Appendix A - Robustness

This appendix presents a series of robustness checks for the main results reported in the main body of the paper.

Subsample Analyses

The main body of the paper reports results from a pooled sample made up of a nationally representative sample and a low-socioeconomic status oversample in order to preserve statistical power. This appendix reports the main results of the paper for each subsample.
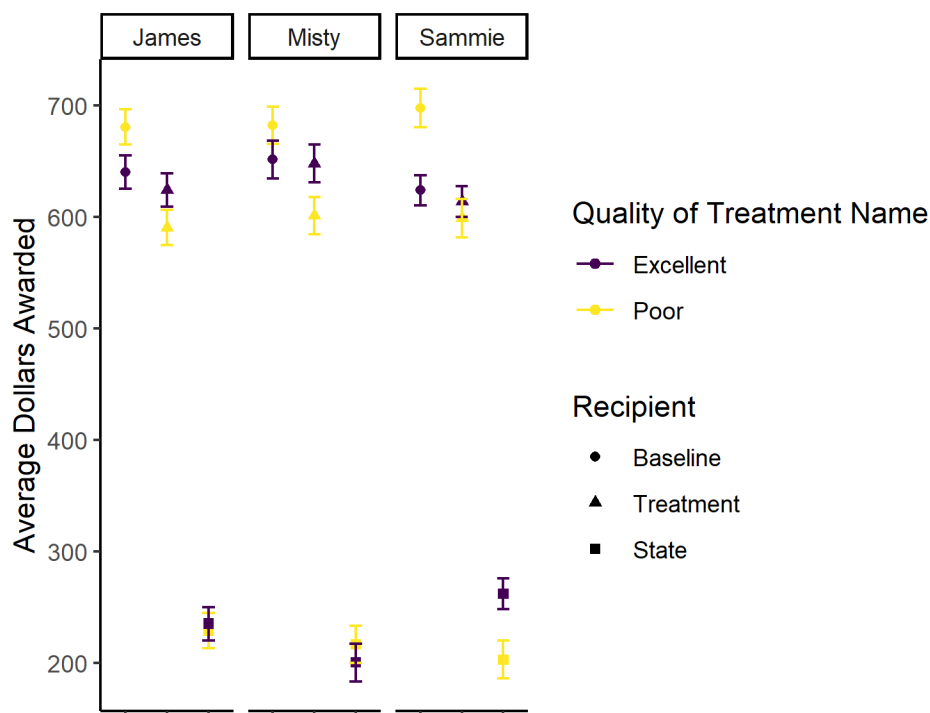

Figure 1: Results Restricted to Nationally Representative Subsample

Restricting the sample to the nationally representative subsample, I find that male names are still awarded less relative to the baseline than female names, but this difference is not statistically significant at conventional levels.

Table 1: Testing H1 with Nationally Representative Sample

Condition	Mean Difference	Upper CI	Lower CI	p-value
Excellent Misty vs Baseline	3.537	-11.467	18.542	0.643
Excellent James vs Baseline	16.065	-8.571	40.7	0.2
Excellent Sammie vs Baseline	10.041	-7.728	27.81	0.266

I find that Excellent Misty is awarded more than Poor Misty and that these differences are statistically significant at conventional levels, whereas the difference in means between Poor and Excellent James and Poor and Excellent Sammie are not statistically significant at conventional levels.

Table 2

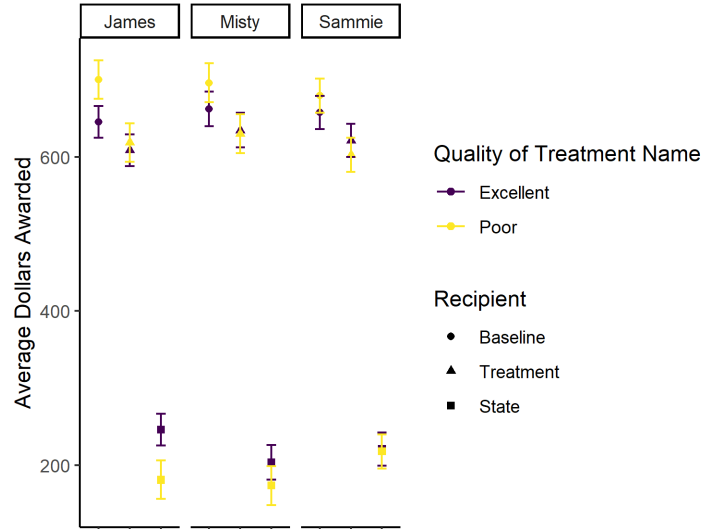| Condition	Mean 1	Mean 2	Upper CI	Lower CI	p-value
Excellent Misty vs Poor Misty	648.09	601.396	7.527	85.86	0.02**
Excellent James vs Poor James	624.393	590.662	-7.489	74.952	0.108
Excellent Sammie vs Poor Sammie	614.041	599.09	-28.278	58.181	0.497

Figure 2: Results Restricted to Low SES Subsample

Restricting the analysis to the low SES subsample, I find that many of my results on the full sample hold. The amount awarded to Excellent Misty is statistically indistinguishable

from the amount awarded to the baseline. Excellent James earns significantly less than the baseline, as does Excellent Sammie.

Table 3: Testing H1 with Low SES Subsample

Condition	Mean Difference	Upper CI	Lower CI	p-value
Excellent Misty vs Baseline	27.491	-9.06	64.042	0.139
Excellent James vs Baseline	36.948	12.719	61.178	0.003***
Excellent Sammie vs Baseline	36.318	8.942	63.694	0.01**

However, Excellent Misty earns only $4 more on average than Poor Misty and this difference is not statistically significant. Excellent James earns *less* than Poor James on average, but this difference is not statistically significant either. Excellent Sammie earns more than Poor Sammie, but this difference is also not statistically significant.

Table 4: Testing H3a-b with Low SES Subsample

Condition	Mean 1	Mean 2	Upper CI	Lower CI	p-value
Excellent Misty vs Poor Misty	634.42	630.119	-45.183	53.783	0.864
Excellent James vs Poor James	608.433	618.682	-61.091	40.592	0.691
Excellent Sammie vs Poor Sammie	621.234	602.602	-38.23	75.494	0.519

These results suggest one of two things. First, it is possible that the treatments effect the two subsamples differently. Perhaps the general population cares more about applicant quality and less about gender, whereas low income populations bring stronger gender stereotypes into deservingness evaluations. However, another consideration is that these differences in results are an artifact of randomization because I randomized treatments over the whole sample, not over the two subsamples individually.

Appendix C - 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 redacted 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 3: Sliding Scale Used by Respondents to Make Allocations