THE
ECONOMIC
JOURNAL

The Economic Journal, **130** (*July*), 1384–1415 DOI: 10.1093/ej/uez054
Advance Access Publication Date: 16 October 2019

NUDGE VERSUS BOOST: AGENCY DYNAMICS UNDER LIBERTARIAN PATERNALISM*

Ralph Hertwig and Michael D. Ryall

Thaler and Sunstein (2008) advance the concept of 'nudge' policies—non-regulatory and non-fiscal mechanisms designed to enlist people's cognitive biases or motivational deficits so as to guide their behaviour in a desired direction. A core assumption of this approach is that policymakers make artful use of people's cognitive biases and motivational deficits in ways that serve the ultimate interests of the nudged individual. We analyse a model of dynamic policymaking in which the policymaker's preferences are not always aligned with those of the individual. One novelty of our set-up is that the policymaker has the option to implement a 'boost' policy, equipping the individual with the competence to overcome the nudge-enabling bias once and for all. Our main result identifies conditions under which the policymaker chooses not to boost in order to preserve the option of using the nudge (and its associated bias) in the future—even though boosting is in the immediate best interests of *both* the policymaker and the individual. We extend our analysis to situations in which the policymaker can be removed (e.g., through an election) and in which the policymaker is similarly prone to bias. We conclude with a discussion of some policy implications of these findings.

Thaler and Sunstein (2008) have explored various ways in which policymakers may design non-regulatory and non-fiscal interventions to help people make better decisions. These 'nudges' come in two main forms: 'educative' (e.g., disclosure requirements, labels, reminders) and 'non-educative' (e.g., ordering of items on a website, cafeteria design, automatic enrolment). Non-educative nudges enlist cognitive biases or motivational deficits (e.g., inertia, procrastination, loss aversion; see Rebonato, 2012) to steer individuals' behaviour in a desired direction. The promising cost–benefit ratio of these non-educative policies, but also their potential for manipulation and paternalism, make them both interesting and controversial.

One key assumption of the nudging approach is that benevolent policymakers dutifully structure non-educative nudges so as to: (*i*) achieve ends consistent with individuals' ultimate and true preferences; and (*ii*) allow individuals to easily reverse the choice implied by the nudge if they so desire.[1] This approach to public policy design is known as 'libertarian paternalism'. But is it reasonable to take at face value the assumption that policymakers consistently apply benevolent self-restraint when nudging? Economics has a rich tradition of examining principal–agent problems in a wide variety of settings. Why not extend this tradition to the premise of guileless libertarian paternalists? Agency issues may be particularly relevant in this context because the policy and its operating mechanism can be—by its very nature—relatively hidden. Some nudges may work only because individuals lack the cognitive competences or the motivation required to detect something as a nudge (e.g., a default) or to see through the underlying psychological

* Corresponding author: Michael D. Ryall, Rotman School of Management, University of Toronto, 105 St. George Street, Toronto M5S 3E6, Canada. Email: mike.ryall@rotman.utoronto.ca

This paper was received on 12 September 2016 and accepted on 22 July 2019. The Editor was Martin Cripps.

We thank F. Ederer, C. Sunstein, R. Bénabou, the participants of the 13th Blankensee Colloquium and the Berlin Institute of Advanced Study. We also thank our research assistant, Natalia Uborceva. Ryall is grateful for funding provided by the Social Sciences and Humanities Research Council of Canada.

[1] Other safeguards such as the Rawlsian publicity principle are also assumed; here, it could mean prohibiting governments from implementing any policies they would not be willing to defend in public.

mechanism. A recent survey indicates little in the way of explicit theoretical work examining this issue (Schnellenbach and Schubert, 2014).

In this article, we examine potential agency issues. Specifically, we analyse policy formulations in which: (*i*) a mismatch of the goals held by policymakers and nudged individuals is possible; and (*ii*) individuals have a cognitive bias. A key novelty of our approach is that policymakers are endowed with the ability to *eliminate a nudge-enabling bias permanently and without cost*. We refer to such options as 'boost' policies (Hertwig and Grüne-Yanoff, 2017; Walton, 2014, for a related approach). By most standards, including those advocated by libertarian paternalists, liberating people from the burden of their cognitive or motivational impediments at zero cost would constitute an unambiguous social good. Yet, as we demonstrate, when the goals of policymakers and individuals are potentially at odds, policymakers may choose to withhold the transfer of durable cognitive competences (i.e., not to boost) in order to retain the option of using nudges in the future. This can happen even when both the policymaker and individual agree that boosting is, presently, the optimal policy.

The essential feature of the boost policy in our model is its durability and generalisability: once the individual decision maker learns how to avoid a particular kind of bias, this new ability can be applied, so the assumption, to all future decisions of a similar nature. In our analysis, the policymaker simply needs to invoke the policy, in which case the boost takes effect instantaneously at zero cost. The effects of boost are permanent. Of course, boosting in the real world can be time-consuming and resource-intensive.[2] The strong assumption that boosts cost the policymaker nothing makes our result a particularly strong one, in the sense that higher costs serve only to make boosts less attractive. The implication is that, if policymakers forgo boosts even under the present idealised conditions, the under-provision of boost interventions may be a serious issue in the real world. The effect of the boost is permanent.

Our main result is to demonstrate that nudge policies have option value to policymakers faced with potential goal conflicts in future policy problems. Specifically, we describe conditions under which the individual falls prey to agency problems on the part of the policymaker, with the latter implementing inefficient nudges—rather than one-off boosts—in the present in order to retain the option of using the nudges again in the future. The dynamic trade-off that arises here is a result of the asymmetry inherent in the two approaches: nudging today leaves open the option of either boosting or nudging tomorrow; boosting today eliminates the bias and, hence, the option to nudge in the future. In our model, an individual faces a stochastic sequence of policy issues that differ in terms of goal alignment. On some policy issues, policymaker and individual agree on desired ends. On others, they do not. Nudges are inefficient in those periods in which boosting is optimal from *both* parties' perspective. The policymaker may nevertheless choose to nudge in such cases in order to preserve the option of using the nudge in future periods in which goals are in conflict.

The idea that policymakers may prefer to design and implement policies for imperfectly rational agents is not new. For example, macro-economic policymakers may prefer to face agents with static rather than rational expectations (see, e.g., Sargent and Wallace, 1975); employers may prefer workers who overestimate or underestimate their abilities (Fang and Moscarini, 2005; Benabou and Tirole, 2003) and shareholders may prefer managers with a certain amount of

[2] Several authors have begun to explore a variety of learning and information-transmission issues in the presence of cognitive biases (Allcott, 2011; Costa and Kahn, 2013; Marteau *et al.*, 2012; Carlin *et al.*, 2013), including the availability of low-cost boosts (see Hertwig and Grüne-Yanoff, 2017).

overconfidence (Gervais *et al.*, 2011).[3] The essential novelty in our treatment is that we model a policy setting in which there is an explicit choice of whether to eliminate an individual's cognitive bias once and for all or to allow it to remain. Our conclusions are important given the enormous interest in nudging as a form of behaviourally informed social policy (e.g., Halpern, 2016; Sunstein, 2013; OECD, 2017). Also, as far as we know, this treatment is the first to consider a dynamic setting in which the policymaker's preferred degree of cognitive bias in the individuals to be nudged or boosted varies stochastically over time (the technical implementation of which is straightforward). This variation is what gives rise to the dynamic tension that is the focus of our main result.

Nudging Towards the Social Good?

The opening premise of Thaler and Sunstein (2008, p. 5) is that 'in many cases, individuals make pretty bad decisions – decisions they would not have made if they had paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control'. In their view, cognitive and motivational deficits are both pervasive and consequential. They argue that people often make mistakes because they 'rely too much on [the] Automatic System' (ibid, p. 21); that biased assessments of risk can 'perversely influence how [people] prepare for and respond to crises, business choices, and the political process' (ibid, p. 25); that people are 'unrealistically optimistic even when the stakes are high' (ibid, p. 32) and 'adopt ... the "yeah, whatever" heuristic' (ibid, p. 35). Last but not least, humans suffer from self-control problems such as 'smoking, alcohol, a failure to exercise, excessive borrowing, and insufficient savings' (ibid, p. 42).

Accepting these deficits as given and hard to remove, Thaler and Sunstein (2008) argue that nudging, under the constraints of libertarian paternalism, is a both efficacious and ethical approach to policy design. A non-educative nudge in the context of libertarian paternalism is a policy in which: (*i*) cognitive biases and motivational deficits are enlisted in the service of policy goals; (*ii*) the policy goals are consistent with the ultimate preferences of the individuals subject to the nudge; and (*iii*) individuals are easily able to reverse the choice instigated by the nudge. Thus, the libertarian paternalist is a policymaker tasked with designing institutional mechanisms that create decision-making contexts—'choice architectures' (ibid, pp. 11–13)—that steer behaviour in ways that improve welfare at both the individual and the group level.[4] Policies that preserve freedom of choice by implementing easily reversible options pass the libertarian test. The idea is to create a 'choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives' (ibid, p. 6). Thus, nudging is intended as a 'relatively weak, soft and non-intrusive type of paternalism' (ibid, p. 6).

The attraction of this approach is, in part, its promise to generate efficient policy mechanisms driven and informed by scientific evidence rather than political dogma. Under nudging, behaviour is influenced without recourse to more forceful injunctions (e.g., regulations) or incentives (e.g., taxes).[5] The notion of evidence-based, libertarian paternalism is novel and appealing and has had

[3] We thank Roland Bénabou for raising this point.

[4] The policy domain is typically public (government-related).

[5] The observations that social influences and pressures can be used to steer a person's behaviour is well established (Cialdini, 2009); so is the observation that the provision of information—i.e., educative nudges—may facilitate better decisions (Thaler and Sunstein, 2008, pp. 93–4).

substantial influence on public policy debates around the globe.[6] At the same time, it has also been subject to critical debate. The issues raised include the pervasiveness and predictability of human irrationality; the ethical implications of nudging; the lack of conceptual clarity around nudging; the extent to which libertarian paternalism is different from other forms of paternalism; whether policy proposals follow from theoretical claims about the limitations of human rationality and self-control; and the effectiveness of nudges as proposed and implemented.[7]

The success of the libertarian paternalism programme is dependent on an absence of agency problems on the part of the policymaker. Yet, the agency issue seems particularly worrisome under this programme, precisely because sometimes the nudge (e.g., positional arranagements) and quite often the operating principle behind the nudge are hidden from the decision maker.[8] Recognising that the approach of using people's cognitive and motivational deficits to influence their decisions may invite abuse, the advocates of libertarian paternalism require such (non-educative) interventions meet the following conditions (adapted from Rebonato, 2012, p. 32):

(*1*) A pervasive bias is attached to a specific decision task. The bias may arise from a property of the environment (e.g., framing of information), limited cognition (e.g., inability to assess probabilities) or a motivational deficit (e.g., procrastination).
(*2*) The consequence of the bias is detrimental to the individual.
(*3*) Enlisting the bias (or some other bias) in such a way as to defeat its undesired consequences is feasible.
(*4*) Policymakers know the relevant preference of the affected individuals.
(*5*) Policy implementation is consistent with the preferences of the targeted individuals.

The final entry on this list of conditions underlines the key role of the policymaker having benevolent intentions. But serious concerns have been raised about this assumption, most pointedly by *The Economist* (Anonymous, 2006): 'Yet from the point of view of liberty, there is a serious danger of overreach, and therefore grounds for caution. Politicians, after all, are hardly strangers to the art of framing the public's choices and rigging its decisions for partisan ends. And what is to stop lobbyists, axe-grinders and busybodies of all kinds hijacking the whole effort?' Earlier, Sunstein and Thaler (2013, p. 1201) suggested an 'opt-out right' as a 'safeguard against confused or improperly motivated policy makers'. Thaler and Sunstein (2008, pp. 239–41) also acknowledged the possibility of 'evil nudgers and bad nudges'. As safeguards, they emphasised the role of transparency, disclosure and regulation ('rules of engagement'; ibid, p. 240). Ultimately, they appear to trust in democratic systems of checks-and-balances, arguing for a 'libertarian check' on bad plans: 'To the extent that individual self-interest is a healthy check on planners, freedom of choice is an important corrective' (ibid, pp. 240–1).

