Journal of Economic Literature 2019, 57(1), 147–160
https://doi.org/10.1257/jel.20171463

The Devil Is in the Details: Implications of Samuel Bowles's *The Moral Economy* for Economics and Policy Research†

Rachel Kranton*

All economists should buy and read The Moral Economy *by Samuel Bowles. The book challenges basic premises of economic theory and questions policies based on monetary incentives. Incentives not only crowd out intrinsic motivations, they erode the ethical and moral codes necessary for the workings of markets. Bowles boldly suggests that successful policies must combine incentives and moral messages, exploiting complementarities between the two. This essay argues that to achieve this objective, economists must study the local institutions and social context and engage untraditional data to uncover the interplay of incentives and identity.* (*JEL* A11, A13, D04, D83, E60, Z13)

1. *Introduction*

The Moral Economy, by Samuel Bowles, should be required reading by all graduate students in economics. Indeed, all economists should buy a copy and read it. The book is a stunning, critical discussion of the interplay between economic incentives and preferences. It challenges basic premises of economic theory and questions policy recommendations based on these theories. The book proposes a path forward: designing policy that combines incentives and moral appeals. Therefore, like such a book should, *The Moral Economy* leaves us with much work to do.

The Moral Economy concerns individual choices and economic policy, particularly microeconomic policies with goals to enhance the collective good. The book takes aim at laws, policies, and business practices that are based on the classic *Homo economicus* model of individual choice. The book first argues in great detail that policies that follow from the *Homo economicus* paradigm can backfire. While most economists would now recognize that people are not purely selfish and self-interested, *The Moral Economy* goes one step further. Incentives can amplify the selfishness of individuals. People might act in more self-interested ways in a system based on incentives and

* Duke University.
† Go to https://doi.org/10.1257/jel.20171463 to visit the article page and view author disclosure statement.

rewards than they would in the absence of such inducements.

The Moral Economy warns economists to be especially wary of incentives because social norms, like trust and honesty, are critical to economic activity. The danger is not only of incentives backfiring in a single instance; monetary incentives can generally erode ethical and moral codes and social motivations people can have toward each other. Rather than rely on models populated by *Homo economicus*, policy makers should therefore consider people as they really are, with possibly intrinsic desires to uphold social norms and civic-mindedness. As Bowles summarizes in the introduction, "incentives cannot alone provide the foundations of good governance" (p. 2). Rather, incentives and social motivations interact and must be considered together. Indeed, *The Moral Economy* argues in its concluding chapters that policy makers attentive to these interactions can render incentives and social mores as complements.

In constructing its arguments, *The Moral Economy* provides at least three public goods for economists themselves. First, the book recounts and frames the pitfalls of economic policy from the point of view of political philosophy, providing a masterful summary on which the rest of us can rely. Looking to Aristotle, Bowles introduces a character who will become central to *The Moral Economy* narrative: a legislator who desires that people take decisions for the collective good. In Aristotle's writings, the legislator accomplishes this goal by making citizens good by "inculcating habits in them." But political philosophy moved on from Aristotle. Niccolo Machiavelli and David Hume present people as acting as if they are good or evil in response to law and incentives. In the epigraph to *The Moral Economy*, Bowles cites Hume directly:

> Every man ought to be supposed to be a *knave* and to have no other end, in all his actions, then in his private interest. By this interest we must govern him, and, by means of it, make him, notwithstanding his insatiable avarice and ambition, cooperate to public good.

Economic policy today largely proceeds according to Hume's maxim. How should the legislator accomplish his goal? Incentives. Price correctly, and the rest will follow.

Most of *The Moral Economy*, however, is dedicated to the experimental and empirical evidence that this maxim is incorrect. Herein lies the book's second public good for economists. *The Moral Economy* provides an exhaustive, yet insightful presentation of the plethora of experiments and empirical findings that show incentives can be detrimental to social goals. The book recounts and unpacks well-known, and not-so-well-known, experiments and case studies that show, mostly, that incentives "crowd out" socially beneficial actions. In pointed contrasts, *The Moral Economy* also presents instances where incentives appear to "crowd in" socially minded behavior.

With this evidence in the background, *The Moral Economy* provides a fundamental critique of microeconomic policy analysis, and herein lies the book's third public good. *The Moral Economy* argues that preferences are endogenous; policies based solely on incentives can erode individuals' intrinsic social motivations. Hence, the analysis of microeconomic policy is subject to the same critique that Robert Lucas levied on macroeconomic policy analysis. Policy changes the underlying structure of the economy, hence ceteris paribus arguments do not hold. Economists can no longer proceed with an unquestioned premise that preferences are innate, individual characteristics. Individuals' preferences are not separable from incentives. For decades, Bowles has been arguing that preferences are endogenous. Finally, it seems that the economics profession has (almost)

caught up, as much of the research that Bowles relays in the book is published in leading, mainstream economic journals.

