

# BEHAVIOURAL ECONOMICS AND PATERNALISM

DANIEL M. HAUSMAN\*

---

**Abstract:** Contemporary behavioural economics has documented common failures of reasoning that apparently make possible policies that benefit individuals by contravening or correcting their judgements. These policies appear to be paternalistic, even though a traditional view would deny that they are paternalistic on the grounds that policies such as nudges do not restrict individual liberty. It appears to many that a new definition of paternalism that takes its cue from behavioural economics is needed. Furthermore, if one revises the definition of paternalism, one must revisit traditional views concerning whether paternalist policies are wise. In *Government Paternalism*, Julian Le Grand and Bill New make a valuable start, although some corrections are needed. This essay provides a general setting for revising the definition of paternalism and suggests a revised definition of paternalism.

**Key Words:** paternalism, nudge, behavioural economics, liberty

Standard welfare economics takes people's preferences to indicate, if not to constitute, people's well-being or welfare (which I shall treat as synonymous). It is obvious that, owing at least to false beliefs, mistakes, and concerns for others, people may prefer what is worse for them over what is better for them. Nevertheless, an optimistic welfare economist might hope that these complications are unimportant and that it is reasonable, as an empirically tractable first approximation, to measure well-being by preference satisfaction.

One advantage of taking preference satisfaction to indicate welfare is that all questions about paternalism are mooted. If what Charlie prefers is

\* Department of Philosophy, University of Wisconsin-Madison, 5197 Helen C. White Hall, 600 N. Park Street, Madison, WI 53706-1474. Email: [dhausman@wisc.edu](mailto:dhausman@wisc.edu). URL: <http://philosophy.wisc.edu/hausman>

best for him, then it is impossible to make him better off by going against his preferences. In substituting your judgement for his with the aim of making him better off, you are guaranteed to fail. It is, of course, absurd to maintain in this way that paternalism is impossible, that the interference of others can never make Charlie better off. Tell that to Charlie after a large stranger pulls him out of the road just before a truck that Charlie did not see runs Charlie over. Although it is obviously mistaken to maintain that paternalism is impossible (as opposed to unwise), economists have found this falsehood to be a convenient simplification that silences any questions about paternalism.

Behavioural economics has disrupted this agreeable state of affairs by documenting both the extent to which individual's preferences are governed by factors irrelevant to an agent's well-being and the systematic biases that result. Thus, the unwelcome question of whether paternalistic policies might be well-advised cannot be evaded, and having disinterred this question, a deeper one arises: just what is this newly reanimated zombie called 'paternalism'?

In this paper I argue that behavioural economics helps reveal the inadequacy of views of paternalism that stipulate that paternalism restricts individual liberty. Moreover, behavioural economics can play a role in justifying paternalism. But I argue against relying on behavioural economics to define paternalism. [Section 1](#) discusses John Stuart Mill's views, in which concerns about the concept of paternalism are, I argue, secondary. [Section 2](#) briefly discusses apparent exceptions Mill countenances to his general rejection of paternalism and the distinction between hard and soft paternalism. [Section 3](#) considers the possibility that the Millian view that paternalism necessarily limits liberty is misconceived, and [Section 4](#) focuses on an important recent book by Julian Le Grand and Bill New, which offers a view of paternalism whose definition and justification is informed by the findings of behavioural economics. [Section 5](#) points out difficulties with Le Grand and New's view, and [Section 6](#) concludes with a defence of a minimalist definition of paternalism that serves the purposes of behavioural economists.

## 1. A MILLIAN VIEW OF PATERNALISM

John Stuart Mill's *On Liberty* is often read as the classic text on paternalism, but its comments on paternalism are part of its argument for individual liberty rather than an analysis or theory of paternalism. Mill does not use the term 'paternalism' in *On Liberty* or in any of his other writings.<sup>1</sup> The closest he gets terminologically is when he speaks of 'the system of

<sup>1</sup> According to the Oxford English Dictionary, the term 'paternalism' first appears in *Appleton's Journal* in 1873. It initially had a positive connotation.