Of course, the idea that the 'right' public choice architects will be selected via the institutions of democracy may be less than fully convincing given that the act of voting is, in principle, subject to

[6] For instance, Sunstein served as the Administrator of the White House Office of Information and Regulatory Affairs in the Obama administration, and Thaler has advised the British government's Behavioural Insights Team (BIT). More generally, numerous governments and international organisations such as the World Bank (2015), the European Commission (Lourenco *et al.*, 2016) and the OECD (2017) have begun to acknowledge the enormous potential of behavioural science evidence in helping to design more effective and efficient public policies.

[7] Examples of recent contributions in this vein include Sugden (2008); Bovens (2009); Berg and Gigerenzer (2010); Allcott (2011); Grüne-Yanoff and Hertwig (2016).

[8] Some nudges, such as Save More Tomorrow, require the explicit attention of the target individual. Typically, however, the mechanistic reason a nudge works is not transparent to the person being nudged.

the very same cognitive and motivational deficits.[9] Why assume that politicians, their appointees and professional bureaucrats are immune to the temptations of pursuing their own agendas and economic interests? By introducing the possibility of goal divergence between policymakers and constituents, we are able to explicitly examine conditions under which the consequences of the nudge policies implemented are inconsistent with individual preferences. 'Evil nudgers' is too strong a description here—even well-intentioned policy goals may be in conflict with individual welfare. We leave open the question of whether agency problems of this kind are adequately controlled by democratic institutions.[10]

Boosting as a Policy Alternative

One way of framing the choice of a well-intentioned public policymaker is as that between selecting traditional policy options ('hard paternalism'), applying nudge policies ('libertarian paternalism') or simply doing nothing at all. However, because the root of the problem in the nudge context is a robust bias on the part of the individual, there is also a fourth possibility: training people to avoid the bias, thus enabling them to maximise their own welfare without the help of a policymaker or other outside party. This option is interesting because, other things being equal, it should dominate nudging.[11] Grüne-Yanoff and Hertwig (2016) have termed this option 'boosting'.

Identifying effective training options has been a key objective of various programmes in psychology (see discussion in Hertwig and Grüne-Yanoff, 2017, p. 7), including the research programme on simple and ecologically rational heuristics (Gigerenzer *et al.*, 2011). The issue at the heart of the discussion between libertarian paternalism and other research programmes on bounded rationality is the extent to which individuals are capable of surmounting their biases. The nudge view is that human beings are essentially hostage to a rapid, automatic system of cognition (Thaler and Sunstein, 2008, p. 21). The boost view (Hertwig and Grüne-Yanoff, 2017), in contrast, is that individual decisional and motivational competences can be improved either directly—by building on and extending a person's repertoire of competences—or indirectly by, for example, making thoughtful changes in the environment (e.g., presenting information in formats that ease computational burden, Hoffrage and Gigerenzer, 1998).[12]

Boost policies have two key features: an effort requirement and a general effect. First, the individual is required to be sufficiently motivated to invest time and effort in acquiring and exercising a new competence. Additional resources may also be required at the policy level (e.g., instruction materials, instructors, an institutional framework). Second, the competence acquired can, at least in theory, be exercised in all instances of decision problems prone to the bias it defeats. Thus, under a boost policy, the individual will be enabled to independently overcome a bias for an entire class of problems—at the expense of an up-front resource investment. In contrast, nudge policies impose little in the way of up-front investments (in the sense that no individual training is required)—at the cost of leaving the individual vulnerable to future instances of the bias. This contrast creates an interesting policy dichotomy within which to study the potential for strategic behaviour on the part of policymakers.

[9] The biases of policymakers themselves is a separate issue (see, e.g., Tasic, 2011; Kuehnhanss *et al.*, 2015).

[10] For possible answers, see, e.g., Besley *et al.* (2011); Besley and Reynal-Querol (2011); Galasso and Nannicini (2011, 2015).

[11] This assertion is consistent with the criterion that policymakers 'create a choice architecture that will make it more likely that people will promote their own ends, as they themselves understand them' (Sunstein, 2014, p. 19).

[12] See Hertwig and Grüne-Yanoff (2017) for a discussion of specific boosts.

We analyse precisely this choice between boosting and nudging in a dynamic context where the policymaker's goals are potentially not aligned with those of the constituents. By boosting competences, policymakers ensure that constituents are, in future, able to choose in accord with their own preferences. Yet, anticipating the potential for goal mismatch in the future, policymakers may choose to withhold such competence-enhancing measures. This may be true even in cases when boosting is consistent with the policymaker's present preferences. We say 'may' because the final resolution will depend on a number of factors, including the patience of the policymaker, the frequency with which goal mismatch occurs and the relative cost considerations.

Throughout the following analysis, we make extensive use of concrete examples to illustrate the intended interpretation and logic of the mathematics as well as to engage those less interested in the mathematical details. In addition, we provide ample references to the relevant work in psychology for the interested reader. In the next section, we lay out the details of our set-up: a stage game that contrasts nudge and boost policies. The set-up is relatively simple, yet sufficiently flexible to capture several categories of cognitive bias. We analyse the stage game in Section 1, with careful attention to the implications of a beneficent policymaker (Subsection 1.3) versus a self-interested one (Subsection 1.4). With this foundation in place, we move on to examine the dynamic case in Section 2. Our main result is Proposition 3, presented in Subsection 2.1. In Subsection 2.3 we explore what happens when the decision maker has the ability to terminate the policymaker due to poor performance. In Section 3, we extend the analysis to consider the possibility that the policymaker is also prone to cognitive bias. We conclude with a discussion of policy implications in Section 4.

1. Stage Game

For the current purpose, we characterise a decision problem as consisting of three elements, two of which are preference related. The first is a state variable that determines which of the decision maker's actions maximises his own utility. The second is a variable that indicates whether the policymaker's preferences are or are not aligned with those of the decision maker. The third element is a variable that indicates a bias-relevant context. We assume that: (*i*) stochastic variation in the preference variables creates uncertainty with respect to the optimal action (e.g., does the decision maker need cancer treatment and, if so, do considerations about costs and a limited healthcare budget put the policymaker at odds with this need?); (*ii*) context arises naturally with respect to the problem at hand (e.g., medical test results); and (*iii*) the decision maker's choice may be influenced by the context (e.g., cancer treatment required if test results are positive).

We also assume that, prior to any policy manipulation, there is an exogenous, stochastic relationship between context and preference variable. Here, "context" represents the contingencies influenced by the policy maker. However, context does not capture all of the prerence-relevant features of the decision problem faced by a specific decision maker. Thus, the model allows for situations in which a given policy works for less than 100% of the population or, alternatively, works for a given decision maker less than 100% of the time. Consider, for illustration, a classic example of a nudge from Thaler and Sunstein (2008): the arrangement of food choices in a cafeteria. The idea is that, through simple changes in the architecture of the cafeteria, people can be nudged to make healthier choices. Salad bars, for instance, can be moved away from the back and placed directly beside the cash registers, whereas crisps can be moved to a more distant location, out of immediate view. Such a choice architecture, combined with a reliance on individual inertia and attentional occupation with items that are close and at eye level, should, in theory, be conducive to healthier food choices. Indeed, 'visibility enhancement' of healthier options (e.g., giving them eye-level shelf position and first place on the menu) and 'convenience enhancements' (i.e., making healthier options easier to select or consume) have been found to increase the likelihood of intended choices (Cadario and Chandon, 2019).

The assumption of a stochastic relationship between context and preference means that the preference for one kind of food is not absolute. For instance, a runner in training for a marathon may load up on peach cobbler because he currently requires high-caloric food; in this case, being nudged towards fruit and vegetables may undermine their health and performance. Suppose the decision maker's net calorie status at the time of a meal determines their optimal action: eat a high-calorie meal when the calorie deficit is high and a low-calorie meal when the deficit is low (or there is a calorie surplus). Here, the food arrangement of the cafeteria creates the context. When running a calorie deficit, the decision maker may (without much thinking) tend towards a cafeteria with arrangements that encourage higher calorie consumption. By the same token, when running a calorie surplus, he may tend towards eateries that encourage low-calorie fare. If so, then co-ordinating meal selection (high or low caloric content) with the food arrangement in the present context is optimal. Alternatively, the individual may frequent the two styles of cafeteria with equal probability, and yet have a high tendency to be in a state of calorie surplus. In this case, choosing a low-calorie meal regardless of the food arrangement context is optimal.

This structure highlights the problem—even for rational decision makers—of making errors under uncertainty. Even if an individual is inclined to frequent cafeterias whose food arrangements correspond to their dietary needs, errors are still possible: they may nevertheless succumb to overeating (e.g., binging) or choose a low-calorie meal when running a calorie deficit. Alternatively, they may have an eating disorder (e.g., anorexia nervosa) that prompts them to frequent 'healthy' cafeterias too often, over-restricting the consumption of high-calorie food. These possibilities highlight that real-world policy issues are rarely solved simply by imposing 'healthy' food arrangements to induce healthy meal choices. In our setting, the policymaker is a well-informed, rational actor who has three options. One is to do nothing, letting individuals make choices according to their own, possibly biased, decision processes. The second is to apply an intervention designed to manipulate the context (through a nudging choice architecture). The third is to bring the individuals' knowledge and competences into line with their own preferences (boost). Each of these policies has different implications. As we now elaborate, which is best depends on the parameters of the decision problem and the nature of the bias.

1.1. *Model Set-up*

The stage game has two players: a Decision Maker, D, and a Policymaker, P. In order to keep things simple, we make extended use of binary variables. We use subscripts to indicate the agents with whom variables are associated. The *decision problem* is a triple $\omega \equiv (x, y, \phi) \in \{-1, 1\}^3$. Thus, there are eight ways for the decision problem to instantiate. When ω is understood from the context, we abuse notation and write $x(\omega)$, $y(\omega)$, and $\phi(\omega)$ to indicate its corresponding component. For example, if $\omega = (1, -1, -1)$, then $x(\omega) = 1$. The first component, x, may be thought of as a state variable associated with the problem at hand that determines D's preferences over his actions. In this model, D's task is to harmonise their action with the decision problem according to x in a fashion described below. The second component determines P's preferences with respect to D's action. The third component, ϕ, is the decision-relevant context variable, observed by both P and D.

Decision problems are generated at the beginning of a stage according to a probability distribution, denoted ρ. For instance, $\rho(\omega)$ is the probability that the problem is ω; $\rho(x=1)$ is the probability that $x=1$, and $\rho(y=1|x=1)$ is the conditional probability that $y=1$ given $x=1$, etc. The distribution ρ may arise from more primitive stochastic processes corresponding to a particular situation. For example, x and y may be independently generated according to distributions $f(x)$ and $f'(y)$, respectively, and ϕ may be a 'signal' generated according to the conditional distribution $f''(\phi|x)$. Then, $\rho(x, y, \phi) = f(x)f'(y)f''(\phi|x)$.