In the last part of the book, though, *The Moral Economy* ventures into new territory. Bowles argues that successful policies must combine incentives and moral messages, exploiting complementarities between the two. Economic policies are conveyed by words, and the wording can be more or less offensive or convincing. Incentives are also not separable from how they are described and explained. *The Moral Economy* leaves us with the conclusion that the messaging is critical and has the power to align individual objectives with those of the legislator.

While *The Moral Economy* is a tour de force, the book ultimately leaves economists with open questions and challenging work ahead. *The Moral Economy* gives the legislator, and his or her economist advisors, little guidance on how to proceed. How should the legislator actually go about designing this policy that engages both economic motives and social mores? As advisors to policy makers, what theory or framework should economists use? What data do we need? This essay suggests that first, economists need to engage in more, and perhaps untraditional, data to uncover the relationships between incentives and social preferences. To date, most work involves experiments and empirical work showing that incentives lead to worse social outcomes, but the mechanism remains largely hidden. Second, and relatedly, economists must study the social context and local institutions. Different societies have different understandings of monetary inducements, and people in different societies have different norms as to appropriate and inappropriate behavior, moral and otherwise. To understand the relation between incentives and behavior, economists must grapple with the full range of human motivations. Economists must be attentive to social details, frameworks for which are only beginning to find a place in our toolkit. This argument runs parallel to Esther Duflo's (2017) Ely Lecture, that successful development policy depends on attentiveness to details and tinkering with the implementation. In the realm of policies to enhance the public good, there is no escape that the devil is in the details and to construct the type of policies *The Moral Economy* envisages, economists must understand the local norms, identities, and institutions.

The remainder of this essay proceeds as follows. Section 2 discusses *The Moral Economy*'s view of economic policy through the lens of political philosophy. In section 3, the essay relays the book's exhaustive presentation of the experimental and empirical evidence by recounting several key experiments that run through the book. Section 4 considers *The Moral Economy*'s view of policy analysis in light of the experimental evidence. Section 5 pushes into *The Moral Economy*'s new and challenging territory of incentives and messaging. Section 6 argues that the devil of constructing successful policies lies in the social details.

2. *Political Philosophy and Economic Policy*

The Moral Economy frames the discussion of incentives and the social good with an insightful and pithy study of political philosophy as it applies to economic policy. *The Moral Economy* thus provides its first, valuable public good. In one place, Bowles gives an accessible and critical description of the philosophical evolution of the assumption of *Homo economicus*—the fictive creature that makes rational choices based on self-interest alone and that populates the classic economic models on which much policy is based. For economists who have not studied the history of economic thought, it might come as a surprise that the founding philosophical traditions of economics recognize the fallacy of self-interest; in these tracts, people are not

only self-interested, but also care about others and the social good. For those who have read bits and pieces, or more, of these works, *The Moral Economy* eloquently and summarily deconstructs the philosophical underpinnings of the neoclassic paradigm.

The book argues that good government and civic behavior go hand in hand. Political philosophers, however, offer different perspectives on this synergy. Chapter 2 of *The Moral Economy* begins with ancient understandings of good people and good government. Aristotle holds that "good laws make good citizens" by inculcating habits and social virtue. And Confucius advises government that people become "upright" when guided by "virtue" and regulated by "ritual" rather than by orders and penalties. These quaint notions, Bowles writes, do not jibe with the current view of economic policy as based on incentives, rewards, and punishments. The current paradigm traces back instead to Machiavelli. In Machiavelli's view, people have basic interests that must be restrained and channeled by laws. Laws do not make people good; rather, laws induce people to do good acts. Fast forward from Machiavelli to Hume, Jeremy Bentham, and Adam Smith, and we have the notion that institutions—governments and markets—harness basic self-interest to achieve public benefits.

Yet, Bowles tells us, we would be wrong to think these philosophers actually thought human beings are amoral and asocial creatures and that only economic self-interest and laws lead people to do good acts. Rather, Bowles writes that Hume studied scoial norms, and in Smith's *Theory of Moral Sentiments* people are naturally interested in the fortunes of others (pp. 16–17). Bowles informs us that even founders of the neoclassical paradigm, such as Francis Edgeworth, understood self-interest as a "handy but empirically false" abstraction. The same is actually also true for Machiavelli, who held that men as "knaves" was a prudent assumption, rather than a true representation or generalization of peoples' characters.

So far, *The Moral Economy* remains on well-trodden territory. Many researchers in economics today understand the paucity of the *Homo economicus* assumption. The new and challenging argument of *The Moral Economy*, what Bowles argues these philosophers did not understand, is that the assumption of self-interest is not innocuous.

While self-interest might be a convenient assumption, *The Moral Economy* argues it is likely dangerous assumption. There is the immediate danger of formulating policy based on a mis-specified model. More insidiously, policies based on models of self-interest can backfire by eroding people's intrinsic desires to do good. As Bowles warns, "moral and other prosocial behavior would be affected—perhaps adversely—by incentive-based policies designed to harness self-interest" (p. 21). The Machiavellian pragmatism of constructing a constitution for knaves is not neutral. A government that constructs policy under the presumption that people are "knaves" will indeed have a citizenry that act like knaves.