despotic, or what is called paternal, government' (1978 [1859]: 101). As this phrase suggests, Mill is concerned about the power that government and social norms exert over individuals, and he objects to the use of that power except where directed toward the protection of other people. He thus condemns restrictions on a person's freedom either for the person's own good or because the person's private behaviour (such as his own unusual relationship with Harriet Taylor) offends others. There are other social objectives that Mill does not mention, such as avoiding cruelty to animals, preserving national monuments, or protecting places of natural beauty and interest. As Mill would, I believe, concede, these may also justify limiting liberty. His concern is to defend liberty from paternalistic and moralistic restrictions. Whether a society restricts homosexual behaviour on the grounds that it is condemned by the Bible, or whether it does so on the (spurious) grounds that homosexual relations are harmful to the individuals involved, what is of concern to Mill is the limitation on individual freedom. Mill is not concerned about whether there might be paternalistic actions that do not limit liberty. Instead of offering a theory of what we now call 'paternalism', Mill condemns intrusive limits on liberty that have either a moralistic or a paternalist basis.

Paternalism might turn out to be nothing other than the paternalistic coercion Mill was concerned about. For example, Gerald Dworkin writes (1972: 65) 'By paternalism I shall understand roughly the interference with a person's liberty of action justified by reasons referring exclusively to the welfare, good, happiness, needs, interest or values of the person being coerced'. Richard Arneson agrees that paternalistic policies restrict liberty with the objective of benefitting the individual or individuals whose liberty is restricted and adds that to count as paternalistic, the policy must be carried out against the will of these individuals (1980: 471). In his recent (2016) revision of the 'Paternalism' entry in the *Stanford Encyclopedia of Philosophy*, Dworkin requires that to count as paternalistic an action or omission must interfere with the liberty or *autonomy* of the person whom the action aims to benefit. Depending on how one understands autonomy, this condition, despite harkening back to Mill, could represent a retreat from the view that paternalism limits *liberty*. The Millian view is roughly that paternalist actions directed to Charlie limit Charlie's liberty or autonomy against his will with the objective of benefitting him.

To simplify the discussion, I shall focus on government paternalism, which is bound to be more problematic, because government knows so much less about what is good for me than do my friends and family. However, much of the discussion applies equally to actions of individuals and of non-governmental groups. Limits to liberty may be implemented via constraints, punishments, costs, or deception, including withholding information. Confining someone who is suicidal so as to make it impossible for her to harm herself exemplifies physical

constraint. Fining drivers who do not use seat belts limits liberty via punishment. Taxing harmful products coerces via costs. Informing would-be masturbators that they risk blindness limits liberty by means of deception, unless the homophobes spreading this false information are themselves misinformed, in which case conveying this falsehood would constitute a misguided effort at rational persuasion.

Millian paternalist policies that coerce individuals for their own benefit often have other consequences and other justifications. For example, requiring motor cyclists to wear helmets may reduce health-care costs that are borne by others, and taxing tobacco raises revenue. In some cases, the non-paternalistic grounds for a policy are far-fetched, while in others it may be uncertain whether the policy is truly paternalistic or merely happens to benefit those whom it coerces. One might suggest a counterfactual test. One possibility is to maintain that a policy is paternalistic only if it would not have been carried out if the policymaker had not believed that it would benefit the individuals whose freedom it limits. This test seems too restrictive. I suggest instead that a policy is paternalistic only if it would still have been carried out if there were no non-paternalistic reasons for it.<sup>2</sup>

## 2. EXCEPTIONS AND 'SOFT' PATERNALISM

Although Mill in *On Liberty* apparently rules out all legal and social limits on freedom that are not grounded in the protection of others, he makes two much-discussed exceptions. He believes that contracts whereby an individual agrees to be a slave, no matter how freely they are entered into, should not be enforced. In that way, we protect the freedom of the individual who is attempting to surrender it (1978 [1859]: 101). Although preventing someone from selling themselves into slavery protects their future freedom, it limits their current freedom and appears to be an instance of paternalism that Mill approves of.