Our analysis begins with ρ as given. The context ϕ is associated with x according to $\rho(\phi|x)$. A rational agent can make inferences about x via Bayes' Rule, e.g.,

$$\rho(x|\phi) = \frac{\rho(\phi|x)\rho(x)}{\rho(\phi|x)\rho(x) + \rho(\phi|-x)\rho(-x)}. \tag{1}$$

To implement various types of bias, we assume that D begins the game with *subjective prior beliefs*, denoted $\tilde{b}$, on a *perceived* set of decision problems $\tilde{\omega} \equiv (\tilde{x}, \tilde{y}, \tilde{\phi}) \in \{-1, 1\}^3$, where we adopt the convention of using tildes to indicate objects for which D's subjective perceptions may depart from reality. We assume that D is a subjective utility maximiser with respect to $\tilde{b}$ based on her perception of the context, $\tilde{\phi}$, where either or both of these subjective objects may be erroneous. In the spirit of the nudge literature, we assume that such departures from reality are robust and persist regardless of any disconfirming data implied by D's experiences. The only remedy in our setting is the permanent correction provided by a boost policy. Finally, to avoid division by zero problems when computing conditional probabilities, assume that both ρ and $\tilde{b}$ are strictly positive. Thus, for example, a beneficent P might be modelled by assuming $\rho(y=x) = .99$.

Once ω is generated, P moves. P knows ϕ, ρ and $\tilde{b}$. P chooses a policy $p \in \{0, -1, 1, B\}$, which corresponds to *no intervention*; *nudge* by manipulating $\tilde{\phi}$ (the context perceived by D) to equal -1; *nudge* by manipulating $\tilde{\phi}$ to equal 1; and *boost*, respectively. We allow the *viability* of nudges—meaning whether P's policy to manipulate the context is effective—to be dependent on the true context. Nudge viability is exogenously determined for each context at the beginning of the game.

For example, suppose the decision is whether or not to sign an organ donor card, where $\phi = 1$ is the context of applying for a driver's licence and $\phi = -1$ is the context of signing up for a kidney transplant wait list. In the first context, a person may be nudged by framing the choice in terms of an 'opt-out' default on the application (see Johnson and Goldstein, 2003). In the second, worries about being implicitly disqualified from a life-or-death transplant by answering 'incorrectly'—i.e., asking for a donation while being unwilling to reciprocate—may prompt a person to sign, regardless of how the donor card is worded. If so, $p = -1$ is viable when $\phi = 1$: implementing this policy nudges $\tilde{\phi}$ from 1 to -1. However, the nudge policy $p = 1$ is not viable in the transplant context, $\phi = -1$. Presumably, no psychological slight-of-hand will be sufficient to budge $\tilde{\phi} = \phi = -1$ to 1 for a desperate patient who is highly alert to avoiding any behaviours that could jeopardise their chances of transplant.

Note, the interpretation is not that nudging when $\phi = 1$ causes D to believe he is signing up for a kidney donation while they are actually filling out forms at the Department of Motor Vehicles (DMV). Rather, it simply implies that they are induced to make the same choice in the DMV context as they would in the kidney waiting-list context. Unless otherwise indicated, assume all nudges are viable.[13] For our purposes, boosts are always viable and free of cost.[14]

The policies of no intervention and nudging leave D's subjective beliefs intact. Boosting aligns D's subjective beliefs with the true distribution and, at the same time, renders artificial manipulation of the context ineffective (which implies that the preceding policy choices are mutually exclusive—it is not possible for P to simultaneously nudge and boost). Boosting transforms D into a rational expected utility maximiser; more generally, this need not be the case (Hertwig and Grüne-Yanoff, 2017) but we assume it here for simplicity.

Following P's choice of p, D perceives the natural (true) context if there was no nudge and the manipulated context if there was a nudge. D then chooses one of two available *actions*, $a \in \{-1, 1\}$. The consequence of D's action given $\omega = (x, y, \phi)$ is the *payoff* $\pi_D(a, \omega) \equiv ax(\omega)$: D's utility is maximised when their action matches x (i.e., is properly aligned with the true problem). Similarly, P receives payoff of $\pi_P(a, \omega) \equiv ay(\omega)$. Notice that relationships between the agents' preferences are introduced through ρ. For example, preferences may be stochastically independent but different on average: $\rho(x, y) = \rho(x)\rho(y)$ and $\mathbb{E}(x) \neq \mathbb{E}(y)$. Alternatively, P's overwhelming concern may be to maximise D's utility, in which case we have $\rho(y = x) = \sim 1$.

Suppose D is fully rational. Then, given ω, D observes $\phi(\omega)$ and maximises

$$\begin{aligned}\mathbb{E}(\pi_D|a, \phi) &= \sum_x ax\rho(x|\phi) \\ &= a\left(\rho(x = 1|\phi) - \rho(x = -1|\phi)\right) \\ &= a\left(2\rho(x = 1|\phi) - 1\right).\end{aligned} \tag{2}$$

Let $a^*(\phi) \in \arg\max_{a\in\{-1,1\}} \mathbb{E}(\pi_D|a, \phi)$ be D's expected utility maximising action given ϕ.[15] Optimising (2) boils down to identifying which value of x is more likely given the context ϕ, then selecting the matching action; e.g., if $\rho(x = 1|\phi) \geq 0.5$, then $a^*(\phi) = 1$. This has the nice implication that

$$\mathbb{E}(\pi_D|a^*, \phi) = \left|2\rho(x = 1|\phi) - 1\right|. \tag{3}$$

The expected utility of implementing the optimal context-contingent strategy is, therefore,

$$\mathbb{E}(\pi_D|a^*) = \textstyle\sum_\phi \left|2\rho(x = 1|\phi) - 1\right|\rho(\phi). \tag{4}$$

The effects of policy p are as follows. Let $\tilde{b}_p$ indicate D's subjective beliefs *following the choice of policy p*. Then, if $p \neq B$, $\tilde{b}_p = \tilde{b}$; otherwise a boost is implemented ($p = B$), which immediately renders $\tilde{b}_B = \rho$. Similarly, let $\tilde{\phi}_p$ indicate the *context perceived by D given p*. If $p = -1$ or 1 (nudges), then $\tilde{\phi}_p = -1$ or 1, respectively. Otherwise, $\tilde{\phi}_p = \phi$. Let

$$\begin{aligned}\tilde{\mathbb{E}}(\pi_D|a, \phi, p) &\equiv \sum_{\tilde{x}} a\tilde{x}\tilde{b}_p(\tilde{x}|\tilde{\phi}_p) \\ &= a\left[2\tilde{b}_p(\tilde{x} = 1|\tilde{\phi}_p) - 1\right]\end{aligned} \tag{5}$$

13 Nudge viability becomes an important issue in the dynamic setting.
14 The point being that a finding of boost failure is strongest under these conditions.
15 Note that expected utility is independent of policy p. D has no preferences over policy choices per se.

indicate D's *subjective expected utility* given the true context, the policy and their choice of action.[16] For example, if $p = 1$, then $\tilde{\mathbb{E}}(\pi_D|a, \phi, 1) = a\left[2\tilde{b}(\tilde{x} = 1|\tilde{\phi} = 1) - 1\right]$. Alternatively, if $p = B$, then $\tilde{\mathbb{E}}(\pi_D|a, \phi, B) = a\,[2\rho(x = 1|\phi) - 1] = \mathbb{E}(\pi_D|a, \phi)$.

Define D's *subjectively optimal strategy* $s_D^*(\phi, p) \in \arg\max_{a\in\{-1,1\}} \tilde{\mathbb{E}}(\pi_D|a, \phi, p)$ as the action that maximises D's subjective expected utility given context ϕ and policy p. D chooses the action that matches the most likely x according to their subjective grasp of the situation. Summing up, $a^*(\phi)$ is the rationally optimal, context-contingent action and $s_D^*(\phi, p)$ is the subjectively optimal stage game strategy given the chosen policy under the true context. Unless indicated otherwise, we assume D always chooses actions according to s_D^*.

What does it mean for D to be 'cognitively biased' in this set-up? The question is not trivial. For example, it is easy to construct cases in which $a^*(\phi(\omega)) \neq x(\omega)$; i.e., the action that maximises *expected* utility may well err with respect to the actual x in the problem at hand. Still, for someone who only observes ϕ, choosing $a^*(\phi)$ is fully rational. Alternatively, the mere fact that $\rho \neq \tilde{b}$ does not imply a departure from rational behaviour: ρ and $\tilde{b}$, though unequal, may yet be sufficiently similar as to result in the same behaviour: $s_D^*(\cdot, 0) = a^*(\cdot)$. Therefore, we define *bias* as a situation in which ρ and $\tilde{b}_0$ are such that $s_D^*(\phi, 0) \neq a^*(\phi)$ in at least one context, ϕ. Implicit in this definition of bias is that boosting ($p = B$) always results in its elimination.

Assuming that D operates according to subjective beliefs that are unconstrained with respect to reality allows us to represent many sorts of cognitive bias. For example, confirmation bias is when decision makers seek or interpret evidence in ways that are partial to existing beliefs, expectations or a hypothesis in hand (Nickerson, 1998). If ϕ is new information, an extreme version of this bias is $\tilde{b}(x) = \rho(x)$ and $\tilde{b}(x|\phi) = \rho(x)$ even though $\rho(x|\phi) \neq \rho(x)$; i.e., D incorrectly dismisses the information. At the other end of the spectrum, anchoring bias is when an individual places too much weight on present (possibly irrelevant) evidence (i.e., anchors; Tversky and Kahneman, 1974). If ϕ represents recent evidence, then a version of this bias might have $\rho(x|\phi) > \rho(x)$ but $\tilde{b}(x|\phi) > \rho(x|\phi)$, indicating that D overreacts to the evidence. A third example is the failure to adjust for base-rate probabilities. Here, ϕ could be the results of a diagnostic test which arises with probability $\rho(\phi|x)$, where x indicates the presence of a disease. Suppose $|\rho(x) - \rho(-\,x)|$ is large (e.g., a small percentage of the population actually has the disease). The bias can be represented by setting $\tilde{b}(\phi|x) = \rho(\phi|x)$ but then further assuming $\tilde{b}(x) = \tilde{b}(-x)$; i.e., D does not distinguish between differences in base-rate probabilities. This results in a flawed calculation of the quantity of interest, $\rho(x|\phi)$, via biased subjective application of Bayes' Rule (1).

Presumably, the extent to which such departures from rationality are sustainable in the face of disconfirming evidence depends on factors such as the robustness of the bias, the importance of the problem, experience and possible incentives. We simply take the bias as given and persistent. Our stochastic set-up does allow us to model situations in which a typically rational decision maker makes an occasional, context-dependent mistake. This adds a dimension of richness as it allows us to cast nudges as a policy tool of varying efficacy independent of policymaker motivations. We explore several numerical examples below.

Given a context ϕ, under the assumption that D plays subjectively optimally, P's choice of policy nets them an expected payoff of

$$\mathbb{E}(\pi_P|\phi, p) = s_D^*(\phi, p)\big(2\rho(y = 1|\phi) - 1\big). \tag{6}$$

[16] We adopt the notational convention that $\mathbb{E}$ indicates correct expectation and $\tilde{\mathbb{E}}$ subjective expectation.

P's stage game *strategy* is a map from observed contexts to policy choices, with $s_P(\phi) = p$ indicating that policy p is chosen when the context is ϕ. Then, the expected value of a policy strategy s_P is

$$\mathbb{E}(\pi_P|s_P) = \sum_\phi s_D^*(\phi, s_P(\phi))\big(2\rho(y=1|\phi) - 1\big)\rho(\phi). \tag{7}$$

We write s_P^* to indicate a policy strategy that maximises (7).