This argument underlies much of the book: not only does the policy matter per se, but also the premise and intent of the policy. Other economists have made and elaborated similar observations about the importance of intent. In Rabin's (1993) seminal article introducing fairness to game theory, monetary payoffs are not the whole story; individual payoffs depend also on whether the other player is helpful or hurtful. In Benabou and Tirole's (2006) theory of prosocial behavior, people want to think of themselves (and others) as motivated not only by the money or but by some intrinsic, socially oriented motivation. People gain utility from thinking of themselves as generous or socially minded and from others' opinions of their generosity. Thus, people will signal this desirable trait or intention through their actions.

The departure in *The Moral Economy* is that individual motivations and any socially minded preferences are not fixed. The same person can either be intrinsically helpful or harmful. The same person could either be motivated to help others or motivated only by money. The motivation, per se, is malleable and depends on the regime under which people live. *The Moral Economy* poses this endogeneity of preferences, or types, as fundamental to the study of policy. Preferences and social mindedness develop within the society and react to and inform the presumptions of governance. *The Moral Economy* bases this proposition on the plethora of experimental and empirical evidence that monetary incentives can crowd out socially minded behavior and that the effects can persist.

3. *Experiments and Evidence of the Effects of Incentives*

The Moral Economy's insightful presentation of the experimental evidence is then the second public good the book provides. The book admirably, exhaustively relays both lab and field experiments, giving the reader confidence not only in the experimental findings, but in their external validity. A few quintessential experiments illustrate the main arguments and run the course of the book; each shows either incentives can crowd out socially beneficial behavior or can crowd in socially beneficial behavior.

For the crowding out, *The Moral Economy* reviews Richard Titmuss' classic work on blood donation (Titmuss 1970), where introducing payments led to a decrease in giving. The leitmotif of *The Moral Economy*, however, is the now well-cited experiment by Gneezy and Rustichini (2000), which introduced a monetary fine for picking up children late at certain Haifa day-care centers.[1] The striking result of the experiment is that after the fine was imposed, the fraction of late pickups doubled relative to the control, and the tardiness persisted even after the fine was lifted. As in the case of blood donations, there are competing explanations for the outcome that incentives crowd out socially minded behavior. The fine, for example, could have removed uncertainty as to the actual cost to the day-care center of tardiness. *The Moral Economy*'s preferred explanation is that the monetary penalty undermined parents' ethical obligations and respect for teachers. Rather than inconsiderate and socially inappropriate, tardiness became a commodity that could be purchased.

The Moral Economy also returns throughout the book to a remarkable experiment by Cardenas, Stranlund, and Willis (2000) that shows how incentives perform relative to a baseline. In rural Columbia, the researchers constructed an experimental task that mimicked the actual common resource problem participants faced in their villages. Each participant chose how much to extract from an (experimental) forest, understanding that lower individual extraction levels would maximize group payoffs. Participants first played the resource game, and extraction levels were well below the individually optimal levels. Then one set of participants was subject to individual fines for over-extraction, while another set of subjects was allowed to discuss the game. By now, readers can predict the result. The communication treatment led to a bit less individual extraction relative to the baseline. As for the fine, individuals extracted less in the first rounds. But then extraction levels rose and reached individually optimal levels; thus in the fine treatment, extraction levels were ultimately well above the baseline. The fines, it is argued, crowded out the social preferences that were evident in the first part of the experiment.

[1] *The Moral Economy* provides an incorrect citation. This experiment appears in Gneezy and Rustichini (2000).

While much of *The Moral Economy* is devoted to the detrimental effects of incentives on socially desirable behavior, Bowles also presents contrary evidence that incentives can work. Indeed, incentives sometimes work extraordinarily well and the success of particularly well-designed policies is critical to the book's overall argument. Echoing Aristotle, while "bad governance"—in the form of economic policy based on incentives—can lead to "bad habits," the converse is also true. "Good governance"—also in the form of economic policy—can instill "good habits." The exemplary case study, here and elsewhere, of incentives that worked, is the small tax placed on plastic bags in Ireland in 2002. Within two weeks, *The Moral Economy* recounts, use of the bags fell precipitously, by 94 percent (Rosenthal 2008).

In light of these experiments and the many others described in the book, *The Moral Economy* provides the theoretical outlines of how a policy maker should take into account such changes when trying to reach a target. The policy maker must consider not only the direct effects of the policy, but the indirect effects on preferences. The outline is really only a sketch; the more substantial contribution is the book's broader critique of policy analysis.

4. *Evaluating Policy—Incentives and the Lucas Critique*

The Moral Economy pulls apart the practice of policy evaluation when the policy can change preferences. The book presents economists with a dire predicament, and here Bowles produces a third public good by relating the problematic of policy's effect on individual preferences to the Lucas critique. While this connection is a bit buried in the middle of the book (pp. 153–56), these dense three pages should, in particular, be required reading for graduate students and practitioners alike.

