PAR PUBLIC ADMINISTRATION REVIEW

Adam Oliver
London School of Economics and Political Science

Viewpoint
Stephen E. Condrey
and Tonya Neaves,
Associate Editors

Towards a New Political Economy of Behavioral Public Policy

Adam Oliver is a behavioral economist at the London School of Economics and Political Science. He has published and taught widely in the areas of health economics and policy, behavioral economics, and behavioral public policy. He is a founding editor in chief of the journal *Behavioural Public Policy*. He is the editor of a volume also titled *Behavioural Public Policy* (Cambridge University Press, 2013) and the author of *The Origins of Behavioural Public Policy* (Cambridge University Press, 2017) and *Reciprocity and the Art of Behavioural Public Policy* (Cambridge University Press, 2019).
E-mail: a.j.oliver@lse.ac.uk

Abstract: *The dominant normative framework in behavioral public policy postulates paternalistic intervention to increase individual utility, epitomized by the so-called nudge approach. In this article, an alternative political economy of behavioral public policy is proposed that sits within, or at least closely aside, the liberal economic tradition. In short, rather than impose utility maximization as the normative ideal, this framework proposes that policy makers provide an environment that is conducive to each person's own conception of a flourishing life, while at the same time regulating against behaviorally informed harms and for behaviorally induced, otherwise forgone, benefits.*

Behavioral public policy is the application of insights from behavioral economics specifically and behavioral science more broadly to public policy design. It is, in any substantive sense, a relatively recent endeavor, although several decades of social science scholarship underpin the approach (for recent accounts of the development of the field, see Oliver 2017; Thaler 2015). Several conceptual behavioral public policy frameworks now exist. Some of these approaches aim to educate—to "boost"—people about their possible behavioral biases so that they may make more savvy decisions (Gigerenzer 2015; Hertwig 2017), and others call for people to engage in more deliberative decision-making—to think more for themselves—so as to minimize reflexive errors (John et al. 2011). Some frameworks instead aim to influence automatic decision-making without appealing directly to deliberation while also retaining the notion of liberty (Thaler and Sunstein 2003), whereas still others allow heavy doses of regulation and even bans (Conly 2013). Some focus on improving the well-being of those that the policy interventions target specifically, whereas others look toward reducing harms to, and increasing benefits for, others (Oliver 2015).

The dominant framework in behavioral public policy to date, however, focuses paternalistically on internalities—that is, it aims to change the behaviors of those targeted for their own benefit—and where the normative goal is to improve welfare, utility, or happiness, an approach epitomized by libertarian paternalism, applications of which are known as "nudges" (Thaler and Sunstein 2003, 2008). In this article, I challenge whether, in relation to influencing the behavior of responsible adults, this ought to be the principal normative approach—that is, the political economic framework—that shapes the future of the field. Rather, I argue for a behavioral public policy framework that sits within the liberal economic tradition of, for example, John Stuart Mill ([1859] 1972), although compared with others in this tradition, the approach posited here aims to tackle more forcefully behaviorally informed harms.

The Paternalistic Embrace

As intimated earlier, libertarian paternalism has been by far the most prominent framework in the behavioral public policy discourse. Like its close cousin asymmetric paternalism (Camerer et al. 2003), libertarian paternalism is an approach that seeks to guide people's behavior in particular directions without the use of force or mandates. People are free to continue with their existing behaviors if they wish; Richard Thaler and Cass Sunstein (2008) contend that retaining freedom of choice is the best safeguard against any misguided policy interventions.

Behavioral economists have observed that many people are guided by particular behavioral influences that appear almost innate, including, for instance, loss aversion (i.e., the heavier weight that people attach to losses than to gains of the same magnitude) and present bias (i.e., the heavy emphasis that people place on the immediate moment) (for a review, see Camerer and Loewenstein 2003). These influences conflict with the assumptions of standard neoclassical economic theory and rational choice theory. Underpinning libertarian paternalism is the assumption that of the many quick and automatic decisions that people make each day—decisions that are invariably guided by these behavioral influences—

Public Administration Review,
Vol. 00, Iss. 00, pp. 00. © 2019 The Authors. *Public Administration Review* published by Wiley Periodicals, Inc. on behalf of The American Society for Public Administration.
DOI:10.1111/puar.13093.

some will lead to actions that if people deliberated a little more, they would not take. The core idea in libertarian paternalism is that with knowledge of the behavioral influences, the context or environment, or what Thaler and Sunstein (2008) term the "choice architecture," that people face can be redesigned such that their automatic decisions better align with their deliberative preferences. In short, for a policy intervention to meet the requirements of libertarian paternalism—that is, for it to be a nudge—it has to be liberty preserving, target internalities, and be informed by behavioral science. Each of these three requirements is represented by an axis in figure 1 (see Oliver 2015).

Movement toward the origin on the vertical axis in the figure indicates that a policy is increasingly liberty preserving rather than regulatory. Similarly, movements toward the origin on the horizontal axis and on the diagonal axis, respectively, indicate that a policy is increasingly addressing internalities rather than externalities, and increasingly informed by behavioral science rather than economic rationality. Consequently, a pure nudge is represented by the point where the axes intersect (i.e., the origin). Examples of pure nudges include placing apples at the front and cheesecake at the back of canteen shelves, painting green footsteps that lead to refuse bins on pedestrian pavements (assuming, of course, that those targeted for behavioral change, upon deliberation, would prefer to dispose of their litter more responsibly), and by appealing to loss aversion, allowing people to deposit money into accounts when they attempt to quit smoking on the understanding that the money will be returned to them if their abstinence is maintained beyond, say, six months.