A policy strategy implies a probability with which D makes a successful decision ('successful' from D's perspective). Let $\gamma(s_P)$ indicate the *probability of a successful decision* implied by s_P, where

$$\gamma(s_P) = \tfrac{1}{2}\sum_\omega \Big|x(\omega) + s_D^*(\phi(\omega), s_P(\phi(\omega))\Big|\rho(\omega). \tag{8}$$

When the action chosen according to s_D^* under s_P is correct, it matches x, making the absolute value in (8) equal to 2; otherwise, it equals 0. Thus, γ measures the performance of P's policy strategy in terms of the expected share of correct decisions that D makes as a result. Given the set-up, this translates directly to payoffs: $\gamma(s_P) < 0.50$ implies the consequence of a policy is a negative expected payoff for D. If $s_P = p$ is a constant strategy (over contexts), we simply write $\gamma(p)$; in particular, $\gamma(0)$ is the percentage of correct decisions arising from the constant no-intervention policy. Finally, let γ^* denote the maximum value of γ attainable via choice of s_P.[17]

1.2. *Example: Cafeteria Design*

Returning to our earlier discussion, let us consider a parametrisation that is consistent with the redesign of a cafeteria. Let x indicate the decision maker's net level of calories burned during the day, where $x = 1$ means *high*, $x = -1$ means *low*. Assume P's preferences are strongly aligned with D's own: $\rho(y = x) = 0.99$. Suppose D leads a sedentary lifestyle such that $\rho(x = -1) = 0.8$. Although D does not own a personal exercise tracker, their preference for a 'normal' cafeteria ($\phi = 1$) versus one with a healthy food arrangement ($\phi = -1$) tends to be correlated with their calorie expenditure: $\rho(\phi = x|x) = 0.6$.

D has a cognitive bias such as overconfidence bias (e.g., Koriat and Bjork, 2005): they mistakenly believe that they have a finely tuned skill to pick a cafeteria in accordance with their actual calorie needs: $\tilde{b}(\phi = x|x) = 0.9$. Otherwise, their beliefs are consistent with reality. Table 1 elaborates both the objective distribution ρ and the subjective distribution $\tilde{b}$ on decision problems $\omega = (x, y, \phi)$.

What can we infer about D's behaviour from $\tilde{b}$? Left alone by P and observing ϕ, D maximises

$$\begin{aligned}\tilde{E}(\pi_D|a,\phi,0) &= a\left[\tilde{b}_0(x=1|\phi) - \tilde{b}_0(x=-1|\phi)\right],\\ &= \begin{cases} a\,[0.03 - 0.97] = -0.94a, & \text{if } \phi = -1,\\ a\,[0.69 - 0.31] = 0.38a, & \text{if } \phi = 1.\end{cases}\end{aligned}$$

The implication is that D makes dietary choices in accordance with the cafeteria setting, $s_D^*(\phi, 0) = \phi$. The behavioural implications are easily seen in Table 2. According to $\tilde{b}$, when $\phi = -1$, D picks $a = -1$ because $0.72 > 0.02$ ($x = -1$ is more likely); and when $\phi = 1$, D picks $a = 1$ because $0.18 > 0.08$. Comparing these values with the true joint probabilities $\rho(x,$

[17] D's expected payoff is $2\gamma(s) - 1$. Therefore, the s_P that maximises γ also maximises (2).

Table 1. *Actual (ρ) and Subjective ($\tilde{b}$) Distributions on Decision Problem Instantiations.*

ω				
x	y	ϕ	ρ	$\tilde{b}$
1	1	1	0.119	0.178
1	1	−1	0.079	0.020
1	−1	1	0.001	0.002
1	−1	−1	0.001	0.000
−1	1	1	0.003	0.001
−1	1	−1	0.005	0.007
−1	−1	1	0.317	0.079
−1	−1	−1	0.475	0.713

Table 2. *Joint Probabilities $\tilde{b}(x, \phi)$.*

	$\tilde{x}$		
$\tilde{\phi}$	−1	1	$\tilde{b}(\tilde{\phi})$
−1	0.72	0.02	0.74
1	0.08	0.18	0.26
$\tilde{b}(\tilde{x})$	0.80	0.20	1.00

Table 3. *Joint Probabilities $\rho(x, \phi)$.*

	x		
ϕ	−1	1	$\rho(\phi)$
−1	0.48	0.08	0.56
1	0.32	0.12	0.44
$\rho(x)$	0.80	0.20	1.00

ϕ) shown in Table 3, we see that this leads D to make correct choices $48 + 12 = 60\%$ of the time and incorrect choices $8 + 32 = 40\%$ of the time: $\gamma(0) = 0.60$.[18]

Given the assumption that P acts in the interests of D, is there anything she can do to improve D's expected outcome via nudging or boosting? The answer is yes. To see this, notice that the rational policy is to always pick the healthy diet, regardless of the value of ϕ, $\rho(x = -1|\phi) > \rho(x = 1|\phi)$. Therefore, nudging by setting $\tilde{\phi} = -1$ ensures that D chooses $\tilde{a}^*(\phi, -1) = -1$, regardless of the value of ϕ (i.e., 100% of the time). This results in 80% correct and 20% incorrect decisions. Indeed, this is the best P can do, provided P does not observe x. Therefore, $\gamma(-1) = 0.80$, which is a substantial improvement over $\gamma(0) = 0.60$.

It is possible for D to be impervious to nudges. To see this, assume $\tilde{b}$ is identical to ρ in every way, except D believes he is much more active than is actually the case, with $\tilde{b}(\tilde{x} = 1) = 0.80$. In this situation, D always orders a large meal, expecting to be pleased 80% of the time and 20% of the time to leave the restaurant feeling bloated from overeating, for a net subjective expected payoff of 0.60. Unfortunately, as indicated by Table 3, this behaviour is actually successful only 20% of the time. D has it backwards and, as a result, engages in behaviour that has a true expected payoff of −0.60. In this case, manipulating $\tilde{\phi}$ has no effect.[19]

[18] A rational D would presumably notice at some point that they are under-performing with respect to their subjective expectations, which imply successful decisions $72 + 18 = 90\%$ of the time.

[19] This mismatch, selecting meals according to $\tilde{b}$ and being surprised at the frequency according to ρ could be interpreted as a form of time-inconsistent preference, another form of bias.

Happily, even in this impervious-to-nudges situation, P has another policy option: implement a boost. Recall, the boost trains D to behave like a rational expected utility maximiser. When $\tilde{b}$ is converted to ρ, performance improves to $\gamma(B) = 0.80$, in accordance with Table 3 because, understanding the true probabilities and contexts, D always chooses $a = -1$. Remember, boosting always eliminates the bias and, therefore, leads to proper expected utility maximisation.

1.3. *Stage Game Result with a Benevolent Policymaker*

We wish to consider stage games in which P is benevolent in the sense that she has preferences that are 'sufficiently' aligned with D's best interests. This language is required due to our assumption that ρ is strictly positive (which rules out perfect alignment, $\rho(x = y) = 1$). The following proposition sums up several policy implications.

PROPOSITION 1. *Assume that D is cognitively biased and define* $\epsilon \equiv 1 - \rho(x = y) > 0$. *Then, the following statements are true:*

(*1*) $\gamma^* = \gamma(B) > \gamma(0)$;
(2) *For* ϵ *sufficiently small,*
  (*a*) *If* $s^*_D(\phi, 0) = k \in \{-1, 1\}$ *for both* ϕ, *then* $s^*_P(\phi) = B$ *for at least one* ϕ;
  (*b*) *If* $s^*_D(\phi, 0) \neq s^*_D(-\phi, 0)$, *then there exists an optimal nudge policy strategy* s^*_P *such that, for each* ϕ, $s_P(\phi) \in \{-1, 1\}$ *and* $s_P(\phi) \neq s_P(-\phi)$.

Statement 1 says boosting is always optimal and that optimal is always strictly superior to the status quo bias. This follows from the definition of bias and because D's optimal performance arises when he maximises expected utility under ρ, which is what happens when D is boosted. Statement (2a) says that, provided that P's expected payoffs are maximised by maximising D's actual expected payoffs (which is the case for ϵ sufficiently small), if D is unresponsive to the context, then any optimal strategy *must* include a boost for at least one context. This is because the assumption of bias implies $a^*(\phi) \neq k$ for at least one ϕ, which cannot be solved by nudges due to D's subjective beliefs being such that he is unresponsive to context.[20] Statement (2b) says that, given sufficient preference alignment, if D is responsive to context, then an optimal strategy can be constructed consisting entirely of nudges. In this last case, keep in mind that such a strategy performs identically to the constant boost strategy. This follows because the assumption of bias and the premise in (2b) imply that D makes objectively suboptimal decisions in both contexts. By manipulating context when D is responsive to it, P can cause D to imitate a rational decision maker without actually boosting.

Proposition 1 points out several key features of the static policy setting. First, boost policies always maximise the welfare of the principal, D. Keep in mind that this strong result is by design: the zero-cost, always-effective nature of boosts is a built-in implication of the model. Proposition 1 formalises this observation.[21] Second, if the bias in play is such that D is unresponsive to the elements of the environment open to manipulation by P, then the *only* way

[20] Note that D can be unresponsive to nudges for two reasons. The first is as a result of viability: i.e., a nudge is declared not viable for a particular context as part of the set-up (an exogenous constraint on manipulating $\tilde{\phi}$, as discussed in Section 1.1.). The second, consistent with the statement (2a), is that $\tilde{b}$ is such that D's natural behaviour does not vary with context.

[21] Keep in mind that the overarching questions at issue here are: (*i*) when do nudge policies do as well as simply fixing the bias?; and (*ii*) given the availability of the always-optimal boost, under what conditions do agency problems arise that lead policymakers to adopt inefficient nudge policies instead?

Table 4. *Joint Probabilities $\rho(x, \phi)$.*

	x		
ϕ	-1	1	$\rho, \tilde{b}$
-1	0.72	0.02	0.74
1	0.08	0.18	0.26
$\rho, \tilde{b}$	0.80	0.20	1.00

to maximise *D*'s welfare is via a policy strategy that includes boosts. This is straightforward. There are a couple of subtleties. One is that, if *D* naturally chooses the objectively optimal action in one context, the boost may not be implemented (i.e., if the non-problematic context is the one that actually arises). The other, related, point is that *P* may implement a nudge policy in the non-problematic context, in which case there will appear to be a correlation between the nudge and successful decision-making even though the nudge in fact has no effect. Finally, when *D* is behaving suboptimally in all contexts and is, simultaneously, responsive to nudges in them, then a nudge policy can be implemented that imitates boosting. In this case, the 'always-nudge' policy is also welfare-maximising.

1.4. *Stage Game Result with Self-Interested Policymaker*

Suppose we relax the assumption that *P* is motivated to act in the best interests of *D*. The first thing to notice is that the mere fact that preferences are not closely aligned is not sufficient to imply policies that are not in *D*'s best interests. Rather, problems arise when the agents' pattern of preferred actions with respect to context are different from one another. From Proposition 1, we know that the boost policy is always effective for achieving the best interests of *D*. Therefore, we wish to characterise the conditions under which the parties' preferences are such that boosts are implemented.

The answer is straightforward: given ϕ, boosting is an optimal choice for *P* if and only if $a^*(\phi)$ maximises their utility. Since $a^*(\phi)$ is optimal for *D*, $a^*(\phi)(2\rho(x = 1|\phi) - 1) > 0$. Given ϕ, $a^*(\phi)$ is also optimal for *P* if and only if $a^*(\phi)(2\rho(y = 1|\phi) - 1) > 0$. Proposition 2 follows immediately.

PROPOSITION 2. *Given ϕ, $s_P(\phi) = B$ is a best reply to s_D^* if and only if*

$$\big(2\rho(x = 1|\phi) - 1\big)\big(2\rho(y = 1|\phi) - 1\big) > 0. \tag{9}$$

To explore this result, let us return to the cafeteria example under the assumptions that: (*i*) $\rho(x = -1) = 0.80$ (*D* is fairly active); and (*ii*) $\rho(\phi = x|x) = 0.90$ (the natural choice of cafeteria type is a good signal of the day's calorie expenditure). In this case, assume that *D* has *no* bias: $\tilde{b} = \rho$. Then, the joint probabilities are as shown in Table 4 for both $\rho(x, \phi)$ and $\tilde{b}(x, \phi)$. *D* co-ordinates dietary choice with cafeteria setting, resulting in $\gamma^* = \gamma(0) = 0.90$. Although *D* is a perfect expected utility maximiser, we continue to assume he is subject to influence via context manipulation.

