took part in (experiment 1 or 2), and the interaction of these two variables. As before, all instruments are clearly informative. Moreover, in light of the previous discussion, the treatment assignment instrument is plausibly exogenous. However, since participants were not assigned to experiments randomly, there is a possibility that this instrument influences our outcome of interest because different types of people took part in the two experiments, rather than just affecting it through beliefs. We therefore control for all observables when conducting our pooled analyses.

As recorded in Table A10, the pooled results closely resemble those of experiment 2. As before, we observe that increasing beliefs about the share of welfare recipients who are Black significantly decreases the share who donate to the pro-welfare organization ($p < 0.01$). The estimate is that a one percentage point increase in beliefs decreases the share who donate by 0.12 percentage points, an effect that is quantitatively very similar to our previous estimate. As in experiment 2, we also do not obtain a statistically significant effect on donations to the anti-welfare organization.

Next, we examine if there is a relationship between beliefs about the ethnic/racial distribution of welfare recipients and support for welfare among Black respondents. To do this, we restrict our attention to experiment 1, which had close to 2,000 Black participants (see Table A11). While we find that the treatments update participants' beliefs in the expected way (*i.e.*, those in the high signal group have higher beliefs than those in the low signal group on average), we do not find that there is a significant relationship between beliefs and welfare support. This may be due to there not being a relationship between these variables for Black Americans, but it may also be due to a lack of statistical power.

To better understand how economically meaningful the relationship between the racial composition of welfare recipients and welfare support is, we can compare these estimates with the associations between welfare support and other variables known to predict such support. Strikingly, beliefs about the ethnic distribution of welfare recipients appear to be a rather important factor compared to these other determinants (see Table A12). For example, according to the estimates obtained from experiment 1, a 10 percentage point shift in beliefs has as large an effect on the share that donate to the anti-welfare nonprofit

as the difference in donation rates between conservative and liberal respondents.[35] These comparisons suggest that the effects that we uncover are not just statistically significant, but also quantitatively meaningful—and could therefore be an important determinant of welfare spending in the United States.

3.3.2 Internal and external validity

The fact that we obtain qualitatively similar effects in two separate experiments suggests that the phenomenon that we uncover is robust. To further study this issue, however, we now conduct several sensitivity analyses. The first set of checks we conduct involves dropping apparently less attentive participants. We begin by dropping participants who completed the survey in less than t minutes, for various values of t, on the grounds that such participants may have failed to pay proper attention to the questions. We also drop participants who spent less than t seconds on the treatment screens, and who spent less than t seconds when providing us with their belief estimates (again doing this for various values of t). Finally, we drop participants who failed to answer the post-treatment quiz correctly (*i.e.*, the question where we asked participants to recite the treatment information), and the very small number of participants who stated that the share of welfare recipients who are Black is 0% or 100%. In general, our estimated coefficients remain quantitatively similar and statistically significant, despite the reduced sample size (see Table A13 and A14 for the results of all robustness checks).

The second set of checks we conduct involves excluding participants who may have failed to provide fully truthful and reflective answers due to experimenter demand or a lack of incentivization. We do this by dropping participants who thought that it was unlikely that they would win the lotteries, on various definitions of 'unlikely'. We also exclude participants who displayed a high degree of 'social desirability bias' (as measured by the Marlow-Crowne scale), and those who understood what the study was about, as judged by their answers to the open-ended debriefing question at the end of the survey. Again, our estimated coefficients remain quantitatively similar, but also become more statistically significant.

We also conduct some additional robustness checks. First, we re-weight our data so that it

---

[35]In experiment 1, we find that a one percentage point increase in beliefs about the share that is Black leads to a 0.72 percentage point increase in the share that donates to the anti-welfare nonprofit. Thus, a 10 percentage point shift in beliefs equates to a 7.2 percentage point shift in donations, and the difference in donation rates between conservative and liberal respondents is 7.81 percentage points (holding gender, age, welfare receipt, education, voting behavior, and beliefs about the share of welfare recipients that is Black constant).

matches the joint distribution of gender, age, and income for the US population. Second, we conduct the analyses while controlling (and instrumenting) for the confidence with which beliefs are held. Third, we re-estimate the regressions using participants' implied beliefs about the share of welfare recipients that is Black (we obtain this using their estimates of the share of the US population that is Black, the share of Black people who are on welfare, and the share of the US population that is on welfare). Finally, we re-estimate our regressions including the control group in the analysis. As in the previous cases, none of these exercises alter our substantive conclusions.

We can also examine whether the effects we uncover can be detected in the descriptive data. To investigate this, we regress donations on participants' beliefs about the share of welfare recipients that is Black, attempting to control for all relevant confounds in the hope that this might give our regression results a causal interpretation.

We also utilize different measures of beliefs (*i.e.*, those elicited prior to treatment assignment and those elicited afterwards) when conducting these regressions. Strikingly, we find that the simple associations between beliefs and support for welfare are close to identical to the estimates that we obtain in our instrumental variable analyses (see Table A15).

To analyze the generalizability of our results, we examine whether our measure of welfare support (whether individuals donate to the pro- or anti-welfare charity) is predicted by the same variables as other measures of welfare support in the literature. As Table A12 shows, we find that women are less likely to donate to the anti-welfare welfare than men (all else being equal). We also find that liberals are more inclined to donate to the pro-welfare charity than conservatives and moderates (and less likely to donate to the anti-welfare charity). Similarly, our measure suggests that the young are more supportive than the old. Reassuringly, all of these findings are consistent with prior work on this topic (see, *e.g.*, Ashok et al. (2015)), which suggests that our incentivized measure is indeed capturing individuals' attitudes towards welfare.

In addition, we examine whether respondents took race into account when deciding whether to donate to the pro- and anti-welfare organisations. We find that around 42% of the White respondents who donated to the anti-welfare organization said that they took race into account when making this decision, which further suggests that race is an important determinant of welfare support.[36] Moreover, 48.5% of respondents (and 58.3% of those who donated against welfare) said that they have thought about the race/ethnicity

---

[36]This data is only from experiment 2, as we did not ask this question in experiment 1.

of welfare recipients before, which suggests that the relevance of this factor extends beyond our survey.

Aside from testing the internal and external validity of our results, it is also important to understand how they should be interpreted. The most natural interpretation is that our treatments have an effect simply by increasing the share of welfare recipients deemed to be Black, and decreasing the share deemed to be White. However, other alternatives are possible: it might be that increasing beliefs about the share that is Black also alters beliefs about the share on welfare who are neither White nor Black (*i.e.*, those who belong to 'other' racial groups). Similarly, our treatment may also shift beliefs about the share of welfare recipients who are Hispanic/Latino. To examine this issue, we measure the effects of treatment assignment on participants' beliefs about the share of welfare recipients who are White, belong to 'other' racial groups, or are Hispanic/Latino.