Some maintain that soft forms of paternalism do not go far enough—that respecting liberty ultimately means that these policy interventions will be insufficiently effective at addressing the actions of those who threaten, for example, their own health and financial well-being (Conly 2013). Followers of this point of view tend to call instead for a harder, coercive form of paternalism. Although they acknowledge that self-harming activities such as smoking may well be driven by behavioral phenomena (e.g., present bias), the ends of these activities are, they contend, so self-evidently bad that they call for the activities to be banned on the demand side.

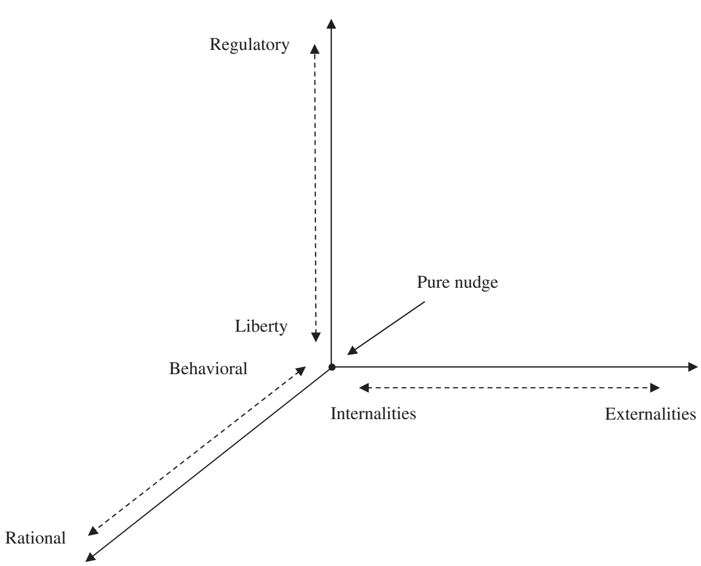

Figure 1 The Requirements of Libertarian Paternalism

As intimated earlier, however, paternalism—soft or hard—is open to the critique that it unreasonably infantilizes the adults it targets and that policy makers ought to have no role—beyond an openly educative one (which I strongly support)—in influencing the behaviors of those of sound mind if their activities impose no ill effects on others, a view akin to Mill's ([1859] 1972) harm principle. A counterargument offered by libertarian paternalists and those whom Robert Sudgen (2018) more broadly defines as behavioral welfare economists is based on their retaining the normative postulate that people should seek to maximize utility. They contend that the behavioral influences will sometimes cause people to fail to act and choose in their own best interests (i.e., to maximize utility). Therefore, following this line of reasoning, libertarian paternalists claim that it is legitimate for policy makers to steer people in the normative direction of utility maximization. But one can question whether this normative direction is, in itself, legitimate.

The View from Nowhere

As noted, rational choice theory, standard economic theory, and behavioral welfare economics all postulate that people ought to maximize utility, which is treated here as synonymous with welfare or happiness. Utilitarianism thus has a prominent place in the intellectual origins of all of these approaches. The widely accepted founding father of British utilitarianism, Jeremy Bentham, believed that humankind is governed by what he referred to as the "sovereign masters of pain," which we seek to minimize, and pleasure, which we seek to maximize (Bentham [1781] 1988). Pain and pleasure, he contended, are feelings that are experienced on a continuum, and thus they can be compared—they guide us on what we ought to do and on what we shall do. From this, he derived his famous dictum that the societal objective ought to be to achieve the greatest happiness for the greatest number.

By the beginning of the twentieth century, most economists and philosophers had resigned themselves to the view that measures of Benthamite cardinal utility—that is, numerical interpersonally comparable indicators of how much pain and pleasure individuals experience in the moment—are impossible to uncover. Nonetheless, modern neoclassical economic theory, which is based on decisions about future experiences rather than experiences in the moment, retained the notion that people ought to, and will want to, maximize their utility. Therefore, aside from allowing for random errors, we can infer that the decisions people make about future episodes, and their retrospective assessments of their previous courses of action, will be consistent with maximizing the amount of utility that they can expect to experience or have experienced. There is now evidence to suggest that this inference is sometimes erroneous.

This evidence, which began to emerge in the early 1990s, shows that when people are asked for their retrospective assessments of different events, there is a tendency for their relative preferences to differ systematically—that is, nonrandomly—from those predicted by utility maximization through a simple aggregation of the moment-to-moment instant utilities as each event is experienced (for a review, see Kahneman, Wakker, and Sarin 1997). Similar observations have been made with respect to prospective evaluations and expected experienced utility. These systematic discrepancies are attributable to factors known as the gestalt characteristics, which are the aspects of an event that respondents perceive as salient, and they include the tendency for people to prefer worse outcomes to precede better outcomes rather than vice versa (Loewenstein and Prelic 1993); an aversion to sudden, steep rates of change in outcomes (Hsee and Abelson 1991); and, most significantly, the tendency for people to place a heavy emphasis on the best, worst (the peaks and troughs), and end moments of an episode (Fredrickson and Kahneman 1993). This last gestalt characteristic is called peak-end evaluation, and it can cause people to underemphasize the duration of the event.