More than forty years ago, Robert Lucas (1976), studying the effects of macroeconomic policy, remarked that policies themselves can change people's beliefs about the actions of others and the government. Since beliefs are critical to their own optimizing decision, the policies essentially change the underlying structure of the economy. Lucas summarizes:

> Given that the structure of an econometric model consists of optimal decision rules of economic agents, and that optimal decision rules vary systematically with changes in the structure of series relevant to the decision maker, it follows that any change in policy will systematically alter the structure of econometric models (p. 41).

The Moral Economy cites this passage in full. Any policy maker must then take into account the change in beliefs that follows from the changes in macroeconomic policies. Substituting the jargon to that of microeconomics, the observation relates directly to the effect of policies on individual beliefs and individual preferences. Bowles also summarizes the thinking of two prominent scholars, Albert Hirschman and Michael Taylor, who were his own personal inspirations. Though somewhat on the edges of mainstream economics, they made such observations relative to microeconomics at around the same time as Lucas. Regulations and legal structures not only mete out punishments, they change people's beliefs and preferences.

The Moral Economy argues that since much of economic activity takes place outside of canonical markets, the effects of incentives on preferences are particularly important to understand. Social norms and preferences support much of economic activity. The markets described in economics textbooks—where prices mediate the exchange of goods and services between anonymous individuals—are actually few

and far between in the real world. *The Moral Economy* describes these ideal, textbook markets as defined by Max Weber, James Buchanan, and others, as requiring "neither personal affection nor long-term relationships among people engaged in exchange" (p. 175). Yet, absent the full set of laws and institutions to completely enforce contingent contracts, exchange takes place between people who have names and faces and who ultimately must trust each other to some extent.

The Moral Economy dedicates much of the second part of the book to arguing synergy between these direct exchanges and social preferences. In many settings, aspects of an exchange—the good or service's quality, as a canonical example—are difficult to verify in advance, or even after the exchange. Incomplete contracts, it is argued, go hand-in-hand with trust, ongoing relations, and sharing surplus. Distrust, short-term relations, and taking as much surplus as possible go hand-in-hand with complete contracts.

Such synergies and complementarities have appeared elsewhere in economics. A relatively large literature in industrial organization and in development economics, not cited in *The Moral Economy,* shows the complementarities between long-term supply relations and high-quality and arms-length supply relations and low quality (Taylor and Wiggins 1997). The availability of markets, or anonymous exchange, can destroy mutually beneficial repeated dealing where partners extend each other trade credit or otherwise enhance the exchange in noncontractible ways (Baker, Gibbons, and Murphy 2002; Kranton 1996; and McMillan and Woodruff 1999).

The Moral Economy presents the strong experimental evidence that demonstrates these synergies in the lab. Brown, Falk, and Fehr (2004) constructed a set of tasks that mimicked legally enforced complete contracts and informally enforced incomplete contracts. Suppliers could produce higher- or lower-quality goods. In a complete contracting treatment, the experimenter enforced the quality level promised by the supplier. In the incomplete contracting treatment, the supplier could renege on any promise. Participants had identification numbers, so it would be possible for participants to keep track of past interactions. The results match the theory well. In the incomplete contracting condition, as the rounds progressed, participants chose more often to interact with the same partners, buyers offered these sellers prices that would more than cover costs, and sellers would forgo higher price offers from other buyers. While the theory of repeated games would also largely predict these results, *The Moral Economy*'s interpretation of social preferences and social norms is consistent with the outcome. In the incomplete contracting condition, participants "learned to trust trading partners" and "remained loyal to them" (p. 179). The incomplete contracting result supports *The Moral Economy*'s overall argument that when people understand actions as explicitly compensated by awards, they behave accordingly. Monetary incentives can "frame" actions.

The book further presents, at length, several experiments that show that monetary incentives, even when no longer in place, can have long-lasting detrimental effects. Thus the imposition of the incentives, to echo Lucas again, changes the underlying structure. The Haifa day-care center experiment described above is the first case in point; when the fine was imposed, parents were more often late picking up their children, and the higher levels of tardiness persisted even when the fine was removed. Bowles also describes the well-known and well-aged psychology experiment by Lepper, Greene, and Nisbett (1973) who tested whether extrinsic rewards might reduce children's intrinsic

motivations in school. In the experiment, young children were rewarded with stars and ribbons for drawings. In a later session, these children spent less time drawing and produced lower quality drawings relative to a control group.

To further its argument, *The Moral Economy* describes in detail an experiment that mimics the adult workplace. Gächter, Kessler, and Königstein (2011) study worker compensation; one participant has the role of employer and another participant has the role of employee, where the employer offers the employee a wage and the employee chooses a level of production. Employers and employees are rematched randomly and anonymously from the subject pool. The first phase of the experiment had three treatments. The first involved no additional incentives. In the second, the employer could specify a wage and a target level of production, and a fine ex post if the production level was not met. In a third condition, the worker would be paid a bonus ex post if the production level was met.