When conducting these six regressions (three per experiment), we use beliefs about the proportion of welfare recipients who are 1) White, 2) Other, 3) Hispanic as our outcome variables. As usual, we exclude the control group from these analyses, and only include participants who identify as White. Our only independent variable is a dummy indicating treatment assignment to the high or low condition.

As can be seen in Table 2, the treatments do not influence beliefs about the share of welfare recipients who belong to 'other ethnic/racial groups', and similarly do not substantially influence the share of welfare recipients who are Hispanic/Latino. This suggests that the most obvious interpretation is the right one: the treatments alter beliefs about the share of welfare recipients who are Black while simultaneously altering beliefs about the share that is non-Hispanic White.

3.3.3 Mechanisms

We now discuss the mechanisms underlying our main result. In some sense, the mechanism is transparent: if White Americans support welfare less when they believe that a higher share of welfare recipients are Black, this must be because they place less value on welfare spending that goes to Black Americans. Nonetheless, one can ask the further question of what explains this underlying preference. In practice, it is very difficult to adjudicate between different explanations (*e.g.*, taste-based vs statistical theories), primarily because individuals might pretend to object to giving welfare on 'statistical' grounds (*e.g.*, because welfare is less effective for some groups than others) even when their objections are rooted in racial animus (Bursztyn et al., 2020c). However, we now attempt to shed

Table 2: Treatment effect on other beliefs

	Experiment 1			Experiment 2		
	Belief % White	Belief % Other	Belief % Hispanic	Belief % White	Belief % Other	Belief % Hispanic
High signal	-3.517***	-0.651***	-0.231	-37.070***	0.224	3.591***
	(0.559)	(0.246)	(0.659)	(0.395)	(0.243)	(0.476)
Constant	59.940***	13.300***	20.310***	68.470***	17.560***	19.320***
	(0.419)	(0.165)	(0.479)	(0.357)	(0.180)	(0.320)
Observations	2,646	2,646	2,646	2,651	2,651	2,651
R^2	0.015	0.003	0.000	0.769	0.000	0.021

Notes. In this table, we present the effect of being allocated to the high signal group (relative to the low signal group) on participants' beliefs about the share of welfare recipients that is White, belongs to Other ethnic/racial groups, or is Hispanic/Latino. We exclude the control group in the analyses. Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

some light on this issue.

We do this by examining the extent to which White people's beliefs about the share of welfare recipients who are Black also influence the extent to which they agree with the following statements: "people who receive welfare are poor through no fault of their own" and "welfare programs help lift Americans out of poverty". We asked respondents about these statements because past studies suggest that White respondents view White and Black welfare recipients differently; and also because beliefs about the efficacy of welfare and worthiness of welfare recipients are key predictors of welfare support (see, *e.g.*, Fong and Luttmer (2009)).

Table 3 presents the results. The analyses are conducted in a similar fashion as before, with treatment assignment acting as an instrument for beliefs. As we can see, we do not find a significant relationship between beliefs and agreement with these statements in either experiment, although our estimated LATE coefficients are negative (*i.e.*, participants are less likely to, for example, think that welfare is effective if the share of welfare recipients who are Black increases). We do, however, find a statistically significant relationship between the share of welfare recipients who are Black and whether participants think that welfare recipients are poor through no fault of their own when pooling the data from the two experiments. This could be taken to suggest that a 'worthiness' channel underlies our main result, particularly as other studies have found a causal effect of

perceived worthiness on charitable donations (*e.g.*, Fong and Luttmer (2011)).[37]

Table 3: Secondary outcomes (LATE estimates)

	Worthiness of welfare recipients		
	Experiment 1	Experiment 2	Pooled analysis
Belief about % Black	-0.00161	-0.0008	-0.0009**
	(0.0047)	(0.0005)	(0.0005)
Constant	0.5740***	0.4500***	0.5540***
	(0.1350)	(0.0195)	(0.0298)
Controls	No	No	Yes
Observations	2,646	2,651	5,297

	Efficacy of welfare spending		
	Experiment 1	Experiment 2	Pooled analysis
Belief about % Black	0.0014	0.0000	-0.0001
	(0.0045)	(0.0005)	(0.0005)
Constant	0.5880***	0.5600***	0.7920***
	(0.1310)	(0.0195)	(0.0300)
Controls	No	No	Yes
Observations	2,646	2,651	5,297

Notes. In this table, we present LATE estimates of the effect of beliefs about the share of welfare recipients who are Black on beliefs about the worthiness of welfare recipients and the efficacy of welfare. We use treatment assignment to the high or low condition to instrument for beliefs. We exclude the control group from the analyses. We control for demographic characteristics when conducting the pooled analysis. Black respondents are excluded from these analyses. Standard errors are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Finally, we test the robustness of the worthiness result using the same checks that we describe in Section 3.3.2 (see Table A16 for the results). In other words, we take the pooled IV regression that estimates the LATE of beliefs on perceived worthiness, and re-run this

---

[37]We cannot identify the impact of worthiness on redistribution preferences since we do not have an instrument that is related to worthiness and unrelated to redistribution preferences.

specification while dropping different participants (e.g., those who seem inattentive). As the table reveals, the result is largely robust to alternative specifications.

3.4 Priming respondents to think about race

The previous section studies how attitudes towards welfare depend on the precise beliefs individuals hold about the ethnic distribution of welfare recipients. Inspired by Alesina et al. (2021), we now study whether simply prompting White respondents to think about the share of welfare recipients who are from different races also changes their attitudes towards redistribution. To do so, we turn our attention to the second experiment, where we randomly varied whether the question about the ethnic distribution of welfare recipients came before or after the questions about welfare support (for those in the control group).