Second, Mill discusses the following case:

If either a public officer or anyone else saw a person attempting to cross a bridge which had been ascertained to be unsafe, and there were no time to warn him of his danger, they might seize him and turn him back without any real infringement of his liberty; for liberty consists in doing what one desires, and he does not desire to fall into the river. Nevertheless, when there is not a certainty, but only a danger of mischief, no one but the person himself can judge of the sufficiency of the motive which may prompt him to incur the risk: in this case, therefore, (unless he is a child, or delirious, or in some state of excitement or absorption incompatible with the full use of the reflecting

<sup>2</sup> I am indebted here to Philippe Mongin.

faculty,) he ought, I conceive, to be only warned of the danger; not forcibly prevented from exposing himself to it. (1978 [1859]: 95)

Mill claims that the limitation on liberty is only apparent, not real, on the grounds that the interference conforms to the true desires of the bridge-crosser. Joel Feinberg and others have interpreted Mill as maintaining that interfering with an action that an agent did not *voluntarily* undertake is either not paternalistic at all (Feinberg 1971, esp. p. 112; 1986: 15) or a different defensible 'soft' variety, in which coercion isn't really coercion when it only prevents someone from doing something involuntarily. As Feinberg puts it,

This seems to lead us to a form of paternalism that is so weak and innocuous that it could be accepted even by Mill, namely that the state has the right to prevent self-regarding harmful conduct only when it is substantially nonvoluntary or when temporary intervention is necessary to establish whether it is voluntary or not. (Feinberg 1971: 113)

Arneson copes with cases such as these by maintaining that even though the crosser is 'seized' and 'turned back' these actions are not against the crosser's will (or are not known to be against his will) (Arneson 1980: 487). If preventing an action is against someone's will if and only if the action is voluntary, then soft-paternalistic actions turn out not to be paternalist at all. On this view, 'hard' or genuine paternalism interferes only with actions that the policymaker knows to be voluntary.<sup>3</sup>

### 3. PATERNALISTIC POLICIES THAT DO NOT LIMIT LIBERTY

The term, 'paternalism' has only been in existence for a century and a half, and during that time it has changed its connotation from positive to negative. Given Mill's influence, it has been associated with limits on liberty (for example, Feinberg 1986: 7). On the other hand, some contemporary philosophers (for example Shiffrin (2000) and Le Grand and New (2014)) maintain that paternalism need not restrict liberty. How should one decide whether they are right? Is the task to uncover the 'essence' of paternalism? Shiffrin suggests that it is not enough to test definitions against intuitions (whose authority over the usage of philosophical jargon is, I suggest, questionable). In addition, she argues that 'it is worthwhile to assess what is central in our normative reactions to paternalism and to employ a conception of paternalism that complements

<sup>3</sup> Note that Mill writes, 'Nevertheless, when there is not a certainty, but only a danger of mischief, no one but the person himself can judge of the sufficiency of the motive which may prompt him to incur the risk'. This suggests that when there is a certainty of injury or death then it is permissible to limit the crosser's liberty. I am not sure whether this is a careless phrase or whether Mill would extend his approval of preventing individuals from selling themselves into slavery to preventing individuals from heedlessly sacrificing their lives or health.

and makes intelligible our sense of paternalism's normative significance' (Shiffrin 2000: 212).

The 'normative significance' of paternalism she has in mind is whether the government (as our agent) or we as individuals treat individuals with proper respect.

The essential motive behind a paternalist act evinces a failure to respect either the capacity of the agent to judge, the capacity of the agent to act, or the propriety of the agent's exerting control over a sphere that is legitimately her domain... Paternalistic behavior is special because it represents a positive effort by another to insert her will and have it exert control merely because of its (perhaps only alleged) superiority. As such, it directly expresses insufficient respect for the underlying valuable capacities, powers, and entitlements of the autonomous agent. (Shiffrin 2000: 220)

From such a perspective, as Shiffrin points out, refusals to aid as well as intrusive actions are sometimes intuitively paternalistic. For example, in refusing to help my son with his math homework, I am not violating his rights or limiting his liberty, but if I think he would be better off working out the difficulties on his own, I am attempting to benefit him by acting against his will.<sup>4</sup> Should we count this as paternalistic on the grounds of intuition, its partial satisfaction of the conditions in the Millian view, and the disrespect it shows toward my son's autonomy, or should we deny that it is paternalistic, on the grounds that there is no restriction on liberty. If what matters to us about paternalism are questions about the justification for limiting liberty, then we might deny that my refusal to help my son is paternalistic. If, on the other hand, we find that 'what is distinctive and worrisome about paternalism' is 'this substitution of judgment and the circumvention of an agent's will' (Shiffrin 2000: 213), then we should seek a definition of paternalism that counts my refusal to help my son as paternalistic. As Dworkin puts it, 'What we must ascertain in each case is whether the act in question constitutes an attempt to substitute one person's judgment for another's, to promote the latter's benefit' (1988: 123).

