

Who Benefits? {Catchy Subtitle Coming Soon}

Sarah “Dot” Warren^{1, 2}

Abstract: ... Given the competing intuitions behind rewarding or punishing men seeking welfare, an experimental test of each hypothesis is presented.

Introduction

Is the value of hard work conditional on one’s gender? We have long known that women experience discrimination in ways that men do not: Women receive different treatment in everything from applying to jobs (Quadlin 2018) to running for office (Hassell and Visalvanich 2019; Chaudoin and Hummel 2020) to the price they pay for basic household goods (Betz, Fortuano, and O’Brien 2021). Economically, women earn less than their male counterparts (Mandel 2013) and comprise the overwhelming majority of those on welfare (Fraser 1987; Smith, Appio, and Cho 2012). Despite this, and the well-documented reality that welfare attitudes are racialized (Hayes and Mitchell 2022; Desante 2013; Winter 2006; Gilens 1999), the *feminization* of welfare attitudes has received limited attention from scholars. Emblematic of this is the fact that much experimental work on the racialization of welfare attitudes use exclusively female names or images when cuing race (Gilliam 1999; Winter 2006; Desante 2013).

Still, there remains uncertainty about whether norms of male favoritism should carry into welfare allotment. The vast literature documenting the more favorable treatment of men relative to women primarily concerns traditionally masculine domains: salary negotiations (Castillo et al. 2013), hiring decisions (Neumark 1996; Goldin and Rouse 2000), and elections (Clayton, Robinson, Muriaas 2019; but see Fraga and Hassell 2021). Welfare is distinct from these domains in that it carries with it social stigma (Soss 1999) and negative stereotypes (Foster 2008; Esping-Andersen 2015; Schofield, Hasalm, and Butterworth 2019). In such a domain—where awarding the “prize” may result in increasing the recipient’s social

¹ Ph.D. Student, University of Rochester, Political Science Department

² This work began while the author was a graduate student in the Department of Political Science at Florida State University. Data collection was generously funded by the Institute of Politics at FSU.

costs³—favorable treatment could reasonably take the traditional form of rewarding men more aid relative to women or it may take the form of subjecting men to less stigma by denying them aid if aid is perceived negatively.

The central purpose of this article is to enter the debate about modern sexism, American values, and bureaucratic responsiveness to show that, when men and women are put in direct competition for scarce resources, gender-based prejudice interacts with American values of hard work to amplify existing, harmful stereotypes about men and women. I find that, on average, male applicants earn less than female applicants with identical needs. Further, I find the amount awarded to women varies conditional on being rated a “Poor” or “Excellent” worker, while there is no difference in the amount awarded to men who are rated “Poor” workers compared to men who are rated “Excellent” workers. Thus, women’s advantage over men applicants extends only so far as women are “deserving,” where deservingness comes in the form of external verification.

The totality of these results suggest that people are more inclined to help poor women rather than poor men, but only *deserving* women, when the quality of women applicants is validated by an external source. This research provides evidence toward benevolent sexism, in which women are viewed as weak and in need of extra care (Glick and Fiske 1996). It additionally informs our understanding of gender differences in considerations of the deserving poor and the consequences of these considerations on bureaucratic responsiveness.

Theory and Hypotheses

Sexism is rooted in the belief that one gender is superior to the other. Typically, sexism is thought of as hostility toward women, perpetrated by men; however, a wealth of evidence suggests that Americans’ views of women are deeply ambivalent (Glick and Fiske 1996). Specifically, sexism toward women may take the form of hostile or benevolent sexism; while

