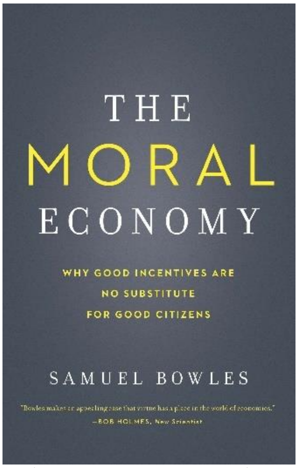
Samuel Bowles. *The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens*. New Haven, CT: Yale University Press, 2016. Pp. xiii-272. ISBN 978-0300163803. Hardcover $27.50.



Depending on one's perspective, Samuel Bowles' *The Moral Economy* is either an unambitious book, or a very ambitious one indeed. On its face, the volume is a kind of how-to manual for policymakers of a technocratic disposition. Its central argument—a resounding, data-driven critique of approaches to public policy that stress purely financial incentives—is thoroughly explicated and engagingly presented. But *The Moral Economy* is also, in its way, a work of political theory. It repeatedly gestures toward far deeper questions about the role of the state in the lives of its citizens, culminating in some startlingly bold theoretical stances.

In the simplest terms, Bowles argues that human patterns of moral reasoning are more than balance sheets. Perhaps the best illustration of this theme is the "case of the Haifa daycare"—an example introduced early on in the book, to which he consistently returns. As the story goes, the daycare adopted a policy imposing fines on parents who were late to collect their children. This new policy was driven by the daycare's frustration with parents who had become increasingly lax about punctuality. The policy, however, had the *opposite* of its intended effect: more parents than before collected their children late, simply opting to pay the fine. It was clear the daycare parents viewed the fine not as a penalty—a punishment for failing to live up to their end of the bargain—but as simply the monetary price of tardiness. And sometimes, depending on the circumstances, that price was worth paying.

As Bowles sees it, this reflects a fundamental problem with the way governments often try to promote certain behaviors. Financial incentives, stemming from a view of human beings as *homo economicus*—fundamentally self-interested and money-motivated—often fail to secure desired outcomes. To support this claim, Bowles provides what amounts to a meta-analysis of behavioral science research in the area, summarizing and evaluating their findings. (This discussion constitutes much of the body of the book.) At bottom, this research reduces down to a straightforward insight: individuals are motivated by more than money, such that money is not a motivational *lingua franca* capable of promoting any behavior one might imagine.

In that spirit, most of *The Moral Economy* makes largely uncontroversial claims. Here it bears mention that it is not clear anyone actually subscribes—or *ever* has subscribed—to the stripped-down, ultra-reductionistic account of incentives that Bowles criticizes. The economic literature is replete with discussions of how nonmonetary values drive individual behavior. Yet Bowles obviously conceives of his book as an critique of *something*—perhaps the *homo economicus* assumptions he believes underpin too much public policy.

And it is here that *The Moral Economy* becomes far more intellectually interesting: it relies heavily on a foundational, yet virtually unexamined, premise that the contemporary state must be the start and end of meaningful analysis in this domain. That assumption trenches on one of the most important disputes in twentieth-century political theory: the clash between the celebrated John Rawls and Robert Nozick.

Rawls began his political analysis from the "top down." His famous "veil of ignorance" thought experiment—the zero point of his political philosophy—*presumed* a certain concept of sovereignty, and worked backwards from that to build out a theory of distributive justice. By contrast, Nozick stressed the importance of thinking politically "from the ground up," conceiving of political theory as an account of development from individuals to tribes to organized confederations. Any other paradigm, Nozick warned, would find itself unmoored from empirical reality, leading to theories unreflective of how human beings actually behave. (The title of his most prominent work—*Anarchy, State, and Utopia*—reflects the conceptual progression at the heart of his study: out of anarchy, a society emerges, which then moves in turn towards optimization.) To its core, *The Moral Economy* embraces Rawls' framing—and it regularly finds itself constrained by that framing's default assumptions.