The Moral Economy's emphasis is on the second phase of the experiment, where all three sets of participants performed the task without any additional incentive. That is, the participants who had been subject to either the fine or the bonus were no longer subject to these inducements. The participants who received a fine or a bonus in the first phase offered near-zero wages and exerted near-zero effort. The participants who had not been exposed to the fine or bonus continued to offer higher wages as employers and exert higher levels of effort as employees (Gächter, Kessler, and Königstein 2011).

Taken together, *The Moral Economy* provides an empirically informed global critique of incentives as policy in the spirit of Lucas. Underlying expectations depend on the institutions; people learn about how to interact with one another within institutions that themselves presume the nature of motivation. Bowles has been studying these issues and endogenous preferences for more than forty years. *The Moral Economy*, presenting this plethora of evidence, much of which is published in leading mainstream journals, shows in a perhaps indirect way that many of these ideas are now more-or-less accepted within the economics discipline. Without minimizing Bowles's effort and *The Moral Economy*'s masterful compendium of this work, the most radical part of the book is, perhaps, in its final chapter.

5. *Incentives and Messaging*

In the final chapter, entitled "A Mandate for Aristotle's Legislator," *The Moral Economy* returns Aristotle's legislator to ancient Athens for lessons on how to construct policy that works—both to provide proper incentives and to instill (or at least not undermine) social preferences. When and why do incentives crowd in rather than crowd out socially desirable behavior? Why did the tardiness fine at the Haifa day-care center backfire? Why did the plastic bag tax in Ireland succeed? In chapter 7, *The Moral Economy* delves more deeply into the policies, taking the reader behind the scenes, into details economists typically do not consider: the political process that leads to the policies and the messaging of the policies themselves.

Chapter 7 begins with an account of how the Athenian citizens' assembly in 325 CE mobilized the city to establish quickly an expensive colony and naval station in the Adriatic Sea to defend the polity. The Athenians, it seems, masterfully designed a mechanism where people would feel obligated to contribute. Ship commanders and providers were drafted, but were allowed to find others to take their place. Those who stepped forward, however, were rewarded. The first, second, and third volunteers received monetary prizes, publicly. Those

who did not contribute were fined, with the proceeds described as going to the overall benefit of the city. *The Moral Economy* relates the history as a tell-tale, of how people can be induced to provide their own labor and resources for the public good. The key is that the monetary inducements were not just incentives per se. The exhortations to contribute to the mission and the monetary rewards and punishments were constructed to be complements, not substitutes.

The Moral Economy then asks the reader to consider how a time-traveling Athenian would construct and communicate the Haifa day-care policy. In the experiment itself, a sign was posted on the door of the day care, simply stating the fine and its approval by the relevant authority:

> Since some parents have been coming in late we (with the approval of the Authority for Private Day Care Centers in Israel) have decided to impose a fine on parents who come late to pick up their children. As of next Sunday a fine of NIS 10 will be charged every time a child is collected after 16.10 (p. 189).

As an experimental condition, the sign is straightforward and neutral. But of course no frame is neutral; the absence of any further explanation or appeals to parents is itself a frame.

Bowles proposes instead the following sign, following the lead of the Athenian polity, with explanations of the tardiness policy, appeals to parental responsibilities toward staff and their own children, and public rewards for timeliness and fines for tardiness, with the proceeds to be allocated to the greater good. The explanation and the mechanism would be quite elaborate and it is worth citing Bowles's suggestion in full to drive home the contrast with the cryptic sign used in the experiment. The following announcement would be posted on the day-care door and notably come from the "Council of Parents," not day-care management:

> The Council of Parents wishes to thank you for arriving on time to pick up your children, since this reduces the anxiety that the children sometimes feel and allows our staff to leave in a timely manner to be with their own families. We will recognize all parents who have a perfect record unblemished by lateness for the next three months with an award of NIS 500, to be given at our annual parents and staff holiday party, with an option to contribute your award to the school's Teacher of the Year celebration.
>
> […] Those who arrive more than ten minutes late, however, will pay a fine of NIS 1,000, with the payment of the fine publicly transmitted also at the holiday party. In the unlikely event that the occasion for such a fine arises, the payment will also support the Teacher of the Year celebration. […] Of course, sometimes it is impossible, for reasons beyond your control to arrive on time; and should this occur, you may explain your circumstances before a committee of parents and staff, and if the lateness was unavoidable or if the fine would cause extreme hardship, the lateness will be publicly reported but no fine will be imposed (pp. 189–90).

Bowles concludes, asking rhetorically "would this Athenian version of the experiment have reversed the crowding out that occurred in the absence of moral framing? It might have."

While we do not have the answer to this question, *The Moral Economy* contrasts the Haifa experiment to the Irish plastic bag case. The plastic bag tax in Ireland was imposed neither suddenly nor cryptically. Rather, the tax arose from a lengthy political process with extensive public deliberation and consultation. A publicity campaign, showing the environmental harm of plastic bags, accompanied the rollout of the tax. Along with the fine were appeals to the public good and the social costs of using plastic bags. The contrast with the messaging of the Haifa fine is stark, as Bowles writes: "In Haifa, the fine seems to have said, 'Lateness is okay as long as you pay for it,' while in Ireland the message was something like, 'Don't trash the Emerald Isle!'" (p. 203).