³ There is ample evidence that social stigma does not decrease aid applications (cites from Vlad paper); however,

both forms of sexism share the assumption that women are inferior and restrict women to a lower social status, they manifest in different behavior patterns. Hostile sexism constitutes punitive behavior toward women who deviate from their prescribed role as heterosexual domestic laborers, whereas benevolent sexism manifests in protective coddling of women who stay within this role. Benevolent sexism reflects evaluations of women that are seemingly positive, but remain caustic to gender equity and restrict women's personal, professional, political, and social opportunities. Examples of benevolently sexist attitudes include the reverence of women in wife, mother, and child caretaker roles, the romanticizing of women as objects of heterosexual affection, and the belief that men have a duty to protect women (Chen and Farhart 2020; de Geus, Ralph-Morrow, and Shorrocks 2022; di Carlo 2020). Failing to remain within the confines of these traditional roles may trigger punishment in the form of hostile sexism. Examples of hostile sexism include beliefs about women as incompetent, unintelligent, overly emotional, and sexually manipulative (McThomas and Tesler 2016; Luks and Schaffner 2019; Cassese and Holman 2019).

Scholarly intuition is mixed, however, regarding how gender attitudes should inform evaluations of welfare recipients. Pateman (1987) argued that women are not perceived as real economic agents and are therefore considered less deserving of aid than their male counterparts. However, it is equally plausible that women in need of government assistance may trigger *benevolent attitudes* of protection. Women consistently regard welfare more positively and as a higher policy priority, in part because women are the primary beneficiaries of welfare spending (). This has additional implications for men applying for aid. While we may expect a woman applying for aid to benefit from benevolent sexism, there is reason to suspect men applying for aid may trigger hostile gendered attitudes. Americans have long believed people “ought to take care of their personal problems by themselves” without relying on the government for aid (Sniderman and Brody 1977, 501), this “pull yourself up by your bootstraps” mentality appears to be concentrated among men.

The feminization of welfare, combined with the perceived failure of men seeking public assistance to pull themselves up by their bootstraps and be breadwinners, may contribute to a distaste for men receiving public benefits: such policies help those who should be helping themselves (Bobocel et al. 1998; Katz and Hass 1988; Sniderman and Tetlock 1986; Sniderman et al. 1996). Building on this body of work, I hypothesize that male aid applicants will be awarded less than female applicants.

H1. On average, male applicants will be awarded less than female applicants.

A second dimension that has characterized past work and public debate on welfare allotments is *deservingness*. The “deserving poor” are those whose impoverishment is due to personal failings or personal qualities of dependence, unlike the “undeserving poor,” but rather those whose personal financial circumstances have been devastated by structural or macroeconomic forces beyond the individual’s control. Thus, conditional on gender, an individual’s *deservingness* should also affect the amount of aid they are awarded. Because notions of deservingness and undeservingness are frequently characterized by ability and willingness to work, we might think that factors like individual competence will affect aid allocations. Specifically, high competence workers should receive more than their low-competence counterparts.

H2. On average, high-competence applicants will be awarded more than low-competence applicants.

Deservingness might also be cued by an external evaluation of worker-quality. If the public truly prioritizes giving aid to good workers whose personal financial circumstances were devastated by forces outside of their control, then high-quality workers should be awarded more than low-quality workers. However, evidence suggests that external or objective “proof” of women’s capabilities may be valuable resources in offsetting negative perceptions of women as workers (Abel, Burger, and Piranio 2020; Dadgar and Trimble 2015; Jepsen, Troske, and Coomes 2014), but offer lower returns for men. I therefore hypothesize:

H3a. For male applicants, there will be no significant difference between amounts awarded to “Excellent” workers as compared to “Poor” workers.

H3b. For female applicants, there will be a significant difference between amounts awarded to “Excellent” workers as compared to “Poor” workers.

