On the Use of Outcome Tests for Detecting Bias in Decision Making

Empirical comparisons of decisions and their consequences across affected groups are not enough to detect bias; careful modeling of decision maker behavior is also required to define bias and clarify what data can and cannot reveal about bias in a given setting.

In his thesis-turned-book, *The Economics of Discrimination* (1957), Gary Becker wrote that a biased decision maker "must act as if he were willing to pay something" to exercise bias. In other words, you own a store but are willing to forgo sales to certain types of people at a cost to your bottom line. Or you refuse to hire certain candidates based on demographic characteristics even though they are the most qualified. These are the prices that you are willing to pay to discriminate. Becker's book jump-started research programs on discrimination that continue today, and "willing[ness] to pay" remains a foundation of that research.

However, how can we learn whether the decisions of employers, teachers, judges, landlords, police officers, and other gatekeepers are discriminatory, rather than reflective of other relevant group-level differences? To answer this question, we must first define what it means for a decision to be unbiased, which requires specifying what unbiased decision makers in a particular setting are supposed to be optimizing, what constraints they face, and what they know at the time they make their decisions. We can then derive optimality conditions for the decision-maker's problem and check whether those conditions are consistent with data for different groups affected by the decision. If these checks suggest that an unbiased decision maker *could do better* by changing how they treat members of a particular group, the analyst may conclude that this group is subject to bias. In other words, in such a case we may have discovered a decision maker willing to pay Becker's price of discrimination.

This paper examines what researchers can learn about bias in decision making by comparing post-decision outcomes across different groups. As in many economic inquiries, it is behavior *at the margin* that matters: when a bail judge, for example, is on the fence about whether to release versus detain a defendant before trial, examining the subsequent pre-trial misconduct outcomes of such a marginal defendant, and comparing the outcomes of marginal defendants of different races, may help reveal a decision maker's differential standards.

But how can we ensure that differential outcomes in those marginal cases reveal decision maker bias? To answer that, the authors make a novel connection between testing for bias and imposing various flavors of Roy models, which have long been employed by economists to analyze decision

Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

making. In his 1951 paper, A.D. Roy describes a world where people choose between hunting and fishing as an occupation, and people differ in their skills in each task. The point of the model is not to observe the aggregated choices, that is, how many choose hunting and how many choose fishing, which is merely a matter of empirics; rather, Roy asks whether those who are relatively more skilled at hunting will hunt, and whether those who are relatively more skilled at fishing will fish, which is a more nuanced question that, like testing for bias, depends on the underlying model of behavior assumed to generate the observed data. Roy models have evolved to incorporate more complexity since A.D. Roy's original formulation, including accommodating additional factors that influence decision making but are not observable to the analyst.

In outcome tests of bias, the authors show that such unobservable factors can render marginal outcomes, even if perfectly known, uninformative about decision maker bias in the most general member of the Roy family—the Generalized Roy Model—which is a workhorse in modern applied economics thanks to its empirical flexibility. The authors then show how a more restricted "Extended" Roy Model delivers a valid test of bias based on the outcomes of marginal cases. This highlights a tradeoff between the flexibility of a decision model and its ability to deliver a valid outcome test of decision maker bias. Indeed, imposing the Extended Roy Model yields a valid test of bias precisely because it rules out other behaviors that may be empirically indistinguishable from bias, like bail judges considering job loss,

family disruption, and other consequences of pre-trial detention beyond the typically measured outcome of pre-trial misconduct.

The authors also discuss ways of taking these models to data across a wide range of real-world settings. They highlight a distinction between econometric assumptions that help identify marginal outcomes using variation across different decision makers, versus modeling assumptions that help derive a valid test of bias based on those marginal outcomes; the former do not necessarily imply the latter. Both types of conditions hold in the Extended Roy Model, however, and due to the restrictions it imposes, it has clear testable implications that may help empirical researchers assess its suitability across empirical settings. The authors also extend their results and discussions to more challenging data environments where variation across different decision makers may not be available, and the analyst attempts to compare average, rather than marginal, outcomes across groups.

Bottom line: empirical description of gatekeeper decisions, and the outcomes that result from those decisions, is not sufficient for detecting bias in decision making; rather, learning about such bias requires specifying and justifying a model that is restrictive enough to deliver testable implications of biased behavior, but rich enough to incorporate the essential elements of the optimization problem faced by decision makers in a given empirical setting.

ABOUT OUR SCHOLARS

Magne Mogstad
Gary S. Becker Professor in Economics and the College, the Kenneth C. Griffin Department of Economics
economics.uchicago.edu/directory/magne-mogstad

Jack Mountjoy
Assistant Professor of Economics and Robert H. Topel Faculty Scholar, Chicago Booth
chicagobooth.edu/faculty/directory/m/jack-mountjoy

THE UNIVERSITY OF CHICAGO