The Moral Economy, then, ultimately concludes with two new and meaningful insights concerning the effect of monetary incentives. Both of these insights mesh with the overall premise of the book concerning the relationship between "good governance" and "good habits." First, monetary incentives cannot be divorced from their policy process. Arbitrary—or seemingly arbitrary—incentives imposed without public consultation and buy-in are unlikely to succeed and could even backfire. Second, monetary incentives cannot be divorced from their messaging. Bonuses and fines should be accompanied by explanations and exhortations that appeal to the common good. The cryptic sign hung on the door of the day care is one framing; the lofty language reminiscent of ancient Athens is another framing. Of course, we will not know the answer to Bowles's question about whether his day-care policy and messaging would work unless intrepid researchers conduct a new set of experiments, with different rewards and fines, different messaging, and different consultative processes. But perhaps these sorts of investigations are exactly what economists should be doing next.

6. *The Devil Is in the Details*

So here we arrive at the end of *The Moral Economy*'s journey. The book presents the substantial body of experimental work that indicates—over and over again—that incentives and preferences interact. The book proposes that a sophisticated legislator should take this interaction into account when formulating policy. The book, however, gives the legislator little practical guidance. Thus, economists have much work to do in order to provide the legislator with what he or she needs to construct successful policy. First of all, the legislator needs better tools to determine whether incentives are incentives per se, whether incentives convey information, or whether indeed incentives also change preferences. Second, the legislator needs better tools to understand the political and social environment in order to better engage the public in policy making and to better construct the message.

The first task ahead is to confront head-on, conceptually and empirically, the relationship between incentives and preferences. The experimental results are convincing that incentives have "surprising" effects, but the underlying mechanisms are not fully understood or tested. While the authors of the studies and Bowles have their preferred explanations, perhaps economists can do better. The first step would be to consider plausible mechanisms in a parsimonious model with fixed preferences, including social preferences. In this setup, new incentives could simply convey information: perhaps in Titmuss' case, people learned that donating blood was not as socially valuable as they thought, and so stopped. Perhaps Haifa parents learned that picking up a child late was not that costly to the day-care center. Perhaps Irish shoppers learned that plastic bags harmed the environment and that small fees add up and capture the costs of pollution. People could be socially minded with exogenous preferences, and the imposition of a reward or fine conveys information as to the true social benefits or social costs of individual actions.

To rule out (or rule in) information and to consider less standard mechanisms such as changes in preferences, economists should devise diverse methods to test alternative hypotheses. Experimental work is certainly one possibility, where incentives can cleanly be introduced into tasks and where information sets could be part of the experimental design. But the real world is messier than an experiment; information sets and individual beliefs are typically neither controlled for, nor observed by, the researcher. The same is obviously true for individual preferences. This paucity of data should end.

Experimental work is beginning to include post-experiment surveys that ask not only for demographic information, but for attitudes and beliefs. Empirical studies could engage less traditional empirical methods, including surveys and hands-on anthropological and historical research. While these methods might not allow researchers to make causal claims, correlations, or the absence thereof, might be able to rule out certain mechanisms. The insights of detailed case studies, with appropriate taxonomies to understand all the moving parts, would be useful to our legislator, who at the moment is groping in the dark (or searching under the lamppost for keys).

The second, and perhaps more radical, work economists should do is study and dissect the political and social environment in order to better formulate and communicate policies. If economists are to design policies that work, economists should know the context and understand people's motivations within that context. Bowles himself writes that economists should be aware that people not only have acquisitive motives—"to *get* things"—but also constitutive motives—"to *be* someone" (emphases in the original, p. 192). This exhortation relates directly to Benabou and Tirole's (2006) schema, where people take actions to signal to themselves who they are. It also relates to the complementary paradigm in my work on identity with Akerlof (Akerlof and Kranton 2000, 2010); people have identities and desire to follow identity-specific norms for how to behave. As Bowles elaborates, sometimes the acquisitive and constitutive motivations coincide and sometimes they clash. The legislator would need to know which is which. To know, the legislator would also need to understand much more about the constitutive motives than economists typically consider, if at all. The legislator needs to know the local norms, customs, identities, and politics.

To illustrate, consider again the fine for tardiness at the day-care center. The announcement in the experiment, as discussed above, perhaps was set to avoid framing the fine. But, it is worth stating again, there is no such thing as a neutral frame. The fine was announced in a context, and this terse notice was understood in that context. How would an Israeli parent, in this particular milieu in Haifa, view the notice, as opposed to, say, an American parent even from a similar socioeconomic stratum? Consider Bowles's new messaging from the time-traveling Athenian. How well does this messaging, and indeed the policy itself, fit the local context? Are annual parent–staff holiday parties a norm? Is it acceptable to select one teacher for a Teacher of the Year award? Would public shaming associated with the fine be at all appropriate? Would the Council of Parents announce the policy? Is there indeed a Council of Parents that has a say in policies? All these details matter and are likely to be specific to the setting.