Given that the gestalt characteristics are quite pervasive influences on human decision-making, it is likely that they have an evolutionary explanation. For instance, the drive for survival in the moment might explain the heavy emphasis that is attached to troughs, and the urge to flourish perhaps partly explains a preference for episodes to improve and to end well. To illustrate peak-end evaluation with two of the most striking examples of the effect, let us first consider evidence reported by Donald Redelmeier and Daniel Kahneman (1996), in which sigmoidoscopy and lithotripsy patients recorded their feelings of discomfort every 60 seconds during their procedures. These were measures of their moments of instant disutility, and by aggregating these moments at the end of the procedures, the authors were able to calculate the total experienced disutility of the procedure for each patient. The duration of these procedures varied considerably across patients, from just a few minutes to about an hour. Following their procedures, the patients' retrospective evaluations of the total discomfort that they felt was also recorded on a 10-point scale. The authors reported that the peak-end effect and duration neglect were observed strongly in the retrospective evaluations, and that the retrospective evaluations did not correlate with total experienced disutility.

In a continuation of this line of research, Redelmeier, Joel Piano Katz, and Kahneman (2003) divided a further 682 sigmoidoscopy patients into two groups. Unbeknown to the patients, the tube was left inserted inside those in one of the groups for an additional short period at the end of their procedures for no clinical reason, causing some additional physical discomfort, albeit of lesser intensity than when the procedure was ongoing. Instant disutility and retrospective assessment were recorded similarly to that reported by Redelmeier and Kahneman (1996). In addition to peak-end evaluation and duration neglect again being observed, the group for whom the instrument was left inserted for longer than was clinically necessary generally tended to remember the overall procedure as less painful than those for whom the instrument was removed, on average, sooner. That the former group must have experienced greater total experienced disutility as a consequence of having the procedure duration extended unnecessarily means that the authors recorded a violation of (utility-based) dominance: in other words, the unambiguously worse experience in terms of total experienced disutility tended to be remembered as less bad than an experience that caused less aggregate pain.

Thus, the gestalt characteristics can cause people to assess options in a manner that conflicts with the assumption of experienced utility maximization. This begs the question of whether people who succumb to the gestalts are making errors of judgment, or whether these characteristics are legitimate influences on their preferences that ought to be respected. If one assumes that the normative assumption of utility maximization that is embedded in standard economic theory and behavioral welfare economics aligns with the Benthamite notion of utility (i.e., pleasure and pain), then the adherents of both of these schools of thought would presumably view the gestalt influences as errors.

The prominent neo-Benthamite Richard Layard (2005) expresses the view that to avoid inconsistency across policy actions, all laws and all rules of morality must be based on a single underlying principle, which he, following his intellectual hero, believes ought to be to secure the greatest overall happiness within any population. Since retrospective and prospective evaluations are unreliable indicators of experienced utility, Layard and his followers distrust preferences that are elicited with these methods and propose instead a variety of alternatives that are meant to measure utility in the moment, including sliding scales and knobs that people can move to record their continuous mood, asking respondents to record their mood every minute or so, and the so-called day reconstruction method, in which participants are required to record their current mood and their activities periodically during the day.

Daniel Hausman (2015) notes that neo-Benthamism has drawn serious interest from several governments, but he identifies some problems with this approach, not least in relation to the way in which many of its advocates measure experienced utility and their claims that these measures have policy relevance. For example, Hausman contends that the evidential connection between a person's mood in the moment and how well his or her life is objectively going in relation to his or her health, educational achievement, financial security, or any other policy consideration is weak. Adaptation to a poor state of affairs—for instance, to objectively defined harms imposed on a person by another party—implies that this state will not impact substantially on mood, and yet if the poor state falls within the purview of public policy, then presumably policy makers would want to address it.

The view that adaptation undermines subjective assessment as a means to inform public policy is reflected in Amartya Sen's (1999) notion of a happy slave. That is, few might choose to be a slave even if it were known and accepted that one would adapt well to that state, because self-determination and the opportunity to flourish as one chooses would be substantively—and objectively—curtailed as a consequence. The flip side of underestimating one's plight is exaggerating it, which can happen, according to Hausman, when transitory states seem more important to people in the moment than a more dispassionate, more objective assessment might view them to be.

An even stronger criticism of the neo-Benthamite approach, however, relates to its proposal to guide policy with an aggregation of the moments of instant utility. The contention, summarized by Hausman, is that "a good life is not a sum of the net goodness of its moments … The same sum of momentary experiences can add up to a wonderful life or an incoherent and mediocre one, depending on how the experiences are ordered and what overall narrative they sustain" (2015, 114). David Hume ([1777] 1983) earlier made a similar point in relation to the value of a life using an analogy of how a building cannot be valued by summing independently the values that one would place on, for example, its roof, windows, doorway, and portico without considering how these components fit together. Earlier still, Aristotle ([350 BCE] 1998) noted that happiness cannot be assessed by a feeling or sensation in the moment, but can only be seen through the quality of a whole life.

Thus, when we look forward to an episode or reflect back on it, whether it be a quite fleeting experience (e.g., watching a film), a more extended one (e.g., a summer holiday), a more extended one still (e.g., our university years), or even our whole life (or, if looking forward, what remains of it), the gestalts may matter because they give meaning to the story. In remembering, a focus on these characteristics perhaps aids a deep desire to identify where and when the peaks occurred, how severe the troughs were (perhaps many of us would wish to trade off some total utility to avoid volatility), whether the experience ended well or badly, whether things improved or got worse, and how quickly they got better or worse (for prospective assessments of future events, the tense in this sentence can be altered accordingly). Following this line of argument, particular experiences in an event and how certain experiences fit together, and not the simple aggregation of everything, gives the event—or a life—its meaning. They partly determine how fulfilling the event was to a person and whether the person feels he or she had, or will have, the opportunity to flourish.