Table 4: The Effect of Priming on Giving

	Pro-welfare donation	Anti-welfare donation
Priming condition	0.0488**	-0.0033
	(0.0245)	(0.0150)
Constant	0.2530***	0.0828***
	(0.0174)	(0.0106)
Observations	1,330	1,330
R^2	0.003	0.000

Notes. In this table, we present the effect of 'priming' respondents to think about race on their propensity to donate to the pro- or anti-welfare organizations. We do this by comparing the donation rates among those who were randomly asked to donate before or after being asked about the ethnic distribution of welfare recipients. Our sample is restricted to White respondents in Experiment 2. These respondents were all allocated to the control group (within which we randomly varied whether the question about the racial composition of welfare recipients was asked before or after we measured their support for welfare). Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 4 shows that those who were not primed to think about race (*i.e.*, those who were asked about the race/ethnicities of welfare recipients after being asked to donate) are 4.9 percentage points more likely to donate to the pro-welfare organization ($p = 0.047$). Echoing the findings of Alesina et al. (2021), this suggests that simply getting individuals to think about race can dampen their support for redistributive programs. This may be because they believe that welfare recipients are disproportionately likely to be Black (see Section 3.1) and specifically do not want to redistribute funds to Black Americans (see

Section 3.3).[38]

3.5   The effects of correcting beliefs

We now examine the effect of providing accurate information about the racial distribution of welfare recipients. In light of our previous findings, it is unclear what effect this might be expected to have. On the one hand, since people generally overestimate the share of welfare recipients who are Black (Section 3.1), information provision should be expected to reduce beliefs about the share. This in turn might be expected to boost support for welfare (Section 3.3). On the other hand, providing such information also makes the issue of race more salient, which might be expected to decrease support for welfare (Section 3.4). It is therefore unclear what effect such information should have on net.

To study this question, we compare support for welfare among those in the control group (*i.e.*, those who received no information) to those in the 'low' treatment group in experiment 1 (as these individuals were presented with an approximately accurate signal about the share of welfare recipients who are Black, namely that this figure is around 20%). As Table 5 reveals, we do not find any significant effect of the treatment on the share who donate to the pro- or anti-welfare organizations. This result suggests that the 'belief updating' and 'priming' effects may be almost exactly canceling one another out.

Table 5: The effects of correcting beliefs

	Pro-welfare donation	Anti-welfare donation
Low signal	0.0060	-0.0149
	(0.0187)	(0.0124)
Constant	0.3530***	0.1210***
	(0.0133)	(0.0088)
Observations	2,627	2,627
R^2	0.0000	0.0010

Notes. In this table, we present the effect of providing White respondents with accurate information about the share of welfare recipients that is Black (*i.e.*, telling them that this figure is 21%). To do this, we compare donation rates among White respondents in the control group and the 'low' signal group in Experiment 1. Standard errors in parentheses (*$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

[38]As one would expect, we do not find that priming influences beliefs about the share of welfare recipients who are Black—suggesting that this is not the channel through which priming effects operate.

4 Conclusion

It has long been hypothesized that beliefs about the racial distribution of welfare recipients influences support for welfare. Our study tests this hypothesis using the tools of randomized inference. We find evidence in two experiments that increasing beliefs about the share that is Black reduces support for welfare among White survey respondents. Moreover, we find that priming White participants to think about race can also reduce their support for redistribution.

Our study does not directly address how these findings could affect welfare policy. Nonetheless, theoretical work suggests that, under political competition, the preferences of voters can have a substantial impact on policies related to redistribution (Lee et al., 2006). It is therefore possible that the preferences and exaggerated beliefs that we uncover here affect actual levels of US welfare spending.

Assuming that welfare preferences do indeed affect welfare policy, our findings suggest that policymakers need to tread a fine line if they wish to use race-related information to alter welfare support. On the one hand, our findings suggest that, in situations where the issue of race is already salient, policymakers can increase support for welfare spending by correcting beliefs about the the racial composition of welfare recipients. On the other hand, our findings also suggest that such a strategy might backfire if it makes the issue of race more salient. In light of these points, policymakers may wish to communicate rather differently depending on whether the issue of race is already a prominent one within debates about welfare.

Our findings highlight several areas for future research. First, it would be worthwhile to attempt to learn more about the precise mechanisms that explain why White respondents prefer redistribution to White welfare recipients. For instance, one might investigate this by attempting to hold relevant factors fixed (*e.g.*, perceived worthiness), while randomly varying beliefs about the ethnic distribution of welfare recipients. Having said this, such studies are likely to be challenging to implement given both the difficulty of influencing the relevant beliefs within an experiment, and the difficulty of measuring such beliefs in an accurate way.

Second, future research could address the extent to which our treatments alter beliefs about the ethnic composition of welfare recipients in the longer run; and also the extent to which such belief revisions persistently alter welfare support. This could be done by employing our basic research design over a longer time period (similar to Abramitzky et al. (2020); Giuliano and Tabellini (2020); Brown et al. (2021); Bursztyn et al. (2021);

Calderon et al. (2021)).

It would also be interesting to study whether the effect of beliefs about the share of welfare recipients that is Black on welfare support depends on *how* one comes to hold these beliefs. This topic could be studied by employing our research design while also varying how the researcher shocks participant beliefs. For example, one could vary who conveys the information and the context within which it is conveyed.

Finally, it would be interesting to better understand how basic beliefs about the racial composition of welfare recipients are formed. In our study, we obtain suggestive evidence that the components of Bayes rule, such as the 'base rate' of different ethnic groups, alter beliefs about the ethnic distribution of welfare recipients in the expected direction. Further research could shed more light on this issue by explicitly randomizing the various components of Bayes rule and examining how this alters beliefs.

References

Abramitzky, R. and Boustan, L. (2017). Immigration in american economic history. Journal of Economic Literature, 55(4):1311–45.

Abramitzky, R., Boustan, L., and Eriksson, K. (2020). Do immigrants assimilate more slowly today than in the past? American Economic Review: Insights, 2(1):125–41.

Alesina, A. and Ferrara, E. L. (2005). Ethnic diversity and economic performance. Journal of Economic Literature, 43(3):762–800.

Alesina, A., Ferroni, M. F., and Stantcheva, S. (2021). Perceptions of racial gaps, their causes, and ways to reduce them. Technical report, National Bureau of Economic Research.

Alesina, A. and Giuliano, P. (2011). Preferences for redistribution. In Handbook of Social Economics, volume 1, pages 93–131. Elsevier.

Alesina, A. and Glaeser, E. L. (2004). Fighting poverty in the US and Europe: A world of difference. Oxford University Press.

Alesina, A., Miano, A., and Stantcheva, S. (2018). Immigration and redistribution. Technical report, National Bureau of Economic Research.

Alesina, A. and Stantcheva, S. (2020). Diversity, immigration, and redistribution. In AEA Papers and Proceedings, volume 110, pages 329–34.

Alesina, A. F., Glaeser, E. L., and Sacerdote, B. (2001). Why doesn't the us have a european-style welfare system?

Altonji, J. G. and Doraszelski, U. (2005). The role of permanent income and demographics in black/white differences in wealth. Journal of Human Resources, 40(1):1–30.

Andre, P., Pizzinelli, C., Roth, C., and Wohlfart, J. (2019). Subjective models of the macroeconomy: Evidence from experts and a representative sample.