Now, suppose *P* is heavily lobbied by the National Association of Spinach Growers. As it happens, spinach salad is always a part of *D*'s low-calorie meal, but never eaten as part of a high-calorie meal. As a result, assume *P*'s preference is for *D* to eat 'healthily' at virtually every meal: say, $\rho(y = -1) = 0.99$. By implementing the constant nudge strategy, $s_P = -1$, this goal is achieved. P receives a near-perfect expected payoff of 0.98. Unfortunately for D, this policy results in a substantial performance drop: $\gamma(-1) = 0.80$, which implies a 20% decrease in expected payoff, from 0.80 to 0.60.

Given the premise that P's preferences differ from D's and that D is susceptible to manipulation by P, there is nothing surprising about this outcome. In the extreme case, in which $\rho(y = x) = 0$, P's policy incentive is even more extreme: to nudge D into behaviour opposite to their natural (and correct) inclinations: $s_P(\phi) = -\phi$. It is not obvious how such a policy would actually be accomplished in the context of cafeteria choices but, if it were, the result would be $\gamma(s_P) = 0.10$ along with a further drop in D's expected payoffs, from 0.80 to -0.80.[22]

2. Dynamic Game

Now, let us turn to a dynamic setting in which the preferences of P and D are not aligned. In the dynamic setting there is an important asymmetry between boost and nudge policies. Once boosted, the individual is endowed with a robust ability to make well-calibrated, self-interested choices—irrespective of whether they comport with the policymaker's objectives. This means that once a person has received a successful boost, the nudge associated with the corrected bias loses its effectiveness. This is not true in the reverse case—nudging does not eliminate the effectiveness of future boosts. How does this tension resolve?

2.1. *The Option Value of Nudging*

Assume the stage game described above is repeated indefinitely. We attach t superscripts to the components of the model to indicate the period. In particular, problem ω^t is generated according to ρ, policy p^t is chosen by P and action a^t is taken by D according to $\tilde{b}_{p^t}$. Both ρ and $\tilde{b}$ are exogenously set at the beginning of time. Note that $p^t = B$ implies that $\tilde{b}_{p^k} = \rho$ for all $k \geq t$. Assume D and P share a common *discount factor*, $\delta \equiv (\frac{1}{1+r})$, where $r \in [0, 1]$ is the *discount rate* (note the implication that $0.50 \leq \delta \leq 1$).

Because the dynamic setting is substantially more complicated than the static one, we narrow attention to situations of particular interest to our investigation. Specifically, assume that: (*i*) D is biased (so that P has some purpose as D's policymaking agent); (*ii*) D's subjective preferences lead to actions other than those preferred by P (so P cannot attain their desired outcomes by doing nothing); (*iii*) D is responsive to nudges (so adding nudges to P's policy repertoire is not meaningless); and (*iv*) if P is indifferent to $p(\phi) = B$ and some other policy, they boost (P is benevolent on the margin). Items (*ii*) and (*iii*) imply that P's preferences are such that either $\rho(y = 1|\phi = 1)$, $\rho(y = -1|\phi = -1) > 0.5$ or $\rho(y = -1|\phi = 1)$, $\rho(y = 1|\phi = -1) > 0.5$. Therefore, without loss of generality (given the preceding assumptions), we focus our attention on the latter case.

Since the stage game is now repeated, we must extend our notation for player strategies. Because D suffers from a persistent bias, their beliefs never change unless boosted. Therefore, since D perceives no time-related tradeoffs to their actions, they play their subjectively rational stage game strategy each period. It is helpful to keep track of the state of D's beliefs with the parameter $\theta^t \in \{U, B\}$, where U indicates that 'D is biased' (i.e., a boost was never implemented)

[22] The ability to accomplish context-specific nudges with respect to restaurant choice is perhaps not so outlandish given the prevalence of digital phone technology. For-profit marketing efforts in this area are consistent with such a phenomenon.

and B indicates 'D is *boosted*' (i.e., was subjected to a boost intervention in some previous period during the history of play). Then, $s_D^*(\phi^t, p^t, \theta^t)$ denotes D's subjectively optimal choice of action in period t given the true context (ϕ^t), P's policy choice in the present period (p^t), which may affect their subjective perception of the context, and the present state of D's beliefs (θ^t). If t is the first period in which $p^t = B$, then $\theta^t, \theta^{t+1}, \ldots = B$, at which point P's ongoing policy choices are irrelevant (D chooses rationally henceforth).

Expected payoffs in a particular period depend on the decision problem in that period, ω^t, the players' actions and whether or not D was ever boosted, θ^t. The ω^ts are independently generated across periods according to ρ, and the value of θ^{t+1} depends only on θ^t and p^t. Therefore, this is a stochastic game with finite states and actions. This implies the existence of Markov perfect equilibria, which will be the focus of the following analysis.

Given the persistence of D's bias, it is sufficient for P to build a dynamic strategy on the values of the context variables and the state of D's beliefs, with $s_P(\phi^t, \theta^t)$ indicating their choice of policy given these contingencies in period t. Let the *value function* $V(\phi^t, \theta^t)$ denote the present value to the policymaker facing (ϕ^t, θ^t) given their choice of a dynamically optimal policy strategy, s_P^*. Then,

$$V\left(\phi^t, \theta^t\right) = \max_{p^t} \mathbb{E}\left(\pi_P | \phi^t, p^t\right) + \delta \sum_{\phi^{t+1}} V\left(\phi^{t+1}, \theta^{t+1}\right) \rho\left(\phi^{t+1}\right). \qquad (10)$$

The present value given (ϕ^t, θ^t) is the expected payoff in the present period given ϕ^t and policy p^t plus the expected discounted value of the next period's situation (given play of p^t).

To begin our analysis, suppose nudges are viable. Under what conditions, if any, does P boost? Clearly, if Condition (9) is met for both values of ϕ, then there is no agency problem and P, indifferent between boosting and nudging, boosts according to being marginally benevolent. Alternatively, if Condition (9) fails for both values of ϕ, then P never boosts. Therefore, let us now restrict attention to the non-trivial setting in which the agents' true preferences agree in one context but not the other. Let $\phi^{☺}$ and $\phi^{☹}$ denote the contexts in which D and P do and do not have preference alignment, respectively.

In this case, P never boosts. To see why, begin with the case in which P never boosts. Then, P manipulates D to take P's preferred action in every period (benefiting D in one context, but not the other). Then, by the simplifying assumption of equality of conditional probabilities, the expected payoff to P in all periods is simply

$$\sum_{\phi} |2\rho(y = 1|\phi) - 1|\rho(\phi). \qquad (11)$$

Alternatively, once D is boosted, P's expected payoff in all future periods is

$$\left|2\rho\left(y = 1|\phi^{☺}\right) - 1\right| \rho\left(\phi^{☺}\right) - \left|\rho\left(y = 1|\phi^{☹}\right) - 1\right| \rho\left(\phi^{☹}\right). \qquad (12)$$

Noting that (12) is strictly less than (11), we see that the pure nudge strategy dominates. Proposition 3 summarises this result.

PROPOSITION 3. *If nudges are viable in both contexts and an agency problems exists in one and only one context, then there will be no t for which $s_P^*(\phi^t, \theta^t) = B$.*

This is an important, if discouraging, result. If there is *any* agency conflict between a policymaker and a decision maker who is responsive to nudges, boosting is never implemented. Even if P is indifferent between boosting and nudging in the present period, P will not boost because

Table 5. *Representative Pattern of Action Preferences.*

	P pref		a^*		$\tilde{a}^*$	
	1	−1	1	−1	1	−1
$\phi^{\smiley}$	–	*	–	*	*	–
$\phi^{\frownie}$	*	–	–	*	–	*

boosting destroys the ability to nudge in a future period when faced with the contentious context. The key insight is that *keeping open the ability to nudge in future periods has option value for P.*

The previous result depended on the assumption that nudges were viable in both contexts. If nudging is not viable under any context, then there is nothing more to be said. Therefore, suppose nudging is not viable under one context. The immediate implication is that, to induce *D* to take the preferred action in that context, *P must* boost. The opportunity cost to gaining this immediate benefit is the loss of being able to nudge in future periods. Let us work out the details of this case.

Assumptions (*ii*) and (*iii*) above plus the focus on settings in which the agency problem exists in exactly one context limits the possible preference patterns. One of these is shown in Table 5. The asterisks show the context-contingent preferences of *D*'s action choice according to: (*i*) *P*; *D*, objectively; and (*iii*) *D*, naturally (i.e., $\tilde{a}^* \equiv s_D^*(\phi^t, 0, U)$). Item (*iii*), nudge responsiveness, implies that *D*'s subjective preferences are along one or the other diagonal. Item (*ii*), that *P* and *D* are at odds over which actions are optimal, implies that *P*'s preferred actions lie on the opposite diagonal. Finally, the focus on settings in which agency problems exist for one context but not the other implies that the objectively optimal actions for *D* are in one of the columns (note that this implies the satisfaction of Item (*i*), that *D* is biased). The allowable possibilities are: (*i*) switch the diagonals for *P* and $\tilde{a}^*$; and (*ii*) switch the column for a^*. This results in four symmetric preference patterns. Without further loss of generality, let us proceed by analysing the pattern in Table 5.

Assume $\phi^{\smiley}$ is the context in which nudging is not viable. Then, in any period in which the context is $\phi^{\frownie}$, the solution to (10) is a nudge (*away* from *D*'s optimal action). Suppose *t* is the first period in which the context is $\phi^{\smiley}$. If the solution to (10) is to boost, then *D* makes rational decisions—irrespective of *P*'s policy decisions—thereafter. If the solution is not to boost in *t*, then it is never to boost. Therefore, suppose boosting is not optimal and, since nudges are not viable in this context anyway, that *P* chooses the no-intervention policy, $p_t = 0$. Then,

$$
\begin{aligned}
V\left(\phi^{\smiley}, U\right) &= \mathbb{E}\left(\pi_P \left|\phi^{\smiley}, 0\right.\right) + \delta \sum_{\phi^{t+1}} V\left(\phi^{t+1}, U\right) \rho\left(\phi^{t+1}\right) \\
&= -\left|2\rho\left(y = 1|\phi^{\smiley}\right) - 1\right| + \\
&\sum_{i=1}^{i=\infty} \delta^i \left[\left|2\rho\left(y = 1\left|\phi^{\frownie}\right.\right) - 1\right| \rho\left(\phi^{\frownie}\right) - \left|2\rho\left(y = 1\left|\phi^{\smiley}\right.\right) - 1\right| \rho\left(\phi^{\smiley}\right)\right] \\
&= \frac{\delta\left|2\rho\left(y = 1\left|\phi^{\frownie}\right.\right) - 1\right| \rho\left(\phi^{\frownie}\right) - \left|2\rho\left(y = 1\left|\phi^{\smiley}\right.\right) - 1\right| \rho\left(\phi^{\smiley}\right)}{1 - \delta}.
\end{aligned}
$$