Research Design and Data

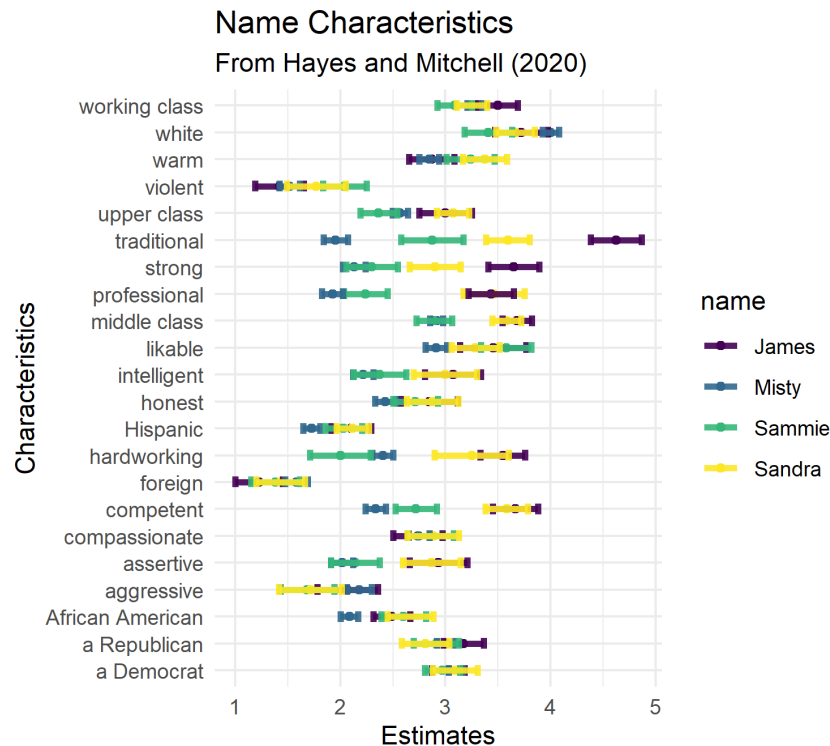
In order to test how sexism, gender, and objective and perceived work ethic shape Americans’ attitudes toward welfare, I conducted a survey experiment in which participants were asked to budget money to different pairs of applicants for state assistance. Following DeSante (2013), I use hand-redacted welfare applications to manipulate targets’ for assistance sex, perceived competence, and objective work-quality rating (Fig. []). To manipulate the objective quality rating, each aid target is randomly assigned a rating of “Excellent” or “Poor” on their aid application. To manipulate sex and perceived competence, I use two male and two female names from Hayes and Mitchell’s (2020) name-characteristics dataset: Sandra, James, Misty, and Sammie. These names were specifically selected to minimize the likelihood that factors other than sex, objective, and perceived work ethic would affect treatment.

Names are used rather than more overt cues to minimize demand effects (De Quidt, Vesterlund, and Wilson 2019).⁴ To mitigate concerns about the effects of race and the racialization of welfare confounding results, all four names chosen were coded as racially distinct white names in the Hayes and Mitchell (2020) names dataset.⁵ Thirdly and finally, these names are matched on characteristics that Americans have long reported are relevant considerations when considering welfare support (Sinderman and Brody 1977; Bobocel et al.

⁴ Experimenter demand effects refer to changes in behavior that result from study participants wanting to help the experimenter confirm her underlying hypothesis.

⁵ To create the dataset, the authors obtained a list of given names among people born in the United States between the years 1955 and 1990 from the U.S. Social Security Administration (SSA). To identify the gender of the names, they used information on sex included in the SSA data.

1998; Katz and Hass 1988; Sniderman and Tetlock 1986; Sniderman et al. 1996). Sandra and James are rated highly in professionalism (), competence (), and work ethic (), while Sammie and Misty are rated lower in all three characteristics (). All names are comparable in estimates of whiteness, honesty, and likeability. Figure [] below shows the complete breakdown of name-characteristics.