Armantier, O., Nelson, S., Topa, G., Van der Klaauw, W., and Zafar, B. (2016). The price is right: Updating inflation expectations in a randomized price information experiment. Review of Economics and Statistics, 98(3):503–523.

Armona, L., Fuster, A., and Zafar, B. (2019). Home price expectations and behaviour: Evidence from a randomized information experiment. Review of Economic Studies, 86(4):1371–1410.

Arthur, D. and Edwards-Levy, A. (2018). American welfare perceptions survey. Huffington Post.

Ashok, V., Kuziemko, I., and Washington, E. (2015). Support for redistribution in an age of rising inequality: New stylized facts and some tentative explanations. Technical report, National Bureau of Economic Research.

Augenblick, N. and Rabin, M. (2021). Belief movement, uncertainty reduction, and rational updating. Quarterly Journal of Economics, 136(2):933–985.

Barsky, R., Bound, J., Charles, K. K., and Lupton, J. P. (2002). Accounting for the black–white wealth gap: a nonparametric approach. Journal of the American Statistical Association, 97(459):663–673.

Bayer, P. and Charles, K. K. (2018). Divergent paths: A new perspective on earnings differences between black and white men since 1940. Quarterly Journal of Economics, 133(3):1459–1501.

Bobo, L. and Kluegel, J. R. (1993). Opposition to race-targeting: self-interest, stratification ideology, or racial attitudes? American Sociological Review, pages 443–464.

Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2019a). Inaccurate statistical discrimination: An identification problem. Technical report, National Bureau of Economic Research.

Bohren, J. A., Hull, P., and Imas, A. (2022). Systemic discrimination: Theory and measurement. Technical report, National Bureau of Economic Research.

Bohren, J. A., Imas, A., and Rosenberg, M. (2019b). The dynamics of discrimination: Theory and evidence. American Economic Review, 109(10):3395–3436.

Bonomi, G., Gennaioli, N., and Tabellini, G. (2021). Identity, beliefs, and political conflict. Quarterly Journal of Economics, 136(4):2371–2411.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. Quarterly Journal of Economics, 131(4):1753–1794.

Boustan, L. P. (2010). Was postwar suburbanization "white flight"? evidence from the black migration. Quarterly Journal of Economics, 125(1):417–443.

Brodish, A. B., Brazy, P. C., and Devine, P. G. (2008). More eyes on the prize: Variability in white americans' perceptions of progress toward racial equality. Personality and Social Psychology Bulletin, 34(4):513–527.

Brown, J. R., Enos, R. D., Feigenbaum, J., and Mazumder, S. (2021). Childhood cross-ethnic exposure predicts political behavior seven decades later: Evidence from linked administrative data. Science Advances, 7(24):eabe8432.

Brown-Iannuzzi, J. L., Dotsch, R., Cooley, E., and Payne, B. K. (2017). The relationship between mental representations of welfare recipients and attitudes toward welfare. Psychological science, 28(1):92–103.

Bursztyn, L., Chaney, T., Hassan, T. A., and Rao, A. (2021). The immigrant next door: Exposure, prejudice, and altruism. Technical report, National Bureau of Economic Research.

Bursztyn, L., Egorov, G., and Fiorin, S. (2020a). From extreme to mainstream: The erosion of social norms. American Economic Review, 110(11):3522–48.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020b). Misperceived social norms: Women working outside the home in saudi arabia. American Economic Review, 110(10):2997–3029.

Bursztyn, L., Haaland, I. K., Rao, A., and Roth, C. P. (2020c). Disguising prejudice: Popular rationales as excuses for intolerant expression. Technical report, National Bureau of Economic Research.

Bursztyn, L. and Yang, D. Y. (2021). Misperceptions about others. Annual Review of Economics, 14.

Calderon, A., Fouka, V., and Tabellini, M. (2021). Racial diversity and racial policy preferences: The great migration and civil rights. Technical report, National Bureau of Economic Research.

Callaghan, B., Harouni, L., Dupree, C. H., Kraus, M. W., and Richeson, J. A. (2021). Testing the efficacy of three informational interventions for reducing misperceptions of the black–white wealth gap. Proceedings of the National Academy of Sciences, 118(38).

Charles, K. K. and Guryan, J. (2008). Prejudice and wages: an empirical assessment of becker's the economics of discrimination. Journal of Political Economy, 116(5):773–809.

Charles, K. K. and Hurst, E. (2002). The transition to home ownership and the black-white wealth gap. Review of Economics and Statistics, 84(2):281–297.

Chiteji, N. S. and Stafford, F. P. (1999). Portfolio choices of parents and their children as young adults: Asset accumulation by african-american families. American Economic Review, 89(2):377–380.

Cook, F. L. and Barrett, E. J. (1992). Support for the American welfare state: The views of Congress and the public. Columbia University Press.

Crowne, D. P. and Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24(4):349.

Dahlberg, M., Edmark, K., and Lundqvist, H. (2012). Ethnic diversity and preferences for redistribution. Journal of Political Economy, 120(1):41–76.

Delaney, A. and Edwards-Levy, A. (2018). Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions. Huffington Post.

Delavande, A. and Zafar, B. (2018). Information and anti-american attitudes. Journal of Economic Behavior & Organization, 149:1–31.

Derenoncourt, E., Kim, C. H., Kuhn, M., and Schularick, M. (2021). The racial wealth gap, 1860-2020. Manuscript, Princeton University and University of Bonn.

Dhar, D., Jain, T., and Jayachandran, S. (2018). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. Technical report, National Bureau of Economic Research.

Dustmann, C., Vasiljeva, K., and Piil Damm, A. (2019). Refugee migration and electoral outcomes. Review of Economic Studies, 86(5):2035–2091.

Eibach, R. P. and Keegan, T. (2006). Free at last? social dominance, loss aversion, and white and black americans' differing assessments of racial progress. Journal of Personality and Social Psychology, 90(3):453.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. American Economic Journal: Microeconomics, 3:114–38.

Feagin, J. (2013). Systemic racism: A theory of oppression. Routledge.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. Journal of Public Economics, 82(2):225–246.

Fong, C. M. and Luttmer, E. F. (2009). What determines giving to hurricane katrina victims? experimental evidence on racial group loyalty. American Economic Journal: Applied Economics, 1(2):64–87.

Fong, C. M. and Luttmer, E. F. (2011). Do fairness and race matter in generosity? evidence from a nationally representative charity experiment. Journal of Public Economics, 95(5-6):372–394.

Fouka, V., Mazumder, S., and Tabellini, M. (2022). From immigrants to americans: Race and assimilation during the great migration. Review of Economic Studies, 89(2):811–842.