Prospective and retrospective evaluations might therefore give a more accurate—that is, more objective—assessment of the impact that a whole experience has had on a person than aggregating the momentary instant utilities because they allow the person to step out of the experience and offer a better view of how good or bad it really was in terms of its positive or negative contribution to a flourishing and fulfilling life. To give a concrete example, perhaps people tend to remember the peak moments of rearing children, because an integral of all of the moments of instant utility might make the proposition of repeating the experience somewhat unattractive. Hausman (2015) writes that he hardly remembers the tedious moments when caring for his children, but he vividly remembers the joy of reading to them and thus implicitly aligns well-being with fulfillment, not with moment-by-moment (un) happiness.

There are, of course, many within the economics profession who would contend that the focus of the gestalts, and the concern with how the moments of an experience fit together, can be encapsulated by adding arguments to the utility function (see, e.g., Binmore 2005). If one accepts this contention, then the gestalt characteristics, although challenging Benthamite utility maximization in retrospective and prospective evaluations, do not necessarily conflict with broader definitions of utility maximization. However, potentially placing anything and everything in the utility function leads to a somewhat empty theoretical framework, in that utility maximization can then be used to justify all actions and choices, and no specific predictions or policy advices are forthcoming. That is, whatever a person does can be attributed to maximizing utility, which, so the argument goes, is what he or she ought to do, so let's leave it at that.

It is, of course, possible that a focus on the gestalts maximizes overall utility for some people; for others, this focus, as well as other choices, actions, and behaviors that they undertake in life, could be driven by other considerations. They may trade off utility, for instance, to feel that their life has meaning or fulfillment of some kind. We cannot know what generally drives human decisions and behavior. People probably have various and varied legitimate reasons for their actions, both interpersonally and, across contexts, intrapersonally, and to assume that utility maximization is the appropriate universal normative requirement is, as proclaimed by Sugden (2018), a view from nowhere. Rather than utility maximization, we may thus conjecture that facilitating the drive for fulfillment, meaning, and a flourishing life as each individual sees fit is the appropriate normative goal. Consequently, the role of policy makers is to secure an environment that helps people to flourish—for them (i.e., the people) to pursue meaning and fulfillment to and in their lives as they see fit, subject to public resource constraints and distributional concerns.

Nourishing Flourishing

If the purpose of policy makers is to help create the conditions for people to flourish as they see fit, then encouraging reciprocity—in short, the act of returning a favor with a favor and a harm with punishment—as an aspect of behavioral public policy (in that it conflicts with the assumption embedded in rational choice theory that people are egoistically selfish in their pursuit of utility) is potentially an important arm of this effort. Reciprocity has evolved as a central social norm in all known cultures because it confers benefits on the group and, through an enlightened form of self-interest, to most of the individuals within each group. There are, of course, many possible negative consequences of reciprocity—for example, its potential to breed resentment, retaliation, retribution, cronyism, fundamentalism, nationalism, and the like—but with care, reciprocal actions can serve as a force for individuals and groups of people to more effectively reach their preferred destinations. Unfortunately, unless care is taken in societal and institutional design, reciprocity as a motivator of human behavior can be crowded out by our baser instincts.

There is a vast multidisciplinary literature on reciprocity, and this is starting to filter through to the public policy discourse, with recent books related to the topic published by, for example, David Sloan Wilson (2015) and Samuel Bowles (2016). A detailed exposition of the importance of reciprocity and how it might be nurtured by policy makers is offered elsewhere (Oliver 2019), but it may include encouraging those with policy influence to emphasize the importance of this motivator of human behavior in their rhetoric, decentralizing decision-making, and ensuring that income and wealth are not concentrated excessively in the hands of a small proportion of the population. Let us briefly consider these three general structural features.

On the importance of emphasizing reciprocity in the policy rhetoric, the underlying messages to "give so as to receive" and to "give to those who have given" are likely to be useful to policy makers, not only in gaining support for policies that they wish to introduce, but also to strengthen an aspect of human motivation that can underpin the collective good. If we take health or social care insurance, for example, we might emphasize to relatively young, healthy people that they contribute toward the costs of people in need of these services now so that they (and those they care about) will benefit in the future.

It is not wise to appeal to pure altruism to construct and sustain a welfare system, since pure altruism is not, for most people, a sustained source of motivation, and it is imprudent to appeal to egoism, because egoism, if encouraged, may thrive and harm. If policy makers wish to create new and protect old institutions, they might sensibly work hard to convince those who finance those systems—systems that substantially benefit "others"—benefit them also. "If some users get all of the benefits and pay few of the costs," writes Elinor Ostrom (2000, 150), "others become unwilling to follow rules over time." Moreover, by appealing to this notion of give-and-take, policy makers are likely to strengthen the social norm of reciprocal altruism, which bodes well for the more widespread collective benefits that a group may accrue by holding true to this substantively prosocial sentiment.

With respect to decentralizing decision-making, in *The Logic of Collective Action,* Mancur Olson (1965) argues that large groups will not be able to organize themselves voluntarily for coordinated and cooperative action, even if they have good reason for doing so. This is partly because the more people there are to share a collective benefit, the less each individual can singularly gain, making the return on cooperation less meaningful, and partly because large groups may have substantial organization costs that have to be subtracted from the expected benefits. Most importantly, Olson maintains that any rational egoist would choose to free ride on large group endeavors, and thus assuming, as he did, that selfish egoism is widespread, there would be insufficient effort within the group to generate the collective good. However, Olson contends that relatively small groups might be able to sustain cooperation, because if a person tries to free ride in a small group, it is more noticeable, and it is thus easier to identify and punish the culprit.