If boosting is optimal, then

$$
\begin{aligned}
V\left(\phi^{\smiley}, U\right) &= \mathbb{E}\left(\pi_P \left|\phi^{\smiley}, B\right.\right) + \delta \sum_{\phi^{t+1}} V\left(\phi^{t+1}, B\right) \rho\left(\phi^{t+1}\right) \\
&= \left|2\rho\left(y = 1|\phi^{\smiley}\right) - 1\right| + \\
&\sum_{i=1}^{i=\infty} \delta^i \left[-\left|2\rho\left(y = 1\left|\phi^{\frownie}\right.\right) - 1\right| \rho\left(\phi^{\frownie}\right) + \left|\rho\left(y = 1\left|\phi^{\smiley}\right.\right) - 1\right| \rho\left(\phi^{\smiley}\right)\right] \\
&= \frac{-\delta\left|2\rho\left(y = 1\left|\phi^{\frownie}\right.\right) - 1\right| \rho\left(\phi^{\frownie}\right) + \left|2\rho\left(y = 1\left|\phi^{\smiley}\right.\right) - 1\right| \rho\left(\phi^{\smiley}\right)}{1 - \delta}.
\end{aligned}
$$

Since this holds for all of the preference patterns allowed under our assumptions (due to their symmetries), we have the following proposition:

PROPOSITION 4. *If t is the first period in which the non-viable context arises, say ϕ, then boosting will be optimal if and only if*

$$
\delta \leq \frac{|2\rho\,(y = 1\,|\phi\,) - 1|\,\rho\,(\phi)}{|2\rho\,(y = 1\,|-\phi\,) - 1|\,\rho\,(-\phi)\,.} \tag{13}
$$

As is conventional in analyses of this kind, Proposition 4 is stated in terms of a condition on the "patience" of the policymaker. Presenting the result in this way highlights the tension faced by the policymaker between cashing in on today's payoff versus maximising expected payoffs in the future. In the case analysed, given $\phi^{\smiley}$, P would like to boost. This, however, imposes an opportunity cost: the inability to nudge D in future periods when $\phi^{\frownie}$ arises. Thus, Condition (13) says that the policymaker will boost when the rate at which they discount future payoffs (r) is 'sufficiently high' (i.e., they place sufficiently greater weight on current than on future payoffs). But how high is 'sufficiently high'? The answer depends on the expected payoff in each context under the optimal policy for that context and the probabilities with which those contexts arise. Notice that, since $\delta \leq 1$, if the numerator in (4)—the probability-weighted expected gain from boosting when the context is $\phi^{\smiley}$—is greater than the denominator—the probability-weighted expected gain from nudging when the context is $\phi^{\frownie}$—then, P boosts at the first instance of $\phi^{\smiley}$ regardless of their patience level. Also, if P is an elected official and δ is interpreted as reflecting the likelihood of being ousted from office sooner rather than later, then Proposition 4 suggests (very loosely) that elected officials who are more worried about their election prospects are more likely to boost. We explore this idea more carefully in the next section.

2.2. *Example 2: Failure to Reason in a Bayesian Fashion*

Now, consider a different cognitive bias, one that may be quite common in the real world and that appears to arise in important situations: the base-rate fallacy. The cognitive problem here is people's apparent difficulty in updating stated probabilities in accordance with Bayes' Rule (1). The base-rate fallacy—meaning that too little weight is placed on base rates of the event in question, relative to the likelihood information—has been documented in numerous experimental studies using written vignettes (e.g., Kahneman and Tversky, 1972; Grether, 1980; Koehler, 1996).[23]

The following details are motivated by a real-life case familiar to the authors. Suppose a decision about medical treatment options is based on the results of a medical test.[24] Specifically, D is an elderly gentleman experiencing symptoms of an intestinal disease that could be colon cancer ($x = 1$) or colitis ($x = -1$). For simplicity, assume that, in the population of patients with these symptoms, the prevalence (base rate) of cancer is $\rho(x = 1) = 0.05$; otherwise the symptoms indicate colitis ($\rho(x = 1) = 0.95$). Here, ϕ represents the results of a sigmoidoscopy, which indicates whether the disease is colitis ($\phi = -1$) or colon cancer ($\phi = 1$). Assume the that the test has a *hit rate* of 90%. That is, the conditional probability that the test indicates disease x when x is the actual disease is: $\rho(\phi = x|x) = 0.90$ for both $x = 1$ and $x = -1$. Further, assume that the *false alarm rate* for both diseases is 10%: i.e., $\rho(\phi = x| - x) = 0.10$ for $x = 1, -1$ (note that, in general, the hit rate and false alarm rate need not sum to 1).

Using Bayes' Rule, this implies that the positive predictive value of the test for a patient with colitis, $\rho(x = \phi|\phi = -1) = 0.99$, is virtually perfect. Unfortunately, the low base rate of cancer implies a high rate of false positives for that disease, with the result that the positive predictive value of a cancer indication by the test is only $\rho(x = \phi|\phi = 1) = 0.32$. Unfortunately, the predictive value is much worse for those with the more deadly disease.

Suppose D can choose to treat his disease using 'natural' methods (diet, homeopathy), $a = -1$, or via conventional medical procedures, $a = 1$. Reflecting a deep aversion to 'Big Medicine', D prefers natural treatments (regardless of their efficacy) for all but the most serious of medical issues. Hence, D would prefer to use natural methods for colitis, but conventional procedures for colon cancer. The preceding assumptions about ρ imply that an unbiased D should always proceed as if he has colitis (use natural methods), irrespective of the test results. This is because the low probability of actually having cancer contributes to a high rate of false positives (9.5% of outcomes) versus true positives (only 4.5% of outcomes).

[23] This bias is also of particular interest to our analysis because there is well-documented evidence of various methods by which people's ability to reason in a Bayesian fashion can be boosted (e.g., Hoffrage *et al.*, 2000; Sedlmeier, 1997; Sedlmeier and Gigerenzer, 2001; Armstrong and Spaniol, 2017). One approach is representation training (e.g., Sedlmeier and Gigerenzer, 2001), which teaches people to use frequency representations rather than to insert probabilities into Bayes' Rule, which is notoriously difficult. Bayesian computations are substantially simpler to perform with natural frequencies. Representation training has been found to have great temporal stability. Alternatively, one might imagine that a better policy is to provide the relevant information (base rate, test results, conditional probabilities) in a table which maps each outcome of, say, a test result for cancer onto the posterior probability of having that disease, along with a copy of the Bayes formula so that individuals have what they need to make the necessary computation. Unlike representation training, however, this approach neither simplifies computation nor provides insight into the formula. More importantly, it does not enable one to transfer a new competence—reasoning in a Bayesian fashion—to new contexts. Rather, it makes one dependent on the information being presented in full, including the posterior probability (often unavailable because experts, such as physicians themselves, often lack the ability to compute the correct Bayesian answers; see Hoffrage *et al.*, 2000; Hoffrage and Gigerenzer, 1998). Furthermore, someone who receives a positive test result is quite likely, in their fearful state of mind, to make mistakes in calculating posterior probabilities, even if they know the Bayes formula; the above policy would reduce the risk of making mistakes.

[24] Traditionally, this was a paternalistic affair: the doctor told the patient what the test result is, what it means, and what do do next. The patient did as they were told. Today, the healthcare environment has profoundly changed. Many patients can and do access information about medical evidence from newspaper articles, books, pamphlets, and the Internet. Unquestioning compliance with doctors' orders has been replaced by notions of *shared decision-making* and *informed consent*, according to which doctors have a duty to fully inform patients about the rationale for their care plan, and decisions are made jointly. (For a brief historical review of these developments see Wegwarth and Gigerenzer, 2013). This new environment raises a number of issues for healthcare policymakers. One is how doctors should best inform their patients about the benefits and harms of medical diagnostics (e.g., the accuracy of cancer-screening tests, the meaning of their results and potential side-effects) and treatments.

Table 6. *Objective (ρ) and Subjective ($\tilde{b}$) Distributions, Base-rate Fallacy.*

x	y	ϕ	$\rho(x, y, \phi)$	$\tilde{b}(x, y, \phi)$
1	1	1	0.009	0.45
1	1	−1	0.001	0.05
1	−1	1	0.00	0.00
1	−1	−1	0.00	0.00
−1	1	1	0.00	0.00
−1	1	−1	0.00	0.00
−1	−1	1	0.099	0.05
−1	−1	−1	0.891	0.45

Further, suppose D's preferences are at odds with the policymaker's due to the effectiveness and costs of conventional medical treatments. Indeed, P operates on the basis of the results of a 2015 comprehensive assessment of evidence by the Australian government, according to which there is no reliable evidence that homeopathy is effective for *any* health condition.[25] For colitis, a well-established pharmaceutical treatment is available which puts the disease into remission 70% of the time, thereby allowing many patients to return to their previous lives. Thus, P prefers this conventional treatment for colitis: $\rho(y = 1|x = -1) = 0.70$. In the other 30% of cases, the drug regimen has no effect. Hence, if x were directly observed, P would prefer conventional treatment when $x = -1$. In the case of colon cancer in older people, P has a strong aversion to conventional treatment: it is very expensive and, even if successful, barely increases a patient's life expectancy. Hence, $\rho(y = 1|x = 1) = 0.01$: P almost always prefers natural methods for treatment of colon cancer in elderly patients like D.

To reflect the base-rate fallacy, we assume $\tilde{b}$ is identical to ρ with the exception of $\tilde{b}(x)$, which we assume is uniform (0.5 probability on both values). This effectively causes D to ignore the base rate of the disease. The behavioural implication is that D responds to the test results as if the probability of having the disease given a positive test result is equal to the probability of a positive test result given the presence of the disease: choose conventional treatment ($a = 1$) if $\phi = 1$ and natural methods otherwise. The objective and subjective distributions are shown in Table 6.

Summing up, according to D's preferences, the objectively optimal choice is to choose natural methods ($a = -1$) irrespective of the test result. Left to make treatment decisions on the basis of his biased subjective beliefs, however, D chooses natural treatment when $\phi = -1$ and conventional treatment ($a = 1$) when $\phi = 1$. P's test-contingent preferences are at odds with both of these: they would prefer that D choose conventional methods when colitis is indicated, $\phi = 1$, and natural methods when cancer is indicated. Hence, referring back to Table 5, $\phi^{☺} = 1$ and $\phi^{☹} = -1$.

Suppose P can implement different nudges for $\phi^{☺}$ and $\phi^{☹}$, respectively. For example, in the former case, P might provide D with slick brochures on the relative ineffectiveness of natural methods versus the relatively safe, symptom-free outcomes of conventional methods. In the latter case, P might go in the opposite direction, vividly describing the pain associated with the conventional approaches, painting palliative care in glowing terms, and drawing the patient's attention to the effects on his loved ones, asking him, 'Do you really want to put *them* through this ordeal?' Assume these nudge techniques—which enlist framing, anchoring and D's tendency to comply with an authority figure—are effective.

25 https://www.nhmrc.gov.au/about-us/resources/homeopathy.

Without viability constraints, s_P^* is: always nudge. Whichever test result arises, nudging yields a positive expected payoff in the current period versus a negative expected payoff from not nudging. Therefore, there is never a reason not to nudge. The expected payoff to P in any given period (i.e., prior to the revelation of ϕ) is 0.34. From D's perspective, s_P^* results in an expected number of correct decisions $\gamma(s_P^*) = 0.10$. This is much worse than simply leaving D to follow his own, biased decision-making: $\gamma(0) = 0.90$. Boosting improves D's performance: $\gamma(B) = 0.95$. Were they to boost, P would experience an expected *loss* per period of 0.33.

