Algorithmic Recommendations and Human Discretion
(with Victoria Angelova and Will Dobbie)

Abstract

Human decision-makers frequently override the recommendations generated by predictive algorithms, but it is unclear whether these discretionary overrides add valuable private information or reintroduce human biases and mistakes. We develop new quasi-experimental tools to measure the impact of human discretion over an algorithm on the accuracy of decisions, even when the outcome of interest is only selectively observed, in the context of bail decisions. We find that 90% of the judges in our setting underperform the algorithm when they make a discretionary override, with most making override decisions that are no better than random. Yet the remaining 10% of judges outperform the algorithm in terms of both accuracy and fairness when they make a discretionary override. We provide suggestive evidence on the behavior underlying these differences in judge performance, showing that the high-performing judges are more likely to use relevant private information and are less likely to overreact to highly salient events compared to the low-performing judges.