Although almost at the opposite end of the ideological spectrum to Olson, Ostrom (1990) concurs that prosocial actions are more likely to thrive at a smaller, more localized level. Ostrom's work provides some support for the above conjecture, where she finds that in common pool resource situations—for example, establishing local fishing rights—collectively desirable outcomes arise when users are left to develop the rules and enforcement mechanisms themselves. In short, Ostrom contends that communities often develop their own cooperative, reciprocal rules without enforcement from the central state and without imposing private property rights. She also believes that prosocial norms, which may have developed from prosocial actions that evolved organically, play an important role in tackling common pool resource challenges. For instance, when there are strong shared norms against opportunistic behavior, people will be less worried about the dangers of defection, and this can reduce the costs of monitoring and sanctioning activities.

Ostrom thus suggests that allowing regulation from the ground up, driven as it often is by mutual interest, may be more effective than trying to impose regulation from the top-down. Ostrom writes that

> It is possible that [people] learn whom to trust, what effects their actions will have on each other and on the CPR [common pool resource], and how to organize themselves to gain benefits and avoid harm. When individuals have lived in such situations for a substantial time and have developed shared norms and patterns of reciprocity, they possess social capital with which they can build institutional arrangements for resolving CPR dilemmas.… Public policies based on the notion that all CPR appropriators are helpless and must have rules imposed on them can destroy institutional capital that has been accumulated during years of experience in particular locations. (1990, 184)

Ostrom therefore contends that local communities will often be best equipped to self-regulate and manage their collective goods and services; she also emphasizes that the threat and application of a set of sanctions agreed by the local community on those who are tempted to or do transgress is a needed component of the regulations. In other words, negative reciprocity is required if cooperation within the group is to be sustained. Without the threat of punishment, selfish egoists, even if they are only small in number, may drive otherwise conditional cooperators to undertake less enlightened, short-sighted actions.

From the above, then, a plausible generally beneficial public policy lesson is for the organization, management, and financing of public sector services to be decentralized, perhaps subject to a nationally imposed minimum standard of service provision, with the caveat that the resources collected at the decentralized level would need to be risk-adjusted at the superregional (or national) level if equity of opportunity across all relevant groups is an objective. This is not to argue that decentralized groupings should be disconnected from each other; indeed, quite the contrary, if we attach importance to cross-regional learning. Nor is it to argue that decentralization is all that one requires to stir reciprocal actions. If one encourages egoism at the decentralized level, for example, then egoism is likely to prevail. It is to contend that decentralizing is a necessary, if not sufficient, condition for reciprocal cooperative arrangements to be given the best chance of thriving.

Finally, on income and wealth concentrations, there is evidence that the very rich in many countries have been awarding themselves indefensible increases in the shares of national income over recent decades. In this respect, reciprocity has perhaps gone awry. As an illustrative example, the share of national income in the United States that went to the richest 1 percent of the population increased from 8 percent in 1980 to 18 percent in 2015 (Pinker 2018); others have estimated the share of net wealth, as opposed to income, enjoyed by the top 1 percent at closer to 40 percent in the United States in 2014 and at 20 percent to 30 percent in an array of other countries, including Japan, France, Canada, Greece, Germany, Denmark, Sweden and the United Kingdom (Credit Suisse 2014). Even if more or less everyone in society saw their income and wealth increase over time, one can easily be made to feel poorer if others appear to have so much more. When a nation's income or wealth is being concentrated increasingly in the hands of a relatively small number of people, the majority, or at least a substantial minority, may have quite legitimate feelings that they are being left behind. A degree of inequality may be a driver of growth; it may be good for the group as a whole. But too much inequality may undermine motivation and other-regarding, group-oriented social norms and might crowd in selfish egoism and its consequent long-term harms.

Widening income inequality is not conducive to sustaining a trusting, reciprocal, cooperative society. As Sapolsky notes, "Trust requires reciprocity, and reciprocity requires equality, whereas hierarchy is about domination and asymmetry" (2017, 292). But we must remember that reciprocity is a two-way street. Redistributive mechanisms may be easier to sustain if those who are in receipt of welfare benefits, for instance, are, if they are able, seen to contribute to the fabric of society. For example, streets, parks, and rivers can be tidied, and loneliness, which is a significant problem, particularly among the elderly, can be alleviated. Tasks that help address these issues, and many others that might also serve to improve skills, could be performed on a part-time basis by many of those in receipt of welfare benefits.

Although it may be unpalatable to some, humans are not naturally pure altruists. If we have a situation in which the non-poor feel that the poor are doing little to alleviate their situation or are offering little in return for assistance, then the non-poor will perhaps feel less inclined to help them. It might well be that in such circumstances the working poor and those in the middle of the income distribution feel squeezed the most, seeing the very rich sail away from them and yet feeling that they are subsidizing those who choose welfare over work. To effectively sustain support for a public policy or an adequate system of welfare, it is wise to emphasize to those who are paying for it what they are getting out of it, and requiring those who benefit to offer something tangible in return (assuming that they are mentally and physically able to do so) would be one way of doing that.

Nurturing reciprocity through policy rhetoric, decentralizing decision-making and addressing excessive income and wealth concentrations maintains for most the great autonomy required for people to pursue their privately held conception of a flourishing life. With great autonomy, and without counteracting measures, there remains considerable opportunity for egoists to act on their instincts. The liberal economic tradition, at least as represented by Mill ([1859] 1972) and his famous harm principle, recognizes this unfortunate potential and allows the regulation of private actions if those actions are imposing harms—or negative externalities—on others. Mill did not, however, recognize explicitly that an individual or organization would use the behavioral influences such as present bias, loss aversion, and so on, to essentially, and often covertly, impose harms on others.