Economists should not despair, however, and throw in the towel in frustration because the real world is messy and local knowledge is critical. Rather, economists should view these observations as an opportunity to rethink our models and our basic paradigms. Economists can develop general principles to better relate our theories to the local social and political context. Understanding the relevant identities and norms is, of course, one step in this direction. Understanding the relevant institutions and informal and formal legal structures is another step. Understanding the language and messaging is yet another step. This is all difficult and unchartered territory.

For inspiration then, *The Moral Economy* should perhaps be read in conjunction with Duflo's American Economic Association 2017 Ely lecture, "The Economist as Plumber" (Duflo 2017). Duflo calls on economists,

as advisors to policy makers and as designers of policy, to focus on the details that prove to be critical for success. Economists might deem these details "uninteresting," but Duflo relates example after example of microeconomic interventions whose success or failure turned on details. In Duflo's analogy, an economist should be a plumber, who finds and plugs the unanticipated leaks as the policy rolls out and "tinkers" with the policy to make the intervention ultimately fit the environment.

The lead example, which Duflo writes actually does relate to plumbing, concerns interest-free loans offered to poor families in Tangier to cover the marginal cost of connecting their homes to the main water system. Household water connections would save people considerable time and relieve considerable stress of collecting water from public taps. The uptake of the loans, however, was less than ten percent. Why? The application process, including going to the municipal office with the necessary documents, was overly burdensome. When the researchers offered procedural assistance, the uptake increased to 69 percent. The lesson, and the lesson from the many other examples in the lecture, is that economists should be both anticipating such bottlenecks and continually tinkering with interventions.

The reason the Tangier uptake was low, and the remedy, fall within the bounds of standard economic thinking, i.e., transactions costs. The numerous policy successes or failures relayed in the lecture are mostly attributed to information asymmetries, bounded rationality, organizational dysfunction, lack of training, or other faults with the "incentive architecture." Indeed, Duflo encourages economists to engage the details, since they have "implications for issues which are an economist's bread and butter: incentives, information, and imperfect rationality, etc." (p. 2).

While the Ely lecture does not dive into details that might fall outside standard models—details related to local social norms, morality, or identity—Duflo does relay two experiments on framing and social motivations. These experiments test different descriptions of public-sector jobs, and they seem to lead to different conclusions about the effect of presumed economic incentives (Duflo 2017, p. 12). Ashraf, Bandiera, and Lee (2015) find in Zambia that health workers responding to an advertisement emphasizing "career advancement" were much more effective on the job than those responding to an ad emphasizing "service to community." On the other hand, Deserranno (forthcoming) finds in Uganda that advertisements that emphasized the maximum, rather than the minimum, salary for a health promoter position attracted more candidates but less socially motivated candidates.

These experiments and their policy details bring us back to *The Moral Economy* and its prescription to marry incentives with moral messages. Economists need to pay attention to the details of messaging and their social meanings, and experimental methods can indeed be brought to bear on these critical aspects of policy. Combining the spirit of Duflo's Ely lecture and the imperative of *The Moral Economy*, a research agenda emerges.

But the hard work does not begin and end with experiments. Would the Athenian's proposed day-care policies and messaging reduce tardiness more effectively than the cryptic announcement? Maybe yes, but maybe no. The Athenian's policy and appeals might be completely inappropriate for a particular context, as discussed above. To formulate the alternative policies and messages, and then to conduct the experiment, the researcher must do much more work, to gain a full grasp of the local, operative social norms and political processes. To uncover the mechanisms behind any finding, the researcher should further engage both

traditional and less-traditional data, including surveys, interviews, and anthropological field work. And a theoretical framework of individual decision making that includes morality, identity, and norms would give structure to the findings, leading to more general conclusions about policy, incentives, and moral (or other) messages.

The Moral Economy concludes with another tell-tale of successful governance: the initiatives of Antanas Mockus, the two-term charismatic mayor of Bogotá from 1995–97 and 2001–03. An astute reader will see the innovative, multifaceted policies that fits with local mores. To tame the chaotic traffic and reduce related deaths, along with the police-enforced regulations and fines, the mayor sent mimes in clown face into the city to shame drivers and jaywalkers. A civic culture program in Mockus' second term gave thumbs-down cards to residents to flash at recalcitrant drivers. A water-saving campaign awarded prizes to those who cut back on their consumption and publicly fined those who used water excessively. The mayor also starred in a television commercial saving water in the shower with his wife. The mayor, like the Athenian assembly, found exactly the right combination, as Bowles writes, of "material interests and moral sentiments, framed so that the two work synergistically rather than at cross-purposes" (p. 220). Thus, a virtuous cycle is possible, when policy makers are attentive to the local environment, and critically themselves have credibility and authenticity, and address issues of public concern with complementarities between incentives and interests.