Fuster, A., Perez-Truglia, R., Wiederholt, M., and Zafar, B. (2018). Expectations with endogenous information acquisition: An experimental investigation. Review of Economics and Statistics, pages 1–54.

Gilens, M. (1995). Racial attitudes and opposition to welfare. Journal of Politics, 57(4):994–1014.

Gilens, M. (1996). "race coding" and white opposition to welfare. American Political Science Review, 90(3):593–604.

Gilens, M. (2009). Why Americans hate welfare: Race, media, and the politics of antipoverty policy. University of Chicago Press.

Gittleman, M. and Wolff, E. N. (2004). Racial differences in patterns of wealth accumulation. Journal of Human Resources, 39(1):193–227.

Giuliano, P. and Tabellini, M. (2020). The seeds of ideology: Historical immigration and political preferences in the united states. Technical report, National Bureau of Economic Research.

Glaeser, E. L. (2005). The political economy of hatred. Quarterly Journal of Economics, 120(1):45–86.

Gross, K. and Wronski, J. (2021). Helping the homeless: The role of empathy, race and deservingness in motivating policy support and charitable giving. Political Behavior, 43(2):585–613.

Gupta, N., Rigotti, L., and Wilson, A. (2021). The experimenters' dilemma: Inferential preferences over populations. arXiv preprint arXiv:2107.05064.

Haaland, I. and Roth, C. (2021). Beliefs about racial discrimination and support for pro-black policies. Review of Economics and Statistics, pages 1–38.

Haaland, I., Roth, C., and Wohlfart, J. (2020). Designing information provision experiments.

Harell, A., Soroka, S., and Iyengar, S. (2016). Race, prejudice and attitudes toward redistribution: A comparative experimental approach. European Journal of Political Research, 55(4):723–744.

Henry, P., Reyna, C., and Weiner, B. (2004). Hate welfare but help the poor: how the attributional content of stereotypes explains the paradox of reactions to the destitute in america 1. Journal of Applied Social Psychology, 34(1):34–58.

Hopkins, D. J., Sides, J., and Citrin, J. (2019). The muted consequences of correct information about immigration. Journal of Politics, 81(1):315–320.

Kraus, M. W., Onyeador, I. N., Daumeyer, N. M., Rucker, J. M., and Richeson, J. A. (2019). The misperception of racial economic inequality. Perspectives on Psychological Science, 14(6):899–921.

Kraus, M. W., Rucker, J. M., and Richeson, J. A. (2017). Americans misperceive racial economic equality. Proceedings of the National Academy of Sciences, 114(39):10324–10331.

Kuo, E. E., Kraus, M. W., and Richeson, J. A. (2020). High-status exemplars and the misperception of the asian-white wealth gap. Social Psychological and Personality Science, 11(3):397–405.

Lee, W., Roemer, J., and Van der Straeten, K. (2006). Racism, xenophobia, and redistribution. Journal of the European Economic Association, 4(2-3):446–454.

Lee, W. and Roemer, J. E. (2006). Racism and redistribution in the united states: A solution to the problem of american exceptionalism. Journal of Public Economics, 90(6-7):1027–1052.

Luttmer, E. F. (2001). Group loyalty and the taste for redistribution. Journal of Political Economy, 109(3):500–528.

Nadeau, R., Niemi, R. G., and Levine, J. (1993). Innumeracy about minority populations. Public Opinion Quarterly, 57(3):332–347.

Onyeador, I. N., Daumeyer, N. M., Rucker, J. M., Duker, A., Kraus, M. W., and Richeson, J. A. (2021). Disrupting beliefs in racial progress: Reminders of persistent racism alter perceptions of past, but not current, racial economic equality. Personality and Social Psychology Bulletin, 47(5):753–765.

O'Brien, R. L. (2017). Redistribution and the new fiscal sociology: Race and the progressivity of state and local taxes. American Journal of Sociology, 122(4):1015–1049.

Peer, E., Brandimarte, L., Samat, S., and Acquisti, A. (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology, 70:153–163.

Quadagno, J. S. et al. (1994). The color of welfare: How racism undermined the war on poverty. Oxford University Press.

Ribar, D. C. and Wilhelm, M. O. (1999). The demand for welfare generosity. Review of Economics and Statistics, 81(1):96–108.

Roth, C., Settele, S., and Wohlfart, J. (2021). Beliefs about public debt and the demand for government spending. Journal of Econometrics.

Roth, C. and Wohlfart, J. (2020). How do expectations about the macroeconomy affect personal expectations and behavior? Review of Economics and Statistics, 102(4):731–748.

Rothstein, R. (2017). The color of law: A forgotten history of how our government segregated America. Liveright Publishing.

Sarsons, H. (2017). Interpreting signals in the labor market: evidence from medical referrals. Job Market Paper, pages 141–45.

Steele, C. M. (2011). Whistling Vivaldi: How stereotypes affect us and what we can do. WW Norton & Company.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. Review of Economic Studies, 87(1):454–486.

Tesler, M. (2015a). The conditions ripe for racial spillover effects. Political Psychology, 36:101–117.

Tesler, M. (2015b). Priming predispositions and changing policy positions: An account of when mass opinion is primed or changed. American Journal of Political Science, 59(4):806–824.

USCB (2018). Survey of income and program participation.

Washington, E. (2006). How black candidates affect voter turnout. Quarterly Journal of Economics, 121(3):973–998.

Wetts, R. and Willer, R. (2018). Privilege on the precipice: Perceived racial status threats lead white americans to oppose welfare programs. Social Forces, 97(2):793–822.

Appendices

A   Tables and figures

Table A1: Balance table (experiment 1)

	[1]	[2]	[3]	[1] vs [2]	[1] vs [3]	[2] vs [3]	Joint test
Female	0.584	0.597	0.597	0.433	0.409	0.965	0.649
Age	34.5	34.4	33.6	0.764	0.023	0.049	0.050
Moderate	0.259	0.267	0.275	0.546	0.251	0.584	0.517
Conservative	0.153	0.168	0.145	0.200	0.478	0.046	0.128
Ever on welfare	0.564	0.542	0.554	0.181	0.540	0.469	0.407
College educated	0.593	0.589	0.587	0.793	0.699	0.900	0.925
Voted	0.840	0.845	0.846	0.702	0.636	0.927	0.881
n	1926	1943	1924				

Notes. This table presents the average characteristics of respondents in the treatment and control groups of experiment 1. Column (1) refers to the control group, column (2) to the low signal group, and column (3) to the high signal group. The subsequent columns present p-values corresponding to the test of equality between each pair of groups. The final column presents the result of a joint test of equality across all groups.