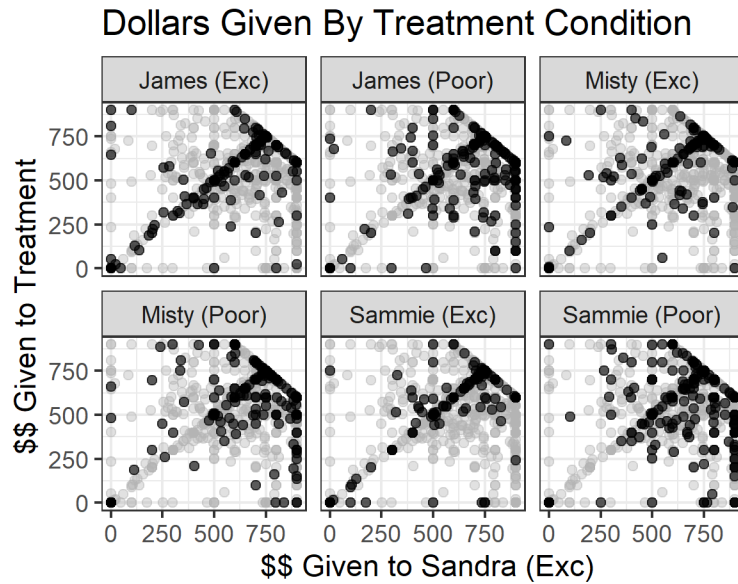
Subjects are given a budgeting task in which they are asked to allocate \$1,500 to two applicants for federal assistance, each of whom has a state-determined need of \$900. Respondents may also choose to give some (or all) of the funds to “offset the state deficit.” Given the budget constraint—both applicants’ full need cannot be met—I use the amount awarded to Applicant 1, Applicant 2, and the Government as an estimate of an applicant’s “deservingness.” These allocations are my main variables of interest. Everything about the applicants remains identical, except for a *worker quality assessment* of Excellent or Poor and their name, which cues both *sex* (male or female) and *competence* (high or low). Given this, if an applicant receives a different allocation across treatments, we can infer that this difference is

due to experimental manipulation. As the applicant's characteristics were manipulated via random assignment, any difference in the relative importance respondents place on fiscal responsibility—illustrated by giving more to offset the budget deficit—can also be traced back to the experimental treatment. This deficit option also allows for individuals to take a principled position, a socially desirable and available option, to decide that the money would be better spent in some other way.

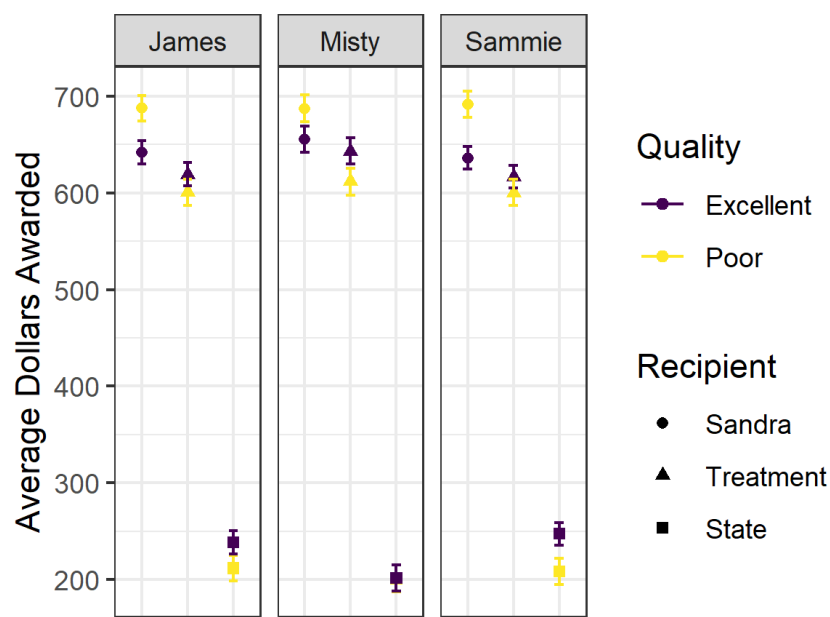
In order to isolate the effects of sex versus the traits people ascribe to different names, I fielded a survey experiment with YouGov (n=2150)⁶ in April 2022 modeled after DeSante (2013). Respondents viewed two applications identical in appearance to the original experiment. Rather than randomizing both applications, all respondents viewed the same baseline application of “Sandra” who was rated as “Excellent” compared to a second application. I used a 2x2x2 factorial design for this second application, randomizing sex (male/female), competence (high/low), and quality assessment (excellent/poor) of the second application, using the names James, Misty, and Sammie as my cue for sex and competence. Respondents were then asked to allocate funding to the two applicants or to offset the state budgetary deficit.