What should such a definition of paternalism say? Shiffrin argues that it need not require that the action be contrary to the will of the person whom it aims to benefit. Suppose, for example, that I want to protect my elderly father from the many fraudulent offers that would probably arrive in his email in-box daily. So I install spam filters on his computer, knowing that he has no idea how to access his 'junk' mail.<sup>5</sup> My action is not contrary to his will, but it is independent of it and circumvents it. I am substituting my judgement for his.

<sup>4</sup> This is not Shiffrin's example, but it has the same structure as the one she gives on p. 213.

<sup>5</sup> This is again not Shiffrin's example, but it has the same structure as one that she gives on p. 214.

In the light of examples such as these, Shiffrin takes it as essential to paternalism that ‘the motive behind [the paternalistic action] is simply that the (putative) paternalist knows better than the agent, or may better implement, what the agent has authority for doing herself’ (Shiffrin 2000: 216).<sup>6</sup> As this statement suggests, Shiffrin counts an action as paternalist if it is directed toward assisting an agent to pursue an objective, even if achieving the objective does not contribute to the agent’s well-being. Since she is concerned about the substitution of judgement and circumvention of an agent’s will, it is not, from her perspective, important whether the intervention is directed toward what benefits the agent or toward the achievement of some other aim of the individual.

But if one is concerned about the limits on government action and one takes for granted that government has an obligation to contribute to the well-being of its citizens, not to the accomplishment of projects unrelated to their well-being, then one needs to modify Shiffrin’s conception of paternalism. Some economists have maintained that welfare economics should be focused on enabling people to achieve their objectives rather than on promoting well-being (Sugden 2004; Vining and Weimer 2010), but for the purposes of this paper, I shall assume the more conventional view that the central goal of normative economics is to enable policy to enhance individual’s welfare.

#### 4. PATERNALISM FOR BEHAVIOURAL ECONOMICS

Although Le Grand and New follow Shiffrin in jettisoning the requirement that paternalist policies restrict liberty and in emphasizing the substitution of judgement, they insist that paternalist policies be directed toward the benefit of those whose judgement they challenge. Here is their definition:

So to our definition. We conclude that a government intervention is paternalistic with respect to an individual if it is intended to

<sup>6</sup> Here is Shiffrin’s definition of paternalism. I shall not discuss its details, which are not germane to this essay’s concern for the connections between paternalism and behavioural economics. Some of the criticisms I shall later make of Le Grand and New’s definition of paternalism apply to Shiffrin’s definition:

So, I suggest that paternalism by A toward B may be characterized as behavior (whether through action or through omission)

- (a) aimed to have (or to avoid) an effect on B or her sphere of legitimate agency
- (b) that involves the substitution of A’s judgment or agency for B’s
- (c) directed at B’s own interests or matters that legitimately lie within B’s control
- (d) undertaken on the grounds that compared to B’s judgment or agency with respect to those interests or other matters, A regards her judgment or agency to be (or as likely to be), in some respect, superior to B’s (Shiffrin 2000: 218).

- address a failure of judgment by that individual
- further the individual's own good. (Le Grand and New 2014: 23)

Although this only states a sufficient condition, they mean the condition to be both necessary and sufficient.<sup>7</sup> The definition concerns only government interventions, because government actions are Le Grand's and New's concern; and it restricts paternalistic policies to those directed toward the benefit of the individual whose judgement fails. Le Grand and New require that there be a 'failure of judgement' on the part of the agent toward whom the paternalistic action is directed, while Shiffrin requires only that the would-be paternalist knows better (or thinks she knows better). By 'failures of judgement' I take Le Grand and New to refer to the foibles that behavioural economics investigates. In support of this interpretation, notice that they deny that mistakes based on false beliefs count as failures of judgement. In apparent contrast to Le Grand and New, Shiffrin allows that a paternalist intervention may respond to people's inability to do what they correctly judge to be best for themselves, rather than to a failure of judgement. In fact, Le Grand and New agree that paternalistic policy may aim to correct weakness of will, but they regard weakness of will as one kind of failure of reasoning or judgement.

