APPENDIX A: ONLINE SUPPLEMENTARY FIGURES AND TABLES

FIG 9. *Comparison of the classification accuracy rate and logistic score test statistics on simulated data*

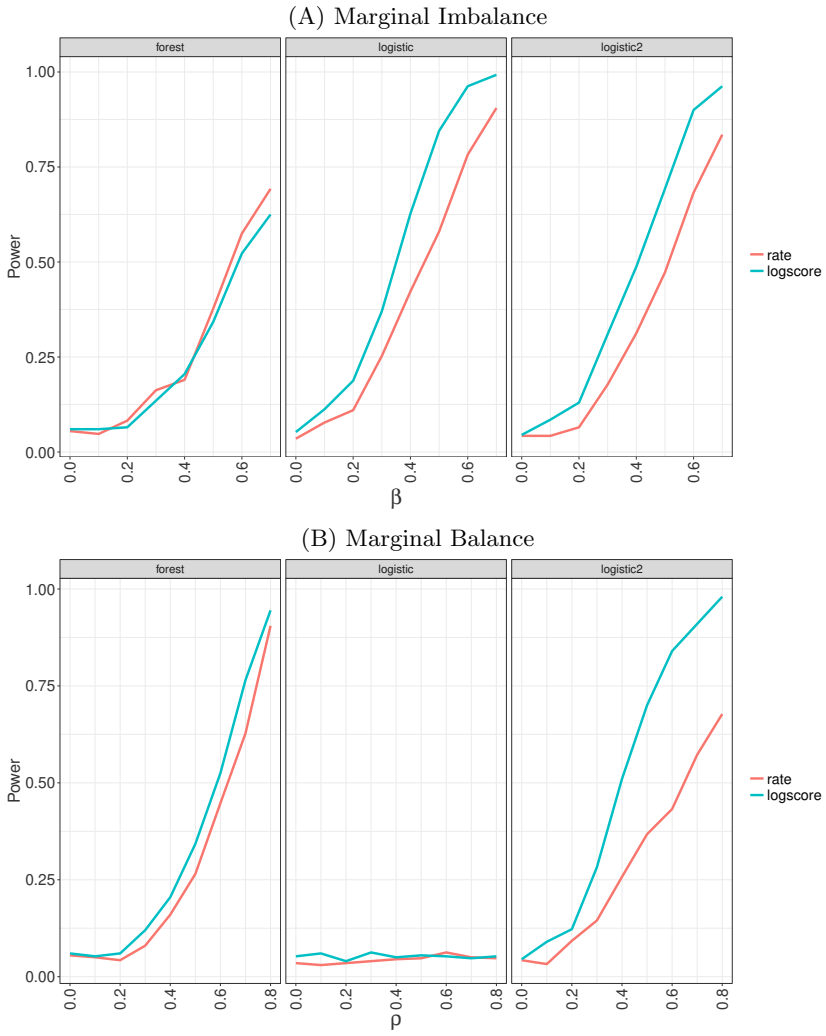


TABLE 4. *Observed defendant characteristics*

```
"Indicator: defendant female"
"Indicator: defendant not black"
"Defendant's age at arrest"

"Indicator: defendant arrested prior to arrest in sampled case"
"Indicator: defendant arrested on felony charge prior to arrest in sampled case"
"Indicator: defendant arrested on drug charge prior to arrest in sample case"
"Indicator: defendant arrested on felony drug charge prior to arrest in sampled"
"Indicator: defendant convicted prior to arrest in sampled case"
"Indicator: defendant convicted on felony charge prior to arrest in sampled case"
"Indicator: defendant convicted on drug charge prior to arrest in sampled case"
"Indicator: defendant convicted on felony drug charge prior to arrest in sampled"

"Indicator: marijuana was drug involved in crime"
"Indicator: powder cocaine was drug involved in crime"
"Indicator: crack cocaine was drug involved in crime"
"Indicator: heroin was drug involved in crime"
"Indicator: PCP was drug involved in crime"
"Indicator: another drug was involved in crime"
"Indicator: defendant charged with felony possession with intent to distribute"
"Indicator: defendant charged with felony distribution"
```

Note: The "Female" and "PCP" variables are `TRUE` for only a small number of observations. As a result, including these two variables in logistic regression models sometimes leads to convergence issues. To avoid these issues, and for sake of consistency, we drop these two variables from *all* analyses in this paper.