Table A2: Balance table (experiment 2)

	[1]	[2]	[3]	[1] vs [2]	[1] vs [3]	[2] vs [3]	Joint test
Female	0.680	0.644	0.663	0.047	0.357	0.288	0.139
Age	35.4	35.9	35.3	0.365	0.949	0.326	0.546
Moderate	0.321	0.336	0.356	0.402	0.052	0.271	0.150
Conservative	0.321	0.305	0.291	0.384	0.094	0.423	0.246
Ever on welfare	0.423	0.418	0.436	0.812	0.513	0.373	0.652
College educated	0.548	0.563	0.553	0.438	0.779	0.620	0.734
Voted	0.842	0.829	0.826	0.369	0.259	0.818	0.494
n	1331	1324	1327				

Notes. This table presents the average characteristics of respondents in the treatment and control groups of experiment 2. Column (1) refers to the control group, column (2) to the low signal group, and column (3) to the high signal group. The subsequent columns present p-values corresponding to the test of equality between each pair of groups. The final column presents the result of a joint test of equality across all groups.

Table A3: Descriptive statistics (pooled)

Variables	n	Mean	σ	Min	Max
Conservative	9,775	0.217	0.412	0	1
Moderate	9,775	0.296	0.456	0	1
College educated	9,775	0.575	0.494	0	1
Voted in 2020 election	9,775	0.839	0.368	0	1
Female	9,775	0.621	0.485	0	1
Age (year)	9,775	34.73	12.93	18	92
Ever been on welfare	9,775	0.501	0.500	0	1
Suspected that study about race	9,775	0.388	0.487	0	1
Belief % of welfare recipients Black (post)	9,775	32.57	16.36	2	100
Belief % of welfare recipients Black (pre)	9,775	37.11	14.70	0	100
Belief % of welfare recipients White (post)	9,775	49.27	19.02	0	90
Belief % of welfare recipients White (pre)	9,775	38.52	16.15	0	93
Confident in post-treatment beliefs	9,775	0.402	0.490	0	1
Confident in pre-treatment beliefs	9,775	0.366	0.482	0	1
Belief % of White on welfare	9,775	40.26	20.62	0	100
Belief % of Black on welfare	9,775	32.70	16.49	0	100
Belief % of Other on welfare	9,775	20.39	15.20	0	100
Belief % of Latino on welfare	9,775	23.53	15.13	0	100
Share of US pop. that is Black	9,775	27.02	12.22	0	100
Share of US pop. that is White	9,775	52.46	14.81	0	94
Share of US pop. that is Other	9,775	20.52	10.72	0	100
Share of US pop. that is Latino	9,775	25.61	13.89	0	100
Implied share on welfare	9,775	0.355	0.138	0	1
Implied share of recipients Black	9,773	27.54	17.72	0	100
Implied share of recipients White	9,773	58.54	22.37	0	98.86
Previously thought about race of recipients	3,982	0.454	0.498	0	1
Took race into account when donating	3,982	0.331	0.471	0	1
Believes welfare low because of racism	3,982	0.298	0.457	0	1
Donated against welfare	9,775	0.102	0.303	0	1
Donated for welfare	9,775	0.323	0.468	0	1
Believes that welfare is effective	9,775	0.595	0.491	0	1
Believes that welfare recipients are worthy	9,775	0.473	0.499	0	1

Notes. This table presents descriptive statistics for all respondents in the two experiments.

Table A4: Beliefs by sub-group

Sub-group	% Black	% White	% Hispanic
White respondents	37.52	38.42	26.72
	(14.54)	(15.87)	(15.25)
Black respondents	35.35	38.92	28.63
	(15.26)	(17.32)	(15.39)
Female respondents	38.16	37.86	27.58
	(14.63)	(15.76)	(15.25)
Male respondents	35.39	39.60	26.26
	(14.65)	(16.72)	(15.34)
Conservative respondents	40.57	35.79	28.22
	(15.71)	(15.40)	(16.65)
Moderate respondents	38.38	36.55	27.74
	(14.59)	(15.49)	(15.38)
Liberal respondents	34.80	40.92	26.17
	(13.89)	(16.52)	(14.55)
With a college degree	36.03	40.34	26.64
	(14.83)	(16.41)	(15.58)
Without a college degree	38.58	36.05	27.68
	(14.40)	(15.46)	(14.88)
Have ever been on welfare	36.00	39.38	28.19
	(14.44)	(16.21)	(16.45)
Have never been on welfare	38.23	37.65	25.97
	(14.88)	(16.05)	(13.96)
Voted	36.62	39.10	26.80
	(14.63)	(16.23)	(15.30)
Did not vote	39.68	35.48	28.55
	(14.80)	(15.40)	(15.16)

Notes. This table presents the average beliefs of different sub-groups regarding the percentage of welfare recipients who are Black, White and Hispanic/Latino respectively. We used data from both experiments when computing these statistics. Standard errors in parentheses.

Table A5: Predictors of pre-treatment beliefs

Variables	Beliefs about % who are Black
Female	3.002***
	(0.298)
Age	-0.0867***
	(0.0116)
Moderate	3.391***
	(0.342)
Conservative	6.056***
	(0.400)
Ever on welfare	-2.043***
	(0.296)
College	-2.137***
	(0.305)
Voted	-1.696***
	(0.421)
Black	-1.822***
	(0.391)
Constant	39.96***
	(0.590)
n	9,775
R^2	0.058

Notes. This table examines the predictors of beliefs about the share of welfare recipients who are Black. The estimates are obtained by regressing beliefs on the demographic variables detailed in the first column. A dummy indicating that someone is liberal is omitted as we have dummies indicating identification as a conservative or moderate. We included the entire sample from both experiments when conducting this regression. Standard errors in parentheses (*$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table A6: Bayesian updating

	log(% welfare recipients that are Black)
log(% that are Black)	0.242***
	(0.00845)
log(% Black Americans on welfare)	0.623***
	(0.00711)
log(% Americans on welfare)	-0.427***
	(0.00842)
n	9,765
R^2	0.984

Notes. This table presents the results of regressing the logarithm of beliefs about the percentage of welfare recipients that are Black on three variables: (1) the logarithm of beliefs about the percentage of the population that are Black (2) the logarithm of beliefs about the percentage of Black Americans who are on welfare (3) the logarithm of beliefs about the percentage of Americans who are on welfare. If all participants use Bayes Rule, the estimated coefficients should be 1, 1, and −1 respectively. Standard errors in parentheses (*$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table A7: Treatment effects on beliefs