For Shiffrin any advantage the would-be paternalist has over the individual subject to the intervention in terms of knowledge or ability to put that knowledge to work could serve to justify a paternalist policy that aims to facilitate the individual's projects. Le Grand and New, in contrast, insist repeatedly that interventions are paternalist only if the subject whom they aim to benefit manifests a 'failure of judgement'. So, for example, Shiffrin's definition would classify as paternalist the policy of an atheist government that closes the nation's monasteries to improve the lives of those who have up to now been monks. The rulers believe that they know better than the monks what makes for a good life. By closing

<sup>7</sup> Interpreting these conditions as merely providing a sufficient condition for government paternalism answers many of the criticisms I shall present later. But Le Grand and New claim explicitly that they are offering a definition here, and they offer no other definition of paternalism. Just before the quotation, they write, 'We are now in a position to summarize the argument so far, and to provide the two conditions that define paternalistic interventions by the government' (p. 22). Earlier in Chapter 2 they write, 'We will discuss many of these complexities before offering a simple definition that nevertheless captures what we see as the essence of paternalism' (p. 8). Their objective in Chapter 2 is explicitly to assess definitions of paternalism and to offer their own. 'It [this book] begins with issues of definition. Before discussing possible justifications for paternalism, it is necessary to specify what is meant by the term' (pp. 1–2). In anticipating their conclusions in Chapter 2, they write in the introduction, 'It therefore seems preferable to define paternalism, not in terms of the intervention itself or of its consequences, but in terms of (a failure of) individual judgement' (p. 2).



the monasteries, the government aims to 'take over or control what is properly within the agent's own legitimate domain of judgment or action' (Shiffrin 2000: 216). On Le Grand and New's definition, in contrast, if the commitment to a monastic life is not due to a failure in judgement but to a difference in values, then the closing of the monasteries would not only be unjustified: it would not be paternalistic. Although an implication of their definition, I am confident that this is not what Le Grand and New intend, because they regard policies that limit people's behaviour on the grounds that their ultimate objectives are unworthy as unjustified paternalism.

In addition, people often make bad choices because of false beliefs. Whether or not there may have been reasoning failures in the acquisition of those beliefs, reasoning *from* false beliefs may be impeccable. 'Having little information *in itself* does not impair our ability to reason' (Le Grand and New 2014: 15). Le Grand and New go on to point out, correctly, that providing information to correct a false belief does not substitute the policymaker's judgement for the agent's judgement. As Mill himself argued, engaging in rational discussion with an individual is not paternalistic; it neither limits liberty nor shows disrespect for the individual's judgement. Interventions that merely supply information are thus ordinarily not paternalistic, and they are, let us suppose, the ideal way to benefit individuals who are doing a poor job of deciding for themselves, owing to their false beliefs.<sup>8</sup> But there are many other ways to intervene in people's lives to save them from making bad choices that result from their false beliefs. And there seems to be no good reason to maintain, as Le Grand and New's definition implies, that an intervention cannot be paternalistic if it is designed to benefit an individual who without any failure of judgement has embarked on a harmful course of action owing to a false belief.

The notion of a 'failure in judgement' or a 'reasoning failure' is crucial to Le Grand and New's view of paternalism, and it is this notion that ties their work to recent advances in psychology and behavioural economics. Their crucial insight is that if an intervention aims to correct what behavioural economics identifies as a reasoning failure (in order to benefit the person whose reasoning fails), then there is a case to be made in defence of the intervention. Justified paternalistic interventions depend on demonstrable reasoning failures rather than disagreements between 'big brother' and individual citizens concerning what objectives are best.