It can perhaps be contended that a key feature of the liberal economic tradition—the market—by fostering reciprocity, can render mute the biasing potential of the behavioral influences, but for more than a century, the private marketing industry has shown that this is not the case. George Akerlof and Robert Shiller, in their book *Phishing for Phools* (2015), reveal that in certain sectors, these activities remain rife, can undermine mutual self-interest, and hence can negatively affect the opportunity to flourish for at least one party to the exchange. Therefore, regulation against these behaviorally informed harms might be warranted.

Budging Phishing

Behaviorally informed regulation, applications of which are called "budges" (Oliver 2013, 2015, 2017), is an alternative conceptual behavioral public policy approach to the forms of paternalism discussed earlier. There are two main types of budges: (1) regulating against behaviorally informed activities that impose harms on others and (2) regulating for behaviorally informed activities that generate benefits for others. Both are focused on the effects experienced by those other than those who are targeted for behavior change and thus differ from the internalities perspective of the nudge approach.

As an example of the first type of budge, consider the marketing divisions of confectionary companies, which through long experience know that salience and immediacy (i.e., present-ness) can have a large effect on consumer buying patterns, and thus they have traditionally paid supermarkets to have their products displayed at child eye-level near checkout counters. If we conclude that consumers purchase more confectionary than they otherwise would—and indeed, more than is good for them (and their children)—as a consequence, then policy makers would have an intellectual justification for regulating against this activity. Indeed, some policy makers appear to be embracing this approach, at least implicitly. In the summer of 2018, for instance, the Secretary of State for Health in England announced plans to ban the sale of confectionary and fatty snacks near supermarket checkout counters, and to introduce tighter regulations on television and online advertising of junk food.

In relation to the second type of budge, consider human inertia—so-called status quo bias—and its potential to sometimes prevent others from realizing benefits because of lack of action by a first party. This may occur, for example, when people do not opt in to be organ donors in countries that require such action, not because they do not want to but because they do not get around to doing so. The budge here, then, may be to regulate openly so as to move the country to either presumed consent or prompted choice—such that people have to choose whether to register or not when renewing their vehicle tax—for organ donor registration.

Budges, like nudges, can be placed in a three-dimensional space, as depicted in figure 2. Compared with figure 1, the vertical and horizontal axes have been inverted, such that a policy placed at the origin would now be a regulation, informed by behavioral economics, to tackle an externality. That would be a pure budge.

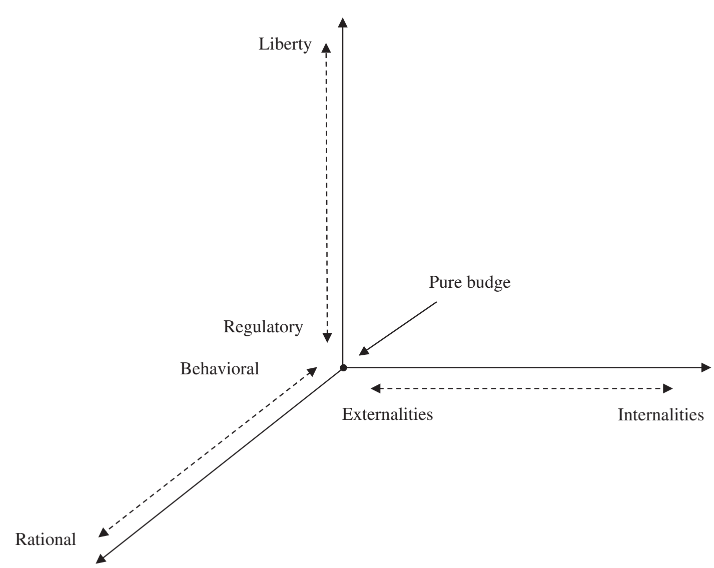

Figure 2 The Requirements of Behavioral Regulation

Countless examples of potentially harm-reducing and benefit-enhancing budges could be offered here, but space constraints allow for only a couple of illustrative examples. Resembling the issue of putting confectionary at checkout counters, supermarkets are often financially induced to place alcoholic beverages at the end of shopping aisles, since it is known that this salient positioning increases sales. Indeed, it has been estimated that this positioning increases alcohol sales by as much as would be realized by a price reduction of 4 percent to 9 percent (Nakamura et al. 2014). If it was broadly concluded that the end-of-aisle placement of alcoholic products was manipulating people's choices such that the egoistic self-interest of supermarkets and the drinks industry was imposing harms (by provoking too much alcohol purchasing and consuming) on consumers vis-à-vis the consumers' perceptions of what it means to live a flourishing life—that is, that it had damaged a mutually beneficial reciprocal exchange—then there would be grounds for regulating against this practice. Such a regulation would be an explicit behaviorally informed intervention to lessen potential harms—in other words, it would be a budge.

As of this writing, public authorities in England appear to be implicitly embracing behavioral regulation with respect to the gambling industry, an aspect of human activity that is ripe for exploiting behavioral affects. Indeed, people may sometimes gamble almost entirely because of the behavioral influences, in that they may overweight the probability of winning and anchor on the jackpot. Of course, responsible gambling may enrich the lives of many people, but the exploitation of the behavioral influences has the potential to inflict serious harms with respect to many conceptions of a flourishing life, particularly for financially vulnerable people. Fixed-odds betting terminals (e.g., roulette machines) that are readily accessible in betting shops on high streets are one way in which the gambling industry profits from its knowledge of these influences. In England, these machines are programmed to allow people to gamble large amounts of money—up to £100 every 20 seconds. The government has woken up to the harms that this can cause and has announced that in 2019 it will introduce regulations to limit the stake (in each 20-second period) that the machines will be able to accept to £2, a proposal that provoked considerable opposition from the gambling industry.