Now, assume that P is prevented from nudging in the $\phi^{☺}$ context (cancer indicated). In this case, D's bias results in the mistaken belief that the test's positive predictive value is $\tilde{b}(x = 1|\phi = 1) = 0.90$. Since $x = 1$ is fatal if left untreated, D cannot be dissuaded from conventional treatment via some nudge. In the $\phi^{☺}$ context, a boost would endow D with the ability to make a rational assessment, and thereby afford P (and themself) a superior payoff. The tradeoff, from P's perspective, is being denied the ability to nudge in the $\phi^{☹}$ context. The tradeoff favours boosting when Condition (13) is satisfied:

$$\delta \leq \frac{\left|2\rho\left(y = 1\left|\phi^{☺}\right.\right) - 1\right|\rho\left(\phi^{☺}\right)}{\left|2\rho\left(y = 1\left|\phi^{☹}\right.\right) - 1\right|\rho\left(\phi^{☹}\right)},$$

$$= \frac{|2 \times 0.52 - 1|\,0.14}{|2 \times 0.30 - 1|\,0.86} = 0.02.$$

Since δ must range between 0.50 and 1.00, this condition is never met. Expected losses in $\phi^{☺}$ periods (occurring only 14% of the time) due to the nudge viability constraint are not enough to motivate boosts, which then entail losses in $\phi^{☹}$ periods (occurring 86% of the time).

This extended numerical example brings a qualitative point to light: failing to maximise D's preferences does not necessarily imply unprincipled behaviour on the part of P. This article casts the overarching issue as an agency problem between P and D. This immediately brings to mind the image of P as the self-serving opportunist. In this example, D simply has a preference for passing on conventional medical treatment when the medical issue is not life-threatening. Having a preference is not the same as having defective cognition: D willingly accepts the cost of physical discomfort, etc., implied by this preference. In this light, P's preference for conventional treatment may stem from a true, perhaps even passionate, desire to improve D's life. P may believe that, once treated, D will come around to their point of view. Even so, if P nudges D accordingly, they are not acting in D's interests as defined by D himself.

2.3. *Nudging with Consequences*

Being an agent for the decision maker (e.g., an elected official), P's essential duty is to ensure good choices by inducing decisions that maximise $\mathbb{E}(\pi_D|\phi, a)$. Although we assume that D cannot fix their decision bias independent of an externally administered policy, it is reasonable to imagine that they do have the capacity to notice when, under P's policy interventions, the consequences of their actions are frustratingly inconsistent with their desired outcomes (otherwise, why agree to be nudged by a policymaker in the first place?). Even so, given a context in which a cognitive bias is present, explanations for bad outcomes may be murky. For example, the decision maker may blame themself or some collection of exogenous factors (accidents, bad luck, or animal spirits). However, in the face of persistently disappointing results, the decision maker could well conclude that the problem *does* lie with the policymaker, perhaps due to incompetence or, worse, intentional misdirection. The stronger such a conclusion is, the more likely D is to remove and replace P. The question we explore next is the extent to which the possibility of premature termination acts as a mechanism to ensure good policy implementation by P.

To make things concrete, suppose that P serves for a fixed term of 3 periods. This number is sufficient to illustrate the issues at play when the consequences of P's policies on D have implications for their own payoffs. If, at the end of P's term, D is satisfied with their performance (in a sense described below) then P receives an end-of-term bonus payment worth $L > 0$. This may be interpreted simply as a bonus payment for a job well done or as the expected net present value of payments associated with being appointed (or elected) for another term.

Let λ^t keep track of D's cumulative payoff through the end of period t, with $\lambda^0 \equiv 0$. As long as P nudges with no prior history of a boost, then $\lambda^t = \lambda^{t-1} + \pi_D^t$.[26] The idea is that, although D may not have a sophisticated system of beliefs with which to evaluate P's performance, they can keep track of their payoffs to obtain a sense of whether P's policies, although not immediately obvious to D, help. Once a boost is implemented, any lack of transparency is removed. D becomes cognisant of the good intentions of P, regardless of the random vagaries of their own actual outcomes. That is, following a boost in t, $\lambda^t = \lambda^{t-1} + 1$ and similarly in all future periods. For our purposes, assume that the bonus L is paid to P if $\lambda^{t=3} \geq 0$ and not otherwise.

To proceed, let us focus on a special class of situations—those in which nudges are viable in both contexts and in which the following symmetries hold:

$$\rho_x^* \equiv \rho(x = 1|\phi_1) = \rho(x = -1|\phi_{-1}) > 0.5, \text{ and}$$

$$\rho_y^* \equiv \rho(y = -1|\phi_1) = \rho(y = 1|\phi_{-1}) > 0.5.$$

Thus, D maximises their expected payoff by matching their action to the context, while P does so by getting D to do the opposite. By equalising the context-contingent probabilities of x and y, we eliminate a substantial amount of algebraic manipulation in what follows without loss of the essential insights into situations of this kind.

Given these simplifications, the expected payoff to P from nudging can be defined as $\eta \equiv (2\rho_y^* - 1)$ and the probability that, following a nudge, λ^t increases by one can be defined as $\alpha \equiv (1 - \rho_x^*)$. Then, the expected, within-period payoff to nudging is η and to boosting is $-\eta$, regardless of the context. Therefore, P's value function (10) in period t can be written as $V(t, \theta^t, \lambda^t)$. The effect of their decision on the present value of their payoff stream depends on the period, whether D was previously boosted, and the cumulative consequences of their policies on D's payoffs. Specifically,

$$V(t, \theta^t, \lambda^t) = \begin{cases} \eta + \delta\Big[\alpha V(t+1, U, \lambda^t + 1) + (1-\alpha)V(t+1, U, \lambda^t - 1)\Big], & \text{if } \theta^t = U, p^t = N \\ -\eta + \delta V(t+1, B, \lambda^t + 1), & \text{otherwise} \end{cases}. \quad (14)$$

Since the bonus payment happens at the end of period 3 (equivalently: the beginning of period 4), we set $V(4, \theta^4, \lambda^4) = L$ if $\lambda^4 \geq 0$ and 0 otherwise. The game ends immediately following the bonus payment.

Figure 1 illustrates the decision tree for P. Decision nodes are squares, random nodes are circles and terminal nodes are diamonds. The expected payoffs associated with each decision and terminal outcome are shown in the ovals. Recall that, once the boost has been selected (B), future policy choices are rendered irrelevant.

[26] Keep in mind, $\pi_D^t = 1$ or -1.

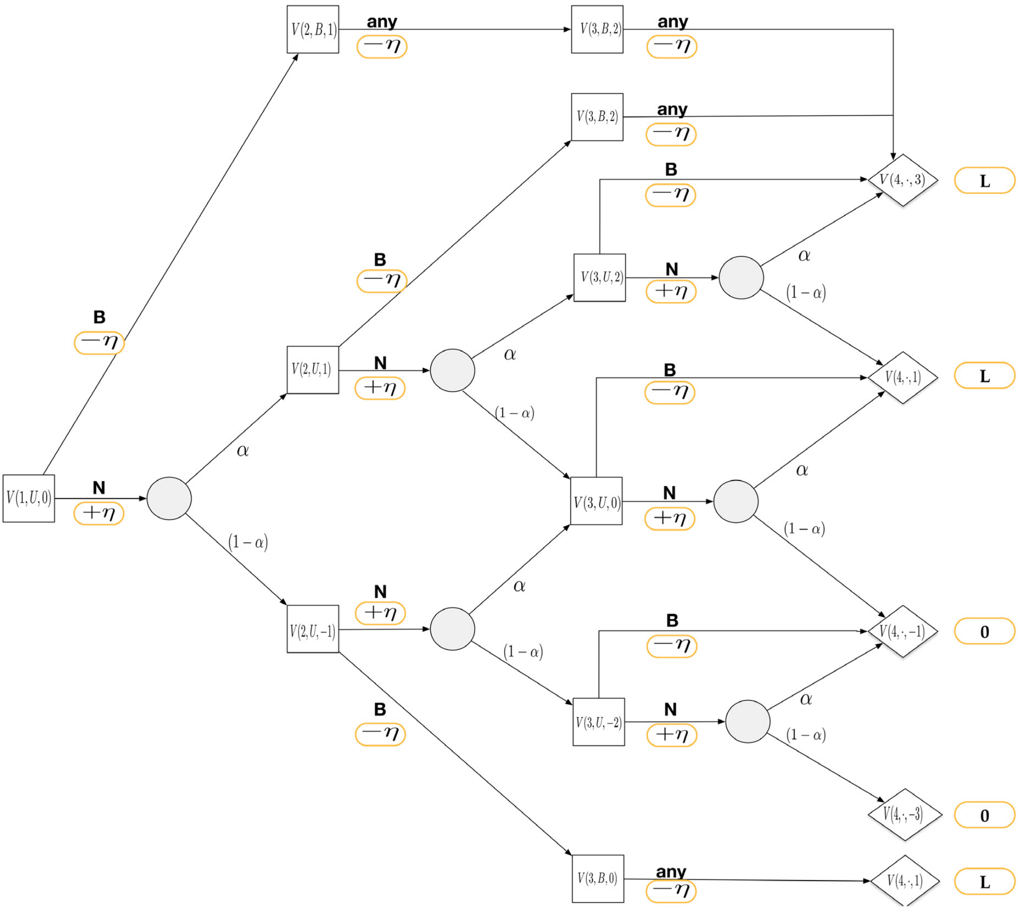

Fig. 1. *P's Decision Tree.*

The problem is solved via backwards induction, starting with the choice faced at $t = 3$. As it turns out, of the six possible decision nodes at which P may find themself, only one involves a non-trivial choice. To see this, note that three involve value functions of the form $V(3, B, \cdot)$, at which P's choice of policy no longer matters because D has been boosted. Therefore, at these nodes, we know $V(3, B, \cdot) = -\eta + \delta L$. At $V(3, U, 2)$, whether P chooses B or N, a final payoff of L is guaranteed (because $\lambda = 2$, the final value is either 3 or 1—either of which results in the bonus being obtained). Therefore, since B and N imply expected payoffs of $-\eta$ and η, respectively, the optimal choice is N. This implies $V(3, U, 2) = \eta + \delta L$. At $V(3, U, -2)$, P cannot obtain the bonus no matter the policy choice. This, again, implies an optimal choice of N and, therefore, $V(3, U, -2) = \eta$.

The only choice requiring some algebra is the node at which the value function is $V(3, U, 0)$. At this node, implementing B results in an expected payoff of $-\eta + \delta L$. Alternatively, N results in $\eta + \delta[\alpha L + (1 - \alpha)0]$. Therefore, B is optimal if $-\eta + \delta L \geq \eta + \delta \alpha L$. Rearranging terms,

$$\delta(1 - \alpha)L \geq 2\eta, \text{ or} \tag{15}$$

$$L \geq \frac{2\eta}{\delta(1-\alpha)}. \tag{16}$$

The interpretation of (15) is straightforward: P should boost if the expected gain from doing so in terms of the discounted, expected terminal payoff is greater than or equal to the forgone, within-period payoff. Alternatively, as stated in (16), the boost is chosen if L is 'sufficiently large' (i.e., relative to the other parameters). Either way, the balance is tipped towards a boost when P is more patient (larger δ), the probability of getting L under a nudge (α) is smaller and the expected payoff from nudging (η) is larger. Thus,

$$V(3, U, 0) = \begin{cases} -\eta + \delta L, & \text{if } L \geq \dfrac{2\eta}{\delta(1-\alpha)} \\ \eta + \delta\alpha L, & \text{otherwise} \end{cases}. \tag{17}$$

This completes the analysis of period 3, with closed-form solutions computed for the value functions at all six nodes. Proceeding to period 2, computing all of the value functions associated with its nodes, and then finishing with period 1 results in a decision strategy for P which specifies the optimal policy at every node in the tree. Relegating these additional details to the Appendix, the following proposition summarises this strategy.