However, they get into trouble by building judgement failures into the definition of paternalism. It is a tempting move to make, because it

<sup>8</sup> If individuals specifically ask not to be given certain information, then matters become more complicated. This complication strengthens the case below for the claim that paternalism does not only aim to correct failures of judgement.

enables them to give a persuasive account of the sense in which so-called 'libertarian paternalism' or nudging directed toward the benefit of those nudged is truly paternalism. Even though libertarian paternalism that avoids deception does not limit liberty and, on Mill's view, would fail to be paternalism at all, it is, on Le Grand and New's definition paternalist. Directed as it often is toward benefitting individuals who suffer from a reasoning failure,<sup>9</sup> it is the very paradigm of paternalist policies.

Le Grand and New (2014: 83f) distinguish four types of reasoning failure: limited technical ability, limited imagination or experience, limited objectivity, and limited will power, the last of which, as mentioned before, is arguably a failure to carry out the conclusions of one's reasoning than a failure of reasoning itself. Because reasoning failures can be objectively established, the boundaries of justified paternalist policies can be drawn clearly. Although the state is substituting its judgement for the judgement of individuals, it escapes Shiffrin's charge (2000: 220) that 'it directly expresses insufficient respect for the underlying valuable capacities, powers, and entitlements of the autonomous agent', because it purports only to correct demonstrable reasoning failures. It is thus unlike policies that aim to discourage or prevent pursuit of individual goals that policymakers regard as harmful to their possessors. The reasoning failures that behavioural economists identify are not, of course, badges of honour; but these failures to live up to a rational ideal are pervasive human foibles rather than grounds for contempt and disrespect. Intervening to contravene the pursuit of some end, in contrast, is arrogant and unjustified. 'There is no objective means of establishing an ends failure, and thus passing judgment on it should not form a part of any government's paternalistic policies' (Le Grand and New 2014: 103).

What distinguishes reasoning failures and makes them a defensible basis for state policy? Le Grand and New write:

To distinguish between, on the one hand, substantive failures of reason and, on the other, divergences from the classical model that are trivial or occasional, we draw on experimental studies, survey evidence, and long-running philosophical inquiry to identify where there is likely to be a genuine and systematic bias in decision making. Put another way, we try to identify situations where an individual would agree to the following statements when she was confronted with a failure of reasoning which she had supposedly committed. Her decision was a non-trivial error that she would probably repeat in similar circumstances; the error was conceptual,

<sup>9</sup> But do the choices of those whom we nudge for their own benefit always derive from a failure of judgement or reasoning. Are those who choose whatever dessert is prominent and appealing in the cafeteria, whether the food be cake or fruit showing a failure of judgement? And if they are, can government know which arrangement would nudge them in the direction of their own objectives?

not merely a verbal or technical misunderstanding; and she *should* have known the correct answer or procedure to find it. (Le Grand and New 2014: 82)

Should one accept the necessary condition on paternalist interventions that this account suggests? What bearing does a failure to satisfy these conditions have on whether a policy counts as paternalistic? For example, suppose that it is not the case that Cindy should have known 'the correct answer or procedure to find it'. Why should one then conclude that the substitution of public judgement for Cindy's (with the objective of benefitting her) is not paternalistic? It seems to me that in their desire to link their account of paternalism to a class of findings in behavioural economics, Le Grand and New have mistakenly attempted to build the findings of behavioural economics into the definition of paternalism.

## 5. FURTHER QUALMS

The discussion thus far of Le Grand and New's definition of paternalism suggests that it is mistaken to insist that a policy can be paternalistic only if it addresses a reasoning or judgement failure. Allowing that paternalistic actions could be directed toward many putative mistakes in addition to the cognitive failings identified by behavioural economics loosens the connection between the understanding of paternalism and the findings of behavioural economics. Although Le Grand and New's definition needs to be modified, their account of when paternalism is *justified* might still rest on the distinction between on the one hand interventions that correct reasoning failures or other mistakes agents make concerning the means of achieving their ends and on the other hand interventions that attempt to impose the values of policymakers on members of the population.