Reading Duflo's Ely lecture and reading Bowles's *The Moral Economy,* it is no wonder that incentives do not always work and that they can backfire. Economists can no longer ignore that incentives are just the tip of the iceberg of economic policy. Incentives come along with, necessarily, a message and are implemented in a local, detailed social and political context, in the real world where people most often interact face to face.

The frontier of economic policy recommendation and implementation is then not incentives per se, but tinkering within the incentives, tinkering within the institutions, and tinkering with the message. Similarly, the frontier of economic theory is the study of social context and messaging, which build directly on the local mores and identities that underlie people's motivations. And the frontier of empirical work is engaging the effect of policy on preferences. With the fruits of such endeavors, economists can conceive successful policies, which are necessarily attentive to the local situation, to the local social context, and complementary to the local norms and customs and people's understandings of obligations, rights, and responsibilities.

References

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115 (3): 715–53.

Akerlof, George A., and Rachel E. Kranton. 2010. "Identity Economics." *The Economists' Voice* 7 (2).

Ashraf, Nava, Oriana Bandiera, and Scott S. Lee. 2015. "Do-Gooders and Go-Getters: Career Incentives, Selection, and Performance in Public Service Delivery." Working Paper.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117 (1): 39–84.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652–78.

Brown, Martin, Armin Falk, and Ernst Fehr. 2004. "Relational Contracts and the Nature of Market Interactions." *Econometrica* 72 (3): 747–80.

Cardenas, Juan Camilo, John Stranlund, and Cleve Willis. 2000. "Local Environmental Control and Institutional Crowding-Out." *World Development* 28 (10): 1719–33.

Deserranno, Erika. Forthcoming. "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda." *American Economic Journal: Applied Economics*.

Duflo, Esther. 2017. "The Economist as Plumber." *American Economic Review* 107 (5): 1–26.

Gächter, Simon, Esther Kessler, and Manfred Königstein. 2011. "The Roles of Incentives and

Voluntary Cooperation for Contractual Compliance." University of Nottingham Centre for Decision Research and Experimental Economics Discussion Paper 2011-06.

Gneezy, Uri, and Aldo Rustichini. 2000. "A Fine Is a Price." *Journal of Legal Studies* 29 (1): 1–17.

Kranton, Rachel E. 1996. "Reciprocal Exchange: A Self-Sustaining System." *American Economic Review* 86 (4): 830–51.

Lepper, Mark R., David Greene, and Richard E. Nisbett. 1973. "Undermining Children's Intrinsic Interest with Extrinsic Reward: A Test of the 'Overjustification' Hypothesis." *Journal of Personality and Social Psychology* 28 (1): 129–37.

Lucas, Robert E. 1976. "Econometric Policy Evaluation: A Critique." *Carnegie–Rochester Conference Series on Public Policy* 1: 19–46.

McMillan, John, and Christopher Woodruff. 1999. "Interfirm Relationships and Informal Credit in Vietnam." *Quarterly Journal of Economics* 114 (4): 1285–320.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5): 1281–302.

Rosenthal, Elisabeth. 2008. "Motivated By a Tax, Irish Spurn Plastic Bags." *New York Times*, February 2, A3.

Taylor, Curtis R., and Steven N. Wiggins. 1997. "Competition or Compensation: Supplier Incentives under the American and Japanese Subcontracting Systems." *American Economic Review* 87 (4): 598–618.

Titmuss, Richard M. 1970. *The Gift Relationship: From Human Blood to Social Policy*. London: Allen and Unwin.

This article has been cited by:

1. John Gibson. 2021. The micro-geography of academic research: How distinctive is economics?. *Scottish Journal of Political Economy* **68**:4, 467-484. [Crossref]
2. Malte F. Dold, Matías Petersen. 2021. Stability of the liberal order, moral learning, and constitutional choice: an unresolved tension in James Buchanan's political economy. *Review of Social Economy* **25**, 1-27. [Crossref]
3. Dominik Rehse, Felix Tremöhlen. 2021. Fostering participation in digital contact tracing. *Information Economics and Policy* **109**, 100938. [Crossref]
4. Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel. 2021. Monitoring institutions in healthcare markets: Experimental evidence. *Health Economics* **30**:5, 951-971. [Crossref]
5. Matthew Hale, Graham Raymond, Catherine Wright. 2020. List of publications on the economic and social history of Great Britain and Ireland published in 2019. *The Economic History Review* **73**:4, 1153-1202. [Crossref]
6. Marianne Johnson. 2020. Where Economics Went Wrong: A Review Essay. *Journal of Economic Literature* **58**:3, 749-776. [Abstract] [View PDF article] [PDF with links]
7. Niklas Elert, Magnus Henrekson. 2019. The collaborative innovation bloc: A reply to our commentators. *The Review of Austrian Economics* **32**:4, 349-361. [Crossref]
8. Magnus Henrekson, Johan Wennström. 2019. 'Post-truth' schooling and marketized education: explaining the decline in Sweden's school quality. *Journal of Institutional Economics* **15**:5, 897-914. [Crossref]