JOHANN GAGNON-BARTSCH
UNIVERSITY OF MICHIGAN
DEPARTMENT OF STATISTICS
E-MAIL: johanngb@umich.edu

YOTAM SHEM-TOV
UNIVERSITY OF CALIFORNIA, BERKELEY
DEPARTMENT OF ECONOMICS
E-MAIL: shemtov@berkeley.edu

FIG 10. *Contour plots of the joint distribution of key covariates in Lyall (2009)*

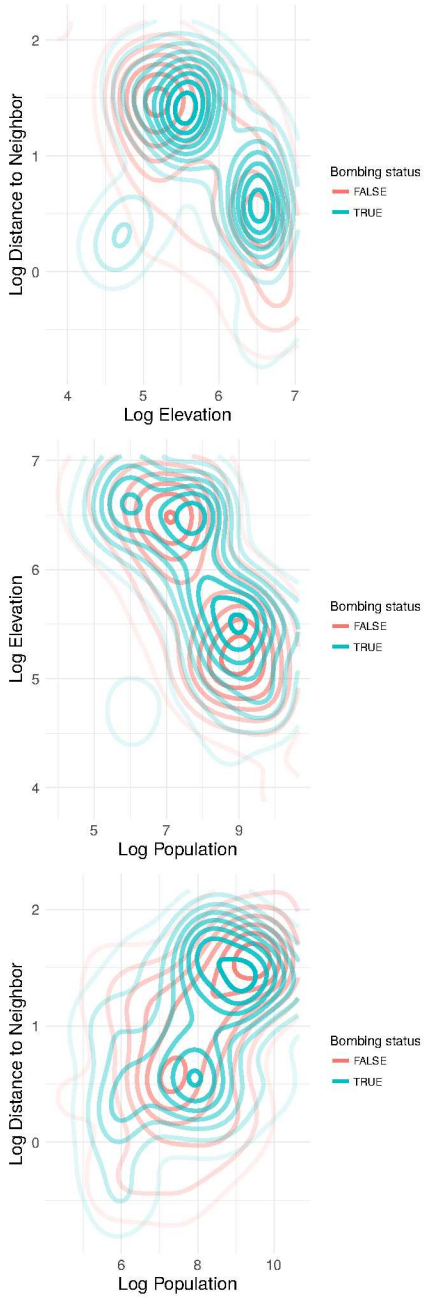
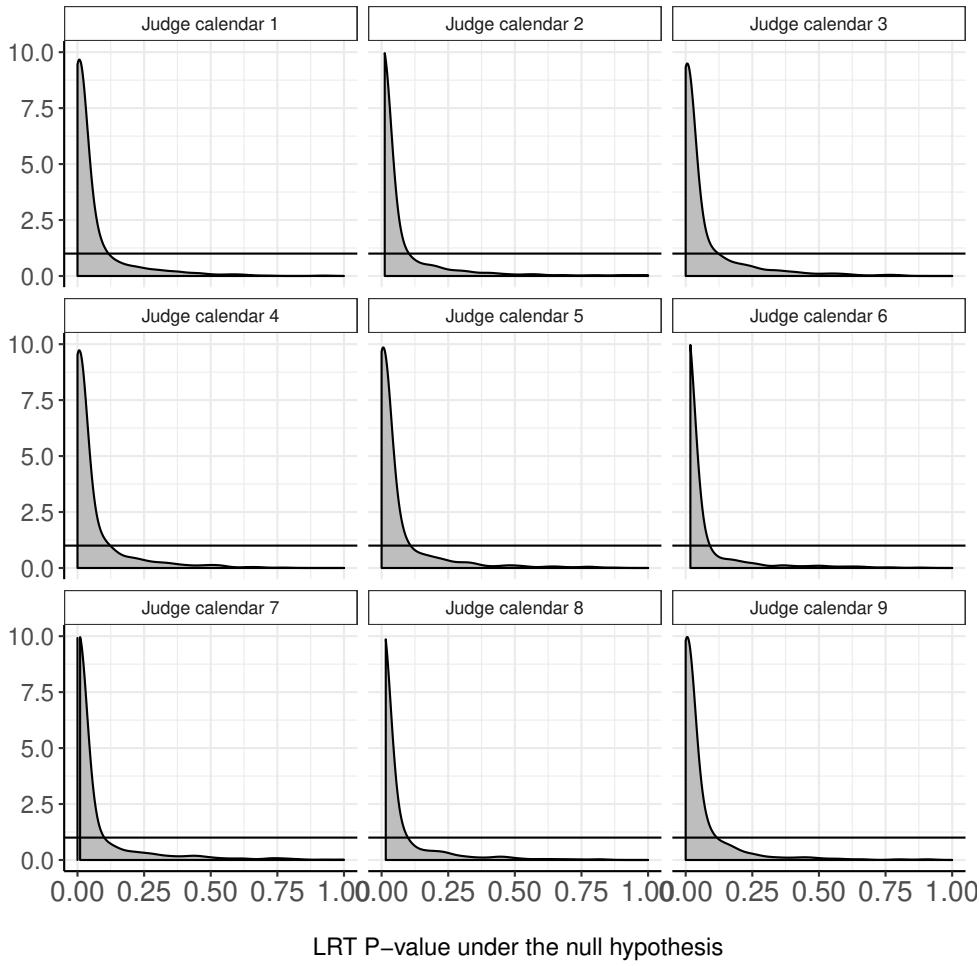
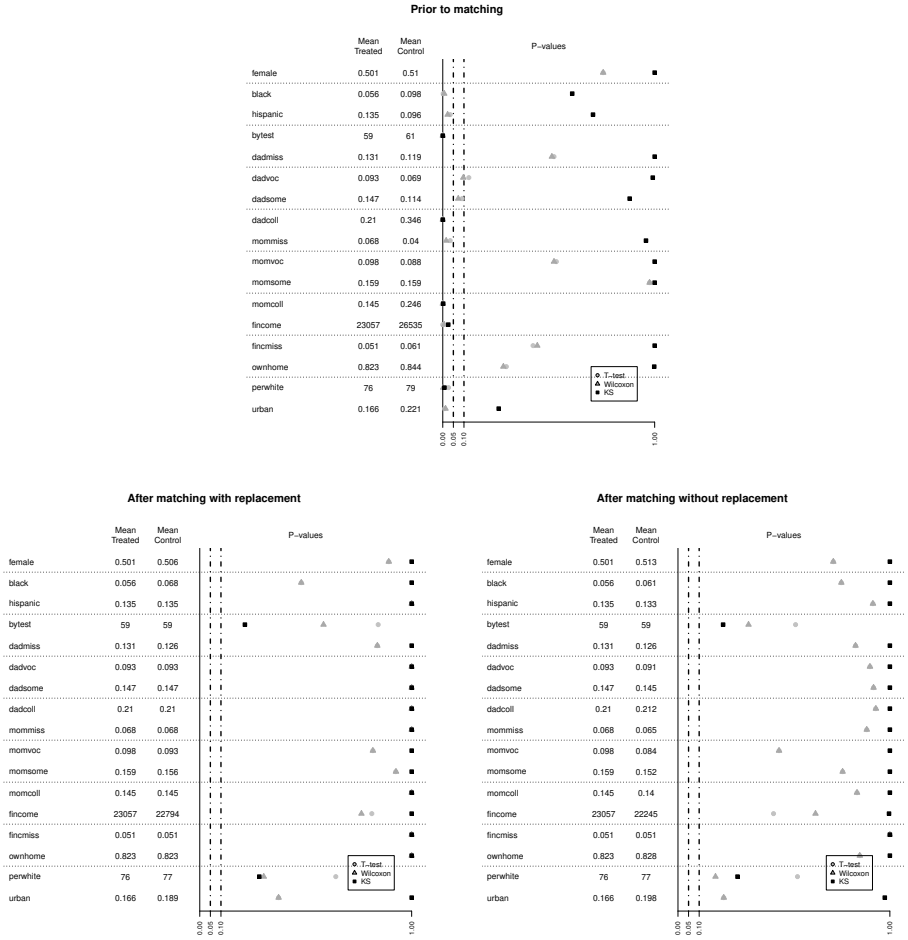


FIG 11. *The distribution of the Likelihood Ratio Test P-value using all two-way interactions*



Notes: Each plot shows the distribution of the LRT P-value under the null hypothesis of random assignment. The horizontal line shows the uniform $[0, 1]$ distribution that is expected under the null.

FIG 12. *Covariate balance before and after matching*



Notes: The figures show the covariate balance in Rouse (1995) data before and after implementing a matching procedure to minimize distances on observable characteristics of students in two-year relative to students in four-year college.