In particular, Bowles follows Rawls in assuming an analytical "zero point" that lacks an obvious real-world correlate. Where (in his telling) too many policymakers begin by assuming that human beings are reducible to *homo economicus*, Bowles substitutes his own vision of *homo plasticus*, where humans are essentially moral blank slates awaiting tutelage by the state. This is, at the very least, a fraught position.

To begin with, Bowles provides a fairly thin account of moral formation itself. In stressing the priority of public policy, *The Moral Economy* largely ignores the centrality of mediating institutions—homes, schools, houses of worship, and so forth—in cultivating habits of virtue. This omission is a large one: governmental appeals to citizens' deepest principles must necessarily assume a shared underlying moral grammar, about which the Rawlsian "veil of ignorance" is remarkably silent. And on a still deeper level, the book has next to nothing to say about the anthropology of moral reasoning—that is, how individuals' deepest commitments emerge and cohere in the first place. To name but one example, Jonathan Haidt's Moral Foundations Theory, which would seem to have important implications for Bowles' project, is nowhere discussed.

These types of questions are far more cognizable on a Nozickian theoretical account. By beginning with the choices and values of individuals, such a model helps tease out the philosophical presuppositions and compromises underlying the modern liberal state. From that perspective, one can more accurately assess a given regime's success or failure in light of its underlying principles—principles that reflect the preferences of its citizenry. Historically speaking, those preferences shape the character of the state, not the other way around.

To be sure, Bowles is not unaware of this "problem of preferences." In perhaps the book's most original and distinctive discussion, Bowles unpacks the problem of the "legislator's trilemma": the inability of any liberal economy to (1) attain Pareto efficiency, while simultaneously (2) allowing economic participation to be voluntary and (3) maintaining neutrality with regard to the preferences of individuals. If the state adopts principles (2) and (3) and does nothing, voluntary market participants will not engage in maximally efficient trade and exchange. If the state adopts principles (1) and (3) and intervenes to promote efficiency, but without addressing individuals' preferences, those individuals' market participation must be compelled (the 2018 volume *Radical Markets*, by Eric Posner and Glenn Weyl, advocated this approach). If the state adopts principles (1) and (2) and intervenes to promote efficiency, but without forcing all individuals to participate in the market, its efficiency-promoting measures must take the form of social policies calculated to shape individuals' preferences.

To resolve this trilemma, Bowles is willing to jettison preference neutrality. He favors, that is, concrete actions by the state to form citizens' moral sensibilities. But this leads to some provocative consequences—none of which, unfortunately, the book chooses to engage. For one thing, it would seem that where individuals' profoundest commitments—those values that are in essence, *theological*—stand in tension with the efficiency goals of Bowles' hypothetical state, that state cannot stand idly by. Instead, it must catechize its recalcitrant citizens to "voluntarily" rethink their stances. Under such a regime, freedom of conscience may remain a formal guarantee, but the state may freely identify and stigmatize beliefs that are deemed undesirable.

It is difficult not to see echoes here of the "illiberal liberalism" probed by conservative thinkers like Patrick Deneen, Adrian Vermeule, and others. And in practical terms, a mushrooming role for the state risks eroding the local and familial settings within which moral education occurs—those contexts through which the actions of a virtue-promoting government become intelligible. In short, it seems that if extended beyond a very narrow context, Bowles' theory risks slipping into a self-destructive snare of its own devising.

But perhaps this bleak reading takes Bowles' position too far. As a fairly narrow account of policymaking optimization within a Rawlsian frame, *The Moral Economy* does provide useful insights. Bureaucrats seeking to promote some particular end or another should probably think beyond their citizens' pocketbooks. And so understood, the book's argument succeeds.

As a robust argument against *homo economicus*, however, it does not. It is no critique of markets, or of market-based policies, to assert that human preferences are complex and multifaceted, and that in many cases a nonfinancial incentive may prevail over a financial one. A concept of *homo economicus* that moves beyond caricature allows for this understanding.

The preference hierarchies of individuals are undoubtedly complex, unstable, and ever-shifting things. The real question posed by *The Moral Economy* is not about how and whether those preferences drive individual behavior, but whether the state should seek to modify them. That latter may carry a rather high cost.

John Ehrett
Washington, D. C.