Appendix A - Survey Questions

Subjects were shown two applicants for state aid with the following prompt:

"Researchers have been hired to consult with a nearby state's welfare agency. Below you will find two applicants for government assistance. The application information has been reducted to hide information that may identify individual applicants.

Each applicant has a state-assessed level of need of \$900 per month. Your task is to allocate \$1,500 between the two applicants. You can allocate any amount between \$0 and \$900 to each applicant. Any remaining funds will be used to offset the state's budget deficit."

Respondents were given three sliding scales to allocate funds. They could slide the bars or type numbers into the boxes on the right. At the bottom of the screen, the total amount they had awarded was shown relative to the full \$1,500 they had to distribute.



Figure 1: Sliding Scale Used by Respondents to Make Allocations

Appendix B - Preregistration

The analysis choices reported in the main text and our research hypotheses were pre-specified in the preregistration. We report our preregistration in this section. There are no departures from the pre-analysis plan. Some exploratory analysis is included in the main text and is denoted as such with a footnote. This experiment was Study 4 in a larger, grant-funded study, all of which was preregistered. As a result, the sections of the pre-registration relevant to this work are included.

EGAP PRE-REGISTRATION
TITLE OF STUDY: Public Aid Project
AUTHORIS), Couch (Dati Warran
AUTHOR(\$): Sarah "Dot" Warren
Is one of the study authors a university faculty member?
No.
Is this Registration Prospective or Retrospective?
Registration prior to researcher access to outcome data.
Is this an experimental study?
Yes
Date of start of study:
•
11/22/2022
Was this design presented at an EGAP meeting?
Was this design presented at an EGAP meeting?

Background and explanation of rationale

Background

Yes (below)

This pre-registration and pre-analysis plan is for a four-study survey to be fielded by YouGov in April of 2022. It is my intention that Studies 1 and 2 will be the basis of one paper, Study 3 the basis of one paper, and Study 4 the basis of one paper, resulting in three total papers. In the spirit of full transparency, this pre-analysis plan is broken down by study.

This study has an N=2150. This includes a nationally representative sample of 1,400 and an oversample of Low SES Americans of 750 to obtain appropriate variation in aid-status.

Explanation of Rationale (Study 4)

This study follows DeSante (2013) and Hayes et al. (2020). It seeks to address the question: "Do Americans punish women who apply for federal aid relative to men?" In other words, when comparing two otherwise identical applicants for federal aid, are Americans more generous toward male applicants? While a wealth of literature considers how, when, and why Black Americans are punished when they apply for welfare (Gilens 1999; Smith 1987; Sears et al. 1997), few have considered if Americans treat male and female aid applicants differently (but see Rabinowitz et al. 2009). Instead, nearly all experiments use female names in order to hold sex constant as they evaluate race-based punishment. Using names that are distinctly white according to Hayes and Mitchel's (2020) name-characteristics dataset, I hold race fixed as white and instead vary names by sex. This allows me to explore a few mechanisms by which evaluations are made, including sex, perceived competence, and quality evaluation.

Study 4

I experimentally manipulate the gender, quality, and competence cues on a pair of hypothetical aid applicants, identical in format to DeSante (2013). Using names that are distinctly white according to Hayes and Mitchel's (2020) name-characteristics dataset, I hold race fixed as white and instead vary names by sex.

This design allows me to examine respondents' generosity to women as compared to men, high quality women (men) as compared to low quality women (men), high (low) quality women as compared to high (low) quality men, low rated women (men) as compared to high rated women (men), and high (low) rated women as compared to high (low) rated men. Thus, I can examine not only to which groups respondents are most giving, but comment on the underlying causes of generosity (and punishment) based on sex, stated quality, or perceived competence and the interactions between these factors.

These names come from Hayes and Mitchell's (2020) name-characteristics dataset and are matched on important characteristics. Specifically, James and Sandra are rated highly in professionalism, competence, and work ethic, while Sammie and Misty are rated lower in all three categories. The figure below shows the full breakdown of name characteristics. Notably, all names are distinctly white. Sandra and James (Misty and Sammie) are well-matched on measures of competence, hard work, and professionalism.