Variables	% who are Black (exp 1)	% who are Black (exp 2)
High estimate	-4.235***	11.34***
	(0.527)	(0.438)
Low estimate	-8.403***	-25.51***
	(0.556)	(0.489)
Constant	35.16***	39.48***
	(0.401)	(0.394)
n	3,948	3,981
R^2	0.059	0.660

Notes. This table presents the effect of treatment assignment on participants' beliefs about the percentage of welfare recipients who are Black. The omitted group in both regressions is the control (*i.e.*, we control for assignment to the high and low condition, but not to the control condition). The constant term represents the average in the control group. The first column presents the effects from the first experiment; and the second column presents the effects from the second experiment. The first-stage regressions used in the IV analysis can be obtained by examining the differences between the high and low signal groups in this table. Standard errors are in parentheses (*$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table A8: Confidence in beliefs

Variables	Confident (exp. 1)	Confident (exp. 2)	Confident (exp. 1)	Confident (exp. 2)
High estimate	-0.0334*	0.0278	0.0529***	0.0655***
	(0.0193)	(0.0188)	(0.0191)	(0.0185)
Low estimate			0.0864***	0.0377**
			(0.0192)	(0.0184)
Constant	0.459***	0.361***	0.373***	0.323***
	(0.0137)	(0.0132)	(0.0134)	(0.0128)
n	2,646	2,651	3,948	3,981
R^2	0.001	0.001	0.005	0.003

Notes. This table presents the effect of treatment assignment on participants' confidence in their beliefs about the ethnic distribution of welfare recipients. Participants are coded as 'confident' if they stated that they are 'confident' or 'very confident' in their beliefs. The first two regressions drop participants in the control group; whereas the last two regressions include it (and use the control group as the omitted category). Standard errors in parentheses ($*p < 0.1$, $** p < 0.05$, $*** p < 0.01$).

Table A9: Heterogeneity by political affiliation

Donated	LATE	σ	Constant	Restriction	Experiment
For welfare	-0.00696	0.00913	0.39171	Conservatives	1
For welfare	-0.00731	0.00499	0.59147	Non-conservatives	1
For welfare	-0.00011	0.00054	0.08357	Conservatives	2
For welfare	-0.00147	0.00059	0.40244	Non-conservatives	2
Against welfare	0.01429	0.01031	-0.23052	Conservatives	1
Against welfare	0.00619	0.00316	-0.07310	Non-conservatives	1
Against welfare	-0.00005	0.00068	0.12956	Conservatives	2
Against welfare	0.000073	0.00031	0.06646	Non-conservatives	2

Notes. This table presents the LATE coefficients for the effects of beliefs about the share of welfare recipients that are Black on the share that either donated to the pro- or anti-welfare organizations, disaggregated by those who identify as conservatives and non-conservatives (i.e., liberal or moderate). Black respondents are excluded from these analyses. These analyses are conducted using the same methods as in Table 1.

Table A10: Pooled estimates

	Pro welfare ITT	Pro welfare LATE	Anti welfare ITT	Anti welfare LATE
Belief about % Black		-0.00120***		-4.37e-05
		(0.000439)		(0.000292)
High estimate	-0.0387**		0.00118	
	(0.0164)		(0.0108)	
Experiment 1	0.0160		0.0298**	
	(0.0176)		(0.0116)	
High estimate × exp 1	0.00556		0.0307*	
	(0.0243)		(0.0166)	
Female	0.0235*	0.0226*	-0.0211**	-0.0268***
	(0.0125)	(0.0124)	(0.00870)	(0.00876)
Age	-0.00107**	-0.00103**	-0.000318	-0.000242
	(0.000489)	(0.000483)	(0.000340)	(0.000338)
Moderate	-0.163***	-0.165***	0.0389***	0.0355***
	(0.0158)	(0.0155)	(0.0103)	(0.00993)
Conservative	-0.304***	-0.305***	0.0781***	0.0814***
	(0.0162)	(0.0141)	(0.0131)	(0.0117)
Ever on welfare	0.0401***	0.0409***	-0.0160*	-0.0100
	(0.0125)	(0.0125)	(0.00855)	(0.00853)
College	0.00716	0.00708	0.0150*	0.0180**
	(0.0128)	(0.0128)	(0.00894)	(0.00893)
Right wing news	-0.00186		0.0300***	
	(0.0144)		(0.0113)	
Voted	0.0109	0.0112	0.0260**	0.0286**
	(0.0179)	(0.0179)	(0.0118)	(0.0119)
Constant	0.433***	0.459***	0.0424**	0.0694***
	(0.0276)	(0.0285)	(0.0185)	(0.0193)
n	5,297	5,297	5,297	5,297
R^2	0.085	0.083	0.024	0.015

Notes. In this table, we present the ITT and LATE estimates of the effects of beliefs about the share of welfare recipients that is Black on support for welfare. All regressions are conducted using all White participants in both experiments (excluding the control groups). When conducting the LATE estimation, we use assignment to the high or low treatment groups, assignment to experiment 1 or 2, and the interaction between treatment assignment and experimental assignment as instruments for beliefs. We control for observables in all regressions (see the list of variables in the table). Standard errors in parentheses (*$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

46

Table A11: ITT and LATE for Black respondents

Variables	Pro-welfare donation ITT	Pro-welfare donation LATE	Anti-welfare donation ITT	Anti-welfare donation LATE
High estimate	-0.00567		-0.00236	
	(0.0276)		(0.0175)	
Beliefs about % who are Black		-0.00151		-0.000628
		(0.00736)		(0.00466)
Constant	0.372***	0.414*	0.105***	0.123
	(0.0195)	(0.218)	(0.0124)	(0.138)
n	1,221	1,221	1,221	1,221
R^2	0.000		0.000	

Notes. In this table, we present ITT and LATE estimates for Black respondents. The ITT estimates reveal the effect of assignment to the high signal group (relative to the low signal group) on participants' propensity to donate to one of the two charities. The LATE estimates reveal the effect of beliefs about the share of welfare recipients who are Black on their propensity to donate to the charities (using treatment assignment as an instrument). We exclude the control group and White respondents from the analysis. Standard errors in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table A12: Predictors of welfare support

	Pro-welfare donation	Anti-welfare donation
Female	-0.00215	-0.0274***
	(0.00944)	(0.00648)
Age	-0.00152***	-0.000344
	(0.000364)	(0.000249)
Moderate	-0.161***	0.0392***
	(0.0112)	(0.00722)
Conservative	-0.279***	0.0832***
	(0.0110)	(0.00893)
Ever on welfare	0.0445***	-0.00343
	(0.00941)	(0.00631)
College	0.0155	0.0182***
	(0.00966)	(0.00647)
Vote	0.0140	0.0192**
	(0.0130)	(0.00855)
Black	0.0142	0.0121
	(0.0128)	(0.00784)
Constant	0.440***	0.0744***
	(0.0189)	(0.0124)
n	9,775	9,775
R^2	0.069	0.015