Experimental Results

⁶ Preregistration link: <https://osf.io/u5wxg/>.



When men and women are put in direct competition for scarce resources, how do women fare compared to their male counterparts? How does perceived competence and external quality ratings effect this relationship? Figure [] presents the main results of the experiment by treatment name, with color denoting the quality rating (Excellent/Poor) of the treatment name, and point-shape denoting the recipient. Recall that the baseline condition is an applicant named Sandra, a high-competence name, who is rated as an “Excellent” quality worker.



H1 predicts that, on average, male applicants will be awarded less than female applicants. The first panel of Figure [] demonstrates that, when she is paired with a high-competence, “Excellent” male name, James, Sandra is awarded significantly more on average (\$642.06 vs. \$619.20, $p = 0.014$). Excellent Sandra also earns more on average than Sammie, a low-competence name, when he is rated “Excellent” (\$635.98 vs \$616.60, $p=0.000$). However, there is also no significant difference between Excellent Sandra and Excellent Misty (\$655.31 vs \$643.20, $p=0.141$). These results are supportive of H1, in that Excellent Misty earns no less than Excellent Sandra, while Excellent James and Excellent Sammie earn substantively and significantly less than the baseline.

Turning to H2, I consider the difference between name competencies and aid amounts awarded. Hayes and Mitchel’s findings suggest that name-characteristics have the potential to minimize differences in aid allotment due to racial prejudice; however, the mitigating effects of name-characteristics do not appear to extend to gender bias. Recall that competence is determined by valence-characteristic scores in the Hayes and Mitchell (2021) names dataset (Fig. []). Sandra and James are both high-competence names, while Misty and Sammie are

relatively low-competence. H2 predicts that high competence names should receive more, on average, than low competence names. Thus, holding applicant quality constant, Sandra should receive more than Misty. James should receive more than Sammie, relative to the baseline. I do not find empirical support for H2. There is no substantively significant difference between Excellent Misty and Excellent Sandra (\$655.31 vs \$643.20, $p=0.141$), nor is there between Excellent James and Excellent Sammie (\$619.20 and \$616.60, $p=0.88$), and Poor James and Poor Sammie (\$600.60 and \$600.32, $p=0.986$). This suggests that, among whites, competence matters less than when making inter-racial comparisons.

Finally, H3a holds that, for male applicants, there will be no significant difference between amounts awarded to workers rated “Excellent” as compared to workers rated “Poor.” H3b holds that the opposite will be true for female applicants; that is, worker quality rating will result in a significant difference in the amount awarded. To evaluate H3, I compare the difference in means for each treatment name to (James, Sammie, and Misty) when they are “Excellent” rated workers to when they are rated “Poor.” Table [] shows the results of this test, which support H3. There is a statistically significant difference in the amount of aid awarded to Excellent Misty compared to Poor Misty; however, there is no such difference between Excellent Sammie or James compared to Poor Sammie or James.

| Treatment Name | Excellent | Poor | Difference | p-value |
|----------------|-----------|----------|------------|---------|
| Misty | \$643.20 | \$611.80 | \$31.40 | 0.044* |
| James | \$619.20 | \$600.60 | \$18.40 | 0.256 |
| Sammie | \$616.60 | \$600.32 | \$16.28 | 0.351 |

WORKING HEREAFTER

Robustness

Compare estimates to the no-name treatment (i.e. compare Sandra vs Name-X to Sandra vs No Name - difference is significant ($p=0.04$, $n = 310$, but the mean of the differences is \$15.51, which is significantly less in magnitude than other significant results.)

Empirical Relevance

Bureaucratic responsiveness – if such a light touch prompts substantively significant treatment effects, how might *increased* interaction (e.g. phone or in-person interviews) amplify these gendered differences in treatment?

Example: *Weinberger v. Wiesenfeld*?

Concluding Discussion