Study 4

H6. On average, male applicants (James/Sammie) will be awarded less than female applicants (Sandra/Misty).

H7. On average, high-competence (Sandra/James) names will be awarded more than low-competence names (Sammie/Misty).

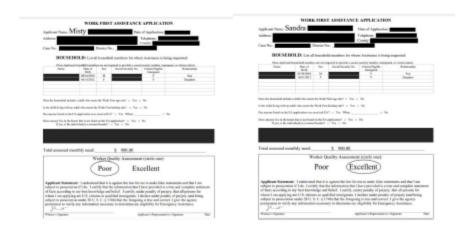
H8a. For female applicants (Sandra/Misty), those rated as "Excellent" workers will be awarded, on average, more than those rated "Poor" workers. For example, "Excellent" Sandra is expected to be awarded more than "Poor" Sandra.

H8b. For male applicants (James/Sammie), there will be no significant difference between amounts awarded to "Excellent" workers as compared to "Poor" workers. For example, "Excellent" James is expected to not earn significantly more or less than "Poor" James.

Randomization Process (Study 4)

Respondents are asked to allocate \$1,500 between two applicants with expressed need of \$900 each or to the state to "offset the federal budget." The exact applicant images that respondents will see are attached in the protocol. The instructions provided to YouGov programmers can be found in the survey protocol attached.

All respondents will see the same Applicant 1, Sandra, who is rated as an "Excellent" worker. There is a 1/7 probability of being assigned each of the possible options for Applicant 2: Excellent James, Poor James, Excellent Misty, Poor Misty, Excellent Sammie, Poor Sammie, Excellent No Name, and Poor No Name. An example of how the applications will appear to respondents is below.



Study 4

Outcome 1: Amount Awarded to Applicant 1

Outcome 2: Amount Awarded to Applicant 2

Outcome 3: Amount Allocated to the State

Study 4

To evaluate H6-8b, I will use difference-in-means testing.

As a first step in evaluating these hypotheses, I will use the average amount awarded to Sandra as the baseline and compare the average amounts awarded to Misty, James, Sammie, and the state. I will use difference in means testing to determine whether the average amounts awarded to Misty, James, and Sammie are statistically significantly different from Sandra.

I will also aggregate the names into Male and Female groups, take the average amount awarded to men (James/Sammie) and the average amount awarded to women (Sandra/Misty). If H6 holds, male applicants should receive significantly less than female applicants.

To evaluate H7, I will group the applicants into high-competence (Sandra/James) and low-competence (Sammie/Misty) groups, take the average amount awarded to each group and perform a difference in means test.

To evaluate H8a, I will compare the average amount awarded to Misty (excellent) with the average amount awarded to Misty (poor) relative to the baseline, Sandra (excellent). If H8a holds, Misty (excellent) should earn significantly more than Misty (poor).

To evaluate H8b, I will compare the average amount awarded to James (excellent) and Sammie (excellent) to James (poor) and Sammie (poor) relative to the baseline, Sandra (excellent). If H8b holds, James (excellent) should not earn significantly more than James (poor) and Sammie (excellent) should not earn significantly more than Sammie (poor).

Appendix C - Discussion of Ethics

This research was reviewed and deemed exempt from review by the [redacted] Institutional Review Board. The experiment was programmed and conducted by YouGov, a large survey firm in the United States. Adult subjects (> 18 years old) were drawn from a database of survey participants maintained and recruited by YouGov. There are no physical, psychological, social, or legal risks beyond the most minimal risks associated with everyday activity. In terms of economic or financial risks, subjects are paid and recruited by YouGov. Subjects do not pay to participate in the recruitment pool and no aspect of the survey or experiments therein require subjects to wager, bet, or contribute financially in any way. Thus, there are no reasonably foreseeable risks related to the subjects' participation in the experiment. No identifying information was shared to the researcher by YouGov. De-identified data were stored electronically on the researchers' server. There was no anticipated expiration of the data storage period. The de-identified data are available for purposes of replication once the research is published.