Notes. This table regresses our measures of welfare support on the demographic variables specified in the first column. The subsequent two columns specify how these variables alter the probability that a participant chooses to donate to the pro- and anti-welfare charities respectively. The sample includes all participants across both experiments. Standard errors in parentheses (*$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table A13: Robustness checks (experiment 1)

Robustness check	Coefficient	Standard error	n
Drop if time taken <6min	0.00629	0.00316	2,571
" 7min	0.00580	0.00318	2,422
" 8min	0.00611	0.00323	2,235
" 9min	0.00512	0.00326	1,996
" 10min	0.00332	0.00341	1,684
Drop if treat time <5sec	0.00649	0.00316	2,554
" 10sec	0.00459	0.00304	1,994
" 15sec	0.00156	0.00379	1,300
" 20sec	-0.000930	0.00423	823
Drop if failed quiz	0.00413	0.0088	1,209
Drop if belief = 0/100	0.00716	0.00315	2,646
Drop if % win lottery <1%	0.00703	0.00337	2,457
" 2%	0.00990	0.00438	1,810
" 3%	0.00975	0.00461	1,591
" 4%	0.00906	0.0049	1,424
" 5%	0.00875	0.00489	1,313
Drop if understood purpose	0.00677	0.00408	1,293
Re-weighted analysis	0.00457	0.0034	2,646
Control for confidence	0.00739	0.00315	2,646
Instrument for confidence	0.0263	0.0465	3,948
Use implied beliefs	0.0300	0.0207	2,645
Include control group	0.00178	0.0015	3,948

Notes. This table presents the results of a series of robustness checks using data from experiment 1. The first column specifies the robustness check, and the second column presents the LATE estimated obtained following the check. The subsequent columns respectively reveal the resulting standard error associated with the LATE estimate and the resulting sample size after the relevant participants have been dropped.

Table A14: Robustness checks (experiment 2)

Robustness check	Coefficient	Standard error	n
Drop if time taken <6min	-0.000972	0.000472	2,626
" 7min	-0.000931	0.000481	2,556
" 8min	-0.000864	0.000500	2,410
" 9min	-0.00111	0.000533	2,196
" 10min	-0.00121	0.000574	1,943
Drop if treat time <5sec	-0.0009517	0.0004742	2,591
" 10sec	-0.0010613	0.0005559	1,982
" 15sec	-0.0003876	0.0007081	1,282
" 20sec	0.0006397	0.0009165	826
Drop if failed quiz	-0.000360	0.000844	800
Drop if belief = 0/100	-0.000979	0.000469	2,651
Drop soc. des. bias	-0.00127	0.000512	2,153
Drop if understood purpose	-0.000928	0.000527	1,981
Re-weighted analysis	-0.00162	0.000547	2,651
Control for confidence	-0.000951	0.000469	2,651
Instrument for confidence	-0.000835	0.000481	3,981
Use implied beliefs	-0.00140	0.000674	2,650
Include control group	-0.00907	0.0004603	3,981

Notes. This table presents the results of a series of robustness checks using data from experiment 2. The first column specifies the robustness check, and the second column presents the LATE estimated obtained following the check. The subsequent columns respectively reveal the resulting standard error associated with the LATE estimate and the resulting sample size after the relevant participants have been dropped.

Table A15: Association between beliefs and welfare support for White respondents

Variables	Pro-welfare donation	Anti-welfare donation	Pro-welfare donation	Anti-welfare donation
Belief % Black (post-treatment)	-0.000679** (0.000293)	8.83e-05 (0.000205)		
Belief % Black (pre-treatment)			-0.00106*** (0.000346)	0.000267 (0.000250)
Female	0.0115 (0.0102)	-0.0253*** (0.00712)	0.0139 (0.0103)	-0.0261*** (0.00713)
Age	-0.00136*** (0.000390)	-0.000239 (0.000278)	-0.00142*** (0.000391)	-0.000219 (0.000280)
Moderate	-0.180*** (0.0126)	0.0398*** (0.00805)	-0.178*** (0.0126)	0.0390*** (0.00810)
Conservative	-0.301*** (0.0116)	0.0851*** (0.00952)	-0.297*** (0.0118)	0.0838*** (0.00970)
Ever on welfare	0.0435*** (0.0102)	-0.00610 (0.00698)	0.0416*** (0.0103)	-0.00555 (0.00694)
College	0.0196* (0.0105)	0.0202*** (0.00721)	0.0184* (0.0105)	0.0207*** (0.00720)
Voted	0.0174 (0.0147)	0.0206** (0.00979)	0.0168 (0.0147)	0.0208** (0.00978)
Constant	0.454*** (0.0224)	0.0647*** (0.0152)	0.473*** (0.0243)	0.0572*** (0.0166)
n	7,929	7,929	7,929	7,929
R^2	0.085	0.016	0.086	0.016

Notes. This table presents the results of regressing support for welfare on beliefs about the share of welfare recipients that is Black, along with several demographic controls. The sample includes all White respondents in both experiments. Standard errors in parentheses (*$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table A16: Robustness checks for effect of beliefs on perceived worthiness

Robustness check	Coefficient	Standard error	n
Drop if time taken <6min	-.0010137**	.0004749	5,197
" 7min	-.0010127**	.000482	4,978
" 8min	-.0009642*	.0004993	4,645
" 9min	-.0012256**	.0005231	4,192
" 10min	-.0010966*	.000561	3,627
Drop if treat time <5sec	-.00093**	.000473	5,205
" 10sec	-.0009376**	.0004732	4,645
" 15sec	-.0008309*	.0004752	3,951
" 20sec	-.0008105*	.0004764	3,474
Drop if failed quiz	-.0002986	.0008984	2,009
Drop if belief = 0/100	-.0009432**	.0004732	5,297
Drop if understood purpose	-.0009979*	.0005485	3,274
Re-weighted analysis	-.0008098	.0005469	5,297
Control for confidence	-.000941**	.0004732	5,297
Use implied beliefs	-.001423**	.0006535	5,295
Include control group	-.0011104*	.000621	7,929

Notes. This table presents the results of a series of robustness checks using data from both experiment. The first column specifies the robustness check, and the second column presents the LATE estimated obtained following the check. The subsequent columns respectively reveal the resulting standard error associated with the LATE estimate and the resulting sample size after the relevant participants have been dropped. The baseline regression that we are testing the robustness of is the pooled specification that examines the effect of beliefs about the share that is Black on the perceived worthiness of welfare recipients.