Part of what motivates Le Grand and New's definition is their appreciation of the range of subsidies and in-kind benefits that government supplies. These are, they argue, intuitively paternalistic:

Archard (1990b: 37) describes a situation where a secret patron, P, provides free tickets for Q to attend various activities – including the opera – which P believes will be for Q's good, but which she suspects Q would not choose to buy ordinarily. Archard notes that it would 'sound perverse to describe P's behavior with regard to Q as paternalist' precisely because P is *adding* to the set of choices Q already has. . . . However, Archard acknowledges that P may have a paternalist *reason* for behaving as she does even if – in his view – the effect is not paternalistic. But it no longer sounds perverse to describe the behavior as paternalistic if one accepts that P's reason is crucial: P will act in this way only if she believes there is something wrong with Q's judgment. That is where the paternalism comes in. (Le Grand and New 2014: 12)

I deny that subsidies to the arts and to museums are always paternalistic.<sup>10</sup> But they sometimes are, and the provision of many in-kind benefits such as food stamps seem to be at least in part paternalistic. Giving people food stamps rather than cash may reflect the judgement of the policymaker or of tax payers that the recipients of aid are subject to reasoning failures and will spend aid given in cash on less beneficial objects than food. So one might think that reflection on the paternalism of in-kind benefits supports Le Grand and New's definition of paternalism. But this is questionable. Unless one assumes that individuals are entirely self-interested, it may be that there is no cognitive failing in the choice of an action that does not promote an individual's well-being. To adapt an old example presented by T.M. Scanlon (1975), suppose that impoverished members of some sect think that it is vital to build a monument to their deity. If given aid in cash, they would use it for the monument. For that reason, policymakers might prefer to distribute aid in kind in the form of food stamps and housing vouchers. Unless one maintains that the objective of building the monument rests on a failure of judgement, Le Grand and New's definition of paternalism would mistakenly imply that the decision to provide aid in kind in this case is not paternalistic. I conclude that whether there is a reasoning failure is relevant only to the question of whether a paternalistic intervention is justified, not to whether a policy is paternalistic.

Notice that the provision of in-kind benefits, although obviously not a limitation on liberty, as compared with providing no benefits at all, does restrict liberty when it is compared to the alternative of providing individuals with cash.<sup>11</sup> This suggests a possible variation on a Millian view of paternalism that requires as a necessary condition not a limitation on liberty but a greater limitation on liberty than some readily available eligible alternative policy.

## 6. CONCLUSIONS: DEFINING PATERNALISM

What is the right definition of paternalism for the purposes of assessing policies that address the complexities of choice behaviour that behavioural economists have identified, and how is behavioural economics relevant to the definition and justification of paternalism? If

<sup>10</sup> P may believe that if Q received the cash value of the free tickets rather than the tickets, Q would spend the money on other things that would be no less beneficial, but P may be concerned for a variety of reasons to bolster the audience for high culture. Moreover, the additional costs of clothing, transportation, baby-sitting, and so forth may make the opera a decidedly worse choice for Q than other possible purchases, unless someone else is subsidizing the tickets. There may be no failure of judgement here.

<sup>11</sup> It is presumably for this reason that Feinberg takes food stamps to be paternalistic, even though he takes some limitation on liberty to be a necessary condition for a policy to count as paternalistic (Feinberg 1986: 7–8).

one is concerned about liberty, as Mill was, then it is reasonable to focus on those actions that limit liberty. If one is concerned about autonomy, like Shiffrin and to a lesser but still significant degree, Le Grand and New, then the focus lies on the circumvention of the judgement of the individual. If one is in addition interested in failures in reasoning that behavioural economists and psychologists have explored, then it is tempting to focus on those actions that aim to benefit individuals whose reasoning is faulty. But I think that a definition of paternalism that does not build in concerns about liberty or concerns about reasoning failures permits a more perspicuous presentation of the disagreements among those with different views of what makes paternalism problematic or what justifies it. So I suggest the following barebones account, which derives from Shiffrin's discussion:

*(Paternalism: barebones):* A policy is paternalist if and only if it aims to take over or control what is properly within the agent's own legitimate domain of judgment or action for the benefit of the agent.<sup>12</sup>

This definition is certainly not uncontentious, relying as it does on an account of what is 'the agent's own legitimate domain of judgment or action'. But it says nothing about motivation, other than that a paternalist policy aims to benefit the individual it acts on. Clearly, there would be no reason ever to carry out a paternalistic policy unless policymakers thought that individuals were either (1) not good judges of what benefitted them, (2) not motivated to benefit themselves or (3) not able to implement what they judged to be beneficial to themselves. But there is no reason why these conditions need to figure in the definition of paternalism. It is not necessary to insist, as this definition does, that paternalistic acts aim to benefit the agents whose actions one aims to control, but doing so maintains the connection to most of the literature concerned with paternalism.