However, regulatory authorities might not always have to do battle with industry for the latter to limit (or reverse) any harms that they impose. In the spirit of reciprocity, it may sometimes be possible for public policy makers to work with those they might otherwise budge, by emphasizing that there may be reputational, and hence commercial, benefits for those who agree more readily to alter their practices, although history suggests that extending this goodwill all the way toward agreeing to industry self-regulation might best be avoided.

Conclusion

To date, the principal overarching normative goal of behavioral public policy in the academic and public policy discourses has been a paternalistic focus on moving internalities toward some conception of utility maximization, exemplified by the nudge approach. It is postulated in this article that in the spirit of the liberal economic tradition, or at least that associated with John Stuart Mill, there is a more appropriate normative path for behavioral public policy to follow. This relatively new public policy approach should seek to foster reciprocity and cooperation so that people are better equipped to flourish and give meaning and fulfillment to their lives in the ways that they see fit—which can include, but is not limited to, seeking personal happiness.

However, the approach proposed here goes further than liberal economics typically does by recognizing that allowing people great freedom to seek their own goals potentially affords more opportunities for the egoistically inclined to use the behavioral influences to manipulate, and impose harms on, others. Therefore, it is suggested that when these behavioral influences, including present bias, loss aversion, probability weighting, and even reciprocity (etc.), are being used toward self-serving, externally harmful ends, then there is a legitimate intellectual justification for regulating against those harms—in short, to budge against these activities. It also suggests that the actions (or lack thereof) of those who are subject to behavioral affects, such as status quo bias, who as a consequence deprive others of substantial potential benefits, can be regulated to realize these benefits, provided that the regulation imposes no substantial costs on those who are regulated.

Thus, in one sentence, the behavioral public policy framework proposed here is for policy makers to secure a behaviorally informed environment that is conducive to helping people to flourish on their own terms and to regulate against and for behaviorally informed externalities. No claim is made that this approach will liberate all untapped potential or solve all of society's ills—no approach to behavioral public policy (or indeed any public policy) can do that. But in facilitating flourishing and impacting upon externalities, this, to me, offers the most promising political economy of behavioral public policy.

Notes

1. The approach that I present in this article is closer to that offered by Mill than that postulated by the Austrian School of liberal economics because, like Mill (only, as stated, more so), I attach particular emphasis to the legitimacy of government regulation against harms. Leaders of the Austrian School, such as Ludwig von Mises ([1927] 2005), disparaged Mill's interventionist approach, labeling him a socialist, although Friedrich Hayek ([1944] 2001) appeared to have a healthy respect for Mill.
2. The references given on the gestalt characteristics are indicative. They do not, of course, prove that the gestalts influence behavior substantively in all circumstances, and there may be an element of publication bias, in that the authors of studies that do not demonstrate these effects might find it difficult to publish their findings. Moreover, although there is additional evidence on these characteristics over and above that which is cited in this article, further studies are, of course, needed to more accurately assess the reach of the gestalts; however, there is sufficient evidence reported on the gestalts to lend force to the belief that they will often have an impact on people's retrospective and prospective assessments of experiences.
3. Sigmoidoscopy is a procedure in which a flexible tube is inserted in a patient's rectum to examine the colon; lithotripsy breaks up kidney stones into small pieces using ultrasound waves.
4. Sugden uses the term "the view from nowhere" pejoratively to indicate that the notion that the same normative standard should be imposed on all people is contestable. If one allows people to pursue the lives that they want to live as they see fit in their private actions—that is, to follow their own conception of a flourishing or meaningful life, which allows them to pursue happiness or even a life that others may think lacks meaning if they so wish—then no view is

imposed on them. There is no omnipresent view from nowhere. However, in public sectors—such as health, education, and so on—there will be many broadly agreed-upon goals, although these are still not imposed from nowhere as they have been decided upon following a period of open explicit debate; these public sector objectives (e.g., good health, knowledge, etc.) are partly intended to help each individual in his or her private pursuit of a flourishing life.

5. As a fundamental motivator of human behavior, reciprocity can also be used to inform the design of interventions that fall under behavioral public policy frameworks that differ from the one that I am proposing here, including libertarian paternalism.
6. It is important to keep in mind that a budge is a behaviorally informed regulation to address a harm or an otherwise forgone benefit. Nudges are non-regulatory behavioral informed interventions that address internalities. Moving from opting in to presumed consent or prompted choice for organ donation more closely fits the criteria of the former than the latter, yet these interventions are frequently mischaracterized as nudges in the literature. Oliver (2015) highlights that many interventions that have been proposed and/or applied as nudges are in fact more closely aligned with the parameters of budging.
7. In the economics literature, externalities refer to harms and benefits experienced by third parties to an exchange. In budge policy, the externalities can refer to the same (so long as the externalities are caused by actions influenced by behavioral phenomena) but are also extended to harms and forgone benefits that a first party can impose on a second party through a behaviorally influenced action or lack of action.
8. There may occasionally be circumstances in which it is legitimate to regulate for a behavioral informed measure to secure benefits for others even when the intervention imposes substantial costs on the regulated party, particularly if the regulated party sells products that are associated with harms as well as benefits. For example, a government may require food manufacturers to place behavioral informed messages and symbols on its products in an attempt to facilitate healthier food purchases by consumers, even if the labeling is costly to the industry.

References

Akerlof, George A., and Robert J. Shiller. 2015. *Phishing for Phools: The Economics of Manipulation and Deception*. Princeton, NJ: Princeton University Press.

Aristotle. [350 BCE] 1998. *The Nicomachean Ethics*. Oxford: Oxford University Press.

Bentham, Jeremy. [1781] 1988. *The Principles of Morals and Legislation*. New York: Prometheus Books.

Binmore, Ken. 2005. Economic Man—Or Straw Man? *Behavioral and Brain Sciences* 28(6): 817–18.

Bowles, Samuel. 2016. *The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens*. New Haven, CT: Yale University Press.

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. 2003. Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism." *University of Pennsylvania Law Review* 151: 1211–54.

Camerer, Colin F., and George Loewenstein. 2003. Behavioral Economics: Past, Present, Future. In *Advances in Behavioral Economics,* edited by Colin F. Camerer, George Loewenstein, and Matthew Rabin, 3–52. Princeton, NJ: Princeton University Press.

Conly, Sarah. 2013. *Against Autonomy: Justifying Coercive Paternalism*. Cambridge: Cambridge University Press.

Credit Suisse. 2014. *Global Wealth Databook 2014*. Zurich: Credit Suisse Research Institute. https://www.credit-suisse.com/corporate/en/research/research-institute/global-wealth-report.html [accessed July 8, 2019].

Fredrickson, Barbara L., and Daniel Kahneman. 1993. Duration Neglect in Retrospective Evaluations of Affective Episodes. *Journal of Personality and Social Psychology* 65(1): 44–55.

Gigerenzer, Gerd. 2015. On the Supposed Evidence for Libertarian Paternalism. *Review of Philosophy and Psychology* 6(3): 361–83.

Hausman, Daniel M. 2015. *Valuing Health: Well-Being, Freedom and Suffering*. Oxford: Oxford University Press.

Hayek, Friedrich A. [1944] 2001. *The Road to Serfdom*. London: Routledge Classics.

Hertwig, Ralph. 2017. When to Consider Boosting: Some Rules for Policy-Makers. *Behavioural Public Policy* 1(2): 143–61.

Hsee, Christopher K., and Robert P. Abelson. 1991. Velocity Relation: Satisfaction as a Function of the First Derivative of Outcome over Time. *Journal of Personality and Social Psychology* 60(3): 341–47.

Hume, David. [1777] 1983. *An Enquiry Concerning the Principles of Morals*. Indianapolis: Hackett Publishing Company.

John, Peter, Sarah Cotterill, Alice Moseley, Liz Richardson, Graham Smith, Gerry Stoker, and Corinne Wales. 2011. *Nudge, Nudge, Think, Think: Experimenting with Ways to Change Civic Behaviour*. London: Bloomsbury Academic.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. Back to Bentham? Explorations of Expected Utility. *Quarterly Journal of Economics* 112(2): 375–405.

Layard, Richard. 2005. *Happiness: Lessons From a New Science*. London: Allen Lane.

Loewenstein, George F., and Dražen Prelec. 1993. Preferences for Sequences of Outcomes. *Psychological Review* 100(1): 91–98.

Mill, John Stuart. [1859] 1972. *On Liberty*. London: Dent.

Nakamura, Ryota, Rachel Pechey, Marc Suhrcke, Susan A. Jebb, and Theresa M. Marteau. 2014. Sales Impact of Displaying Alcoholic and Non-alcoholic Beverages in End-of-Aisle Locations: An Observational Study. *Social Science & Medicine* 108: 68–73.

Oliver, Adam. 2013. From Nudging to Budging: Using Behavioural Economics to Inform Public Sector Policy. *Journal of Social Policy* 42(4): 685–700.

———. 2015. Nudging, Shoving and Budging: Behavioural Economic-Informed Policy. *Public Administration* 93(3): 700–714.

———. 2017. *The Origins of Behavioural Public Policy*. Cambridge: Cambridge University Press.

———. 2019. *Reciprocity and the Art of Behavioural Public Policy*. Cambridge: Cambridge University Press.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

———. 2000. Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives* 14(3): 137–58.

Pinker, Steven. 2018. *Enlightenment Now: The Case for Reason, Science, Humanism and Progress*. London: Allen Lane.

Redelmeier, Donald A., and Daniel Kahneman. 1996. Patient's Memories of Painful Medical Treatments—Real-Time and Retrospective Evaluations of Two Minimally Invasive Procedures. *Pain* 66(1): 3–8.

Redelmeier, Donald A., Joel Piano Katz, and Daniel Kahneman. 2003. Memories of Colonoscopy: A Randomized Trial. *Pain* 104(1–2): 187–194.

Sapolsky, Robert. 2017. *Behave: The Biology of Humans at Our Best and Worst*. London: Bodley Head.

Sen, Amartya. 1999. *Development as Freedom*. New York: Random House.

Sugden, Robert. 2018. *The Community of Advantage: A Behavioural Economist's Defence of the Market*. Oxford: Oxford University Press.

Thaler, Richard H. 2015. *Misbehaving: The Making of Behavioral Economics*. New York: Penguin Random House.

Thaler, Richard H., and Cass R. Sunstein. 2003. Libertarian Paternalism. *American Economic Review* 93(2): 175–79.

———. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

von Mises, Ludwig. [1927] 2005. *Liberalism*. Indianapolis, IN: Liberty Fund.

Wilson, David Sloane. 2015. *Does Altruism Exist? Culture, Genes, and the Welfare of Others*. New Haven, CT: Yale University Press.