The conclusions one should draw concerning when, if at all, paternalistic policies are justified cannot be stated briefly and are beyond the scope of this essay. Clearly they depend on philosophical

<sup>12</sup> Compare this definition to Dworkin's remark, quoted earlier, that to ascertain whether an action is paternalistic one must determine, 'whether the act in question constitutes an attempt to substitute one person's judgment for another's, to promote the latter's benefit' (1988: 123). This definition is also largely compatible with the one defended by Brynn Welch and myself. That definition read:

A policy is paternalistic if and only if it aims to advance the interests of some person *P* either (a) via influencing *P*'s choices by shaping how *P* chooses or limiting what *P* can choose or (b) by some means that will take effect regardless of what *P* does and against *P*'s will (Hausman and Welch 2010: 129).

The definition here is simpler than the one Welch and I offer, and it avoids the mistake of insisting that a paternalistic action must be against the will of the person it aims to benefit.

commitments, especially views of freedom, autonomy, and well-being and the weights that should be placed on these. Conclusions concerning when paternalistic policies are justified will also depend on empirical judgements concerning people's competence to control their own lives and concerning the ability of government to identify people's mistakes and to know how to correct them. Even though behavioural economics does not help to define paternalism, it does help one to judge when paternalistic policies are justified.

#### ACKNOWLEDGEMENTS

I am indebted to Julian Le Grand and Philippe Mongin for comments on a previous version.

#### REFERENCES

- Arneson, R. 1980. Mill versus paternalism. *Ethics* 90: 470–489.
- Dworkin, G. 1972. Paternalism. *Monist* 56: 64–84.
- Dworkin, G. 1988. Paternalism: some second thoughts. In *The Theory and Practice of Autonomy*, 121–129. Cambridge: Cambridge University Press.
- Dworkin, G. 2006/2016. Paternalism. In *Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. <https://plato.stanford.edu/entries/paternalism/>.
- Feinberg, J. 1971. Legal paternalism. *Canadian Journal of Philosophy* 1: 105–124.
- Feinberg, J. 1986. *Harm to Self*. New York, NY: Oxford University Press.
- Hausman, D. and B. Welch. 2010. To nudge or not to nudge. *Journal of Political Philosophy* 18: 123–136.
- Le Grand, J. and B. New. 2014. *Government Paternalism: Helpful Friend or Nanny State?* Princeton, NJ: Princeton University Press.
- Mill, J. S. 1978 [1859]. *On Liberty*. Rpt. Indianapolis, IN: Hackett Publishing.
- Scanlon, T. M. 1975. Preference and urgency. *Journal of Philosophy* 72: 655–70.
- Shiffrin, S. 2000. Paternalism, unconscionability doctrine, and accommodation. *Philosophy and Public Affairs* 29: 205–50.
- Sugden, R. 2004. The opportunity criterion: consumer sovereignty without the assumption of coherent preferences. *American Economic Review* 94: 1014–33.
- Vining, A. and D. Weimer. 2010. An assessment of important issues concerning the application of benefit-cost analysis to social policy. *Journal of Benefit-Cost Analysis* 1. ISSN (Online) 2152–2812. doi: <https://doi.org/10.2202/2152-2812.1013>.

#### BIOGRAPHICAL INFORMATION

**Daniel M. Hausman** is the Herbert A. Simon and Hilldale Professor of Philosophy at the University of Wisconsin-Madison. A founding editor of this journal, his research has centred on epistemological, metaphysical and ethical issues lying at the boundaries between economics and philosophy. His most recent books are *Preference, Value, Choice and Welfare* (2012), *Valuing Health: Well-Being, Freedom, and Suffering* (2015) and *Economic Analysis, Moral Philosophy, and Public Policy*, 3rd edn (co-authored with Michael McPherson and Debra Satz, 2017).