PROPOSITION 5. *P's optimal decision strategy for nodes at which $\theta^t = U$ is:*

Period 1: Nudge
Period 2: Nudge at $V(2, U, 1)$. Boost at $V(2, U, -1)$ if and only if:

$$\delta^2(1-\alpha)L \geq (1 + \delta(1-\alpha))2\eta. \tag{18}$$

Period 3: Nudge at $V(3, U, 2)$ and $V(3, U, -2)$. Boost at $V(3, U, 0)$ if and only if:

$$\delta(1-\alpha)L \geq 2\eta. \tag{19}$$

Readers may notice that the decision tree in Figure 1 exhibits the classic structure of a finite-length real option. As always, nudging keeps open the option of nudging or boosting in the future. Boosting closes off the option of nudging. What Proposition 5 says is that the spectre of D imposing a substantive punishment on P, should the accumulation of bad outcomes become too great, can, indeed, induce them to boost in some dynamic situations (rather than none, as demonstrated in the previous section). Such boosting occurs when the accumulation of negative results: (*i*) makes the bonus sufficiently unlikely under a continued nudge regime; and (*ii*) is not yet so great as to swamp the boost's ability to induce that bonus. As long as D happens to experience good outcomes despite being nudged against their interests, P will continue to nudge them against their interests.

3. Policymakers Are Human Too

The preceding analysis examines nudge versus boost policies in settings where agency issues may exist between policymakers and decision makers. As such, it departs from the usual line of work in this area by relaxing the assumption of preference alignment between the two parties. While nudges that are not in the individual's best interests may be driven by malicious or self-serving intent, misaligned preferences may arise from an honest disagreement over what constitutes the individual's best interests. As discussed at the end of the preceding example, P may be motivated by an earnest belief that D, in some sense, has the 'wrong' preferences for their own good. In this section, we examine another reason why the policymaker may fail to implement efficient nudges—namely, that they suffer from cognitive biases of their own (e.g., self-enhancement bias, hindsight bias, anchoring bias, overconfidence bias, false-consensus effect; Krueger and Funder, 2004; Pronin and Schmidt, 2013).[27]

Suppose we now eliminate divergent preferences by eliminating y and assuming both players receive payoffs equal to $a^t x^t$ in period t. This implies ρ is a distribution on the four combinations of (x, ϕ). This situation is complicated by the fact that we must now entertain the possibility that P's bias may extend to their assessment of D's behaviour in response to the context ϕ and their policy choice p. Let $\tilde{b}_P$ denote P's subjective beliefs, where $\tilde{b}_P(x, \phi)$ is their assessment of the probability that (x, ϕ) is generated by Nature at the start of the period, and (abusing notation) $\tilde{b}_P(a|\phi, p)$ is their belief that D chooses a given context ϕ and policy p. Then, given their observation of ϕ (the observed context) and policy choice of p, they believe that the relevant outcome (a, x) occurs with probability $\tilde{b}_P(a, x|\theta, p) = \tilde{b}_P(a|\phi, p)\tilde{b}_P(x|\phi)$.[28] Now, P's subjective expected, context-contingent stage game payoff for choosing p is

$$\tilde{\mathbb{E}}(\pi_P|\phi, p) \equiv \sum_{(a,x)\in\{-1,1\}^2} ax\tilde{b}_P(a, x|\phi, p). \tag{20}$$

Again, in keeping with the spirit of this literature, we assume that $\tilde{b}_P$ is robust to learning in response to evidence.[29]

At this level of generality, it should be clear that the analysis can be taken in any number of directions.[30] Here, because this is not the primary focus of the article, we limit ourselves to a couple of immediate points. One important class of situations is that in which P suffers from the same bias as does D. These are settings in which P is both: (*i*) well calibrated with respect to D's behaviour, $\tilde{b}_P(a|\phi, p) = 1$ if $a = s_D^*(\phi, p)$ and 0 otherwise; and (*ii*) shares D's bias, $\tilde{b}_P(x, \phi) = \tilde{b}_D(x, \phi)$, where $\tilde{b}_D$ now indicates D's biased subjective beliefs. Hearkening back to the previous example, both D and P may suffer from the failure to reason in a Bayesian fashion (e.g., adjust for the base rate of an event). In such cases, doing nothing ($p = 0$) is always optimal: bias alignment leads P to conclude that D chooses optimally—without any need for intervention on their part.

Second, compare P's policies in earlier sections (preference divergence) with those here (cognitive bias). In the earlier setting, P chooses p to maximise (6). For the moment, assume that, in the present setting, D is well calibrated with respect to D's behaviour (as described above). Then, P maximises (20), which can be rewritten

$$\begin{aligned}\tilde{\mathbb{E}}(\pi_P|\phi, p) &= s_D^*(\phi, p)\big(\tilde{b}_P(x = 1|\phi) - \tilde{b}_P(x = -1|\phi)\big),\\ &= s_D^*(\phi, p)\big(2\tilde{b}_P(x = 1|\phi) - 1\big).\end{aligned}$$

Therefore, the set of optimal policies for P in each setting are equal when

$$s_D^*(\phi, p)\big(2\tilde{b}_P(x = 1|\phi) - 1\big) = s_D^*(\phi, p)\big(2\rho(y = 1|\phi) - 1\big), \text{ or}$$
$$\tilde{b}_P(x = 1|\phi) = \rho(y = 1|\phi).$$

[27] We thank an anonymous referee for encouraging us to consider this issue.

[28] It is worth noting that this is in the form of a *subjective environment response function* (Kalai and Lehrer, 1993): P observes ϕ, takes policy action p and experiences the payoff consequence (a, x) with probability $\tilde{b}_P(a, x|\theta, p)$.

[29] Hence, evidence-based learning results, such as those presented in Kalai and Lehrer (1993), do not hold.

[30] A thorough analysis of joint biases between policymaker and subject warrants at least an entire article.

In other words, for every situation in which a socially inefficient nudge policy is implemented due to divergent interests in the earlier setting, there exists, in theory, a cognitive bias in the well-intentioned policymaker in this setting that also induces the implementation of the same inefficient policy. By 'situation' we mean that the $\rho(x, \phi)$s in each setting are equal. Thus, the results of the previous sections do not vanish under the presumption of beneficent policymaking, provided one is willing to attribute cognitive biases to decision makers whose decisions represent policy (see discussion in Grüne-Yanoff and Hertwig, 2016).

4. Conclusions

The nudge approach to public policy (Thaler and Sunstein, 2008) is important not only as a compelling call to policymakers to take evidence from psychology and behavioural economics seriously but also, perhaps even more so, because it advances a conceptual framework for doing so: libertarian paternalism. The myriad examples of real-world policymakers around the globe now adopting this framework speak to its inherent attractiveness. This popularity notwithstanding, more evidence is needed on the efficacy of constructing social policy around citizens' assumed cognitive biases and motivational deficits (House of Lords, 2011). At a higher level still, issues surrounding the ethics of such policies—and, relatedly, the agency problems that may arise—need to be addressed. Given that numerous policymakers have already adopted the libertarian paternalism framework, it is crucial that these issues be tackled in a timely manner.

This article explores the issue of agency at the level of the policymaker, assuming the existence of readily available and inexpensive policy options with the capacity to eliminate constituents' biases and deficits. This approach places the agency problem in stark relief: empowering individuals to liberate themselves from biased judgement is an uncontroversial social good. However, as we show, the asymmetric nature of boosting versus nudging—in the sense that nudging always leaves open the option for boosting, but not vice versa—means that nudging has option value to the policymaker. Thus, a sufficiently patient policymaker may prefer not to eliminate a bias, even when doing so would be in their immediate interests. A policymaker who forgoes boosting secures future opportunities to employ nudges. The risk is that, in such situations, a libertarian paternalist leans much more heavily in the direction of paternalism.

Our analysis assumes that either a nudge or a boost can be employed in the same setting; otherwise, the policymaker would not face a (strategic) choice between the two types of intervention. This condition applies in numerous real-world settings, such as our examples of medical treatments and cafeterias (see also Dallacker *et al.*, 2019); it also applies in other real-world domains that involve, for instance, risk literacy, financial literacy and health literacy (for examples see Hertwig and Grüne-Yanoff, 2017). Yet, admittedly, there will also be numerous cases in which only one type of intervention is applicable (for examples see Hertwig, 2017; Sunstein, 2014; Walton, 2014) such as maths anxiety (Berkowitz *et al.*, 2015).

Furthermore, the alternative to a nudge such as automatic enrolment in a savings plan is sometimes not a boost but an intervention such as 'active choice'; e.g., requiring individuals to make an explicit choice for themselves (Carroll *et al.*, 2009), or enhanced active choice (Keller *et al.*, 2011). Boost and active choice, unlike nudging, aim to return agency to the decision maker—but active choice does so without first boosting a specific consequence. Some of the agency issues we considered here may also arise in the decision between, say, a default (automatic enrolment) and active choice. Once active choice has been implemented in a domain, it may prove impossible to take this 'right' (obligation) back; consequently, nudging would become mute.

Last but not least, let us point out that, despite the widespread appeal of nudging, there are also limits (see Hertwig, 2017). For instance, it is hard to imagine how without empowering people one could offer lasting and robust remedies to the problems of bias, intentional misinformation and micro-targeting presently faced by today's media consumers. Indeed, this is one area in which 'commercial' nudging is actually the source of the problem (see, e.g., Groseclose, 2011). Equipping citizens to make judgements of information quality independently and competently—despite the manipulation efforts—is indispensable to maintaining democratic forms of government (see Groseclose, 2011; Cook *et al.*, 2017; van der Linden *et al.*, 2017; Lazer *et al.*, 2018).

Finally, our results raise a number of related issues. For example, a common complaint in Western democracies is that politicians, because they seek re-election every few years, consider only the short-term consequences of policies. This implies that the discount rate of elected policymakers is very low indeed. However, in our setting, a short-run horizon on the part of the policymaker may counter-intuitively imply the adoption of policies that are actually efficient in the long term—with boosted constituents becoming *somebody else's* problem. Another issue is the existence or development of policy interventions that overcome biases. The discovery of competence-boosting and bias-correcting mental tools implies the creation of markets for such services. What are the policy dynamics in the presence of privately offered services? Which segments of the population are more likely to seek and pay for such services? What are the policy implications when libertarian paternalism affects some segments more than others? How would one interpret the imposition of regulations defining 'quality requirements' on educational programmes designed to overcome the very biases enlisted by paternalistic policymakers? These are but a handful of novel and interesting issues raised by the current analysis and results.

Appendix A

A Proof of Proposition 2

PROOF. Recalling (2) and that D is assumed to be a subjective expected utility maximiser (plays S_D^*), the expected payoff to D given ϕ when boosted is

$$\mathbb{E}(\pi_D|\phi, B) = a^*(\phi)\big[\rho(x=1|\phi) - \rho(x=-1|\phi)\big], \tag{A1}$$

$$= \big|\rho(x=1|\phi) - \rho(x=-1|\phi)\big|, \tag{A2}$$

with the second step arising because a^* matches the sign of the quantity in brackets. The expected payoff to P given a boost, then, is equal to (A1) when $y = 1$ and is the opposite of it when $y = -1$. That is

$$\mathbb{E}(\pi_P|\phi, B) = \mathbb{E}(\pi_D|\phi, B, y=1)\rho(y=1) - \mathbb{E}(\pi_D|\phi, B, y=-1)\rho(y=-1), \tag{A3}$$

where $\mathbb{E}(\pi_P|\phi, B, y) = \big|\rho(x=1|y,\phi) - \rho(x=-1|y,\phi)\big|$. Expanding (A3) delivers the result.

Alternatively, the result may be demonstrated in a more tedious but direct manner as follows. Under s_P^*, P's expected payoff to boosting is

$$\mathbb{E}(\pi_P|\phi, B) = a^*(\phi)\big[\rho(y=x|\phi) - \rho(y\neq x|\phi)\big]. \tag{A4}$$

When P's preferences are sufficiently aligned with D's, then it is in their interest to implement a boost. Alignment occurs when the bracketed terms in both (A1) and (A4) have the same sign.

That is, given ϕ, P and D prefer the same action to be taken if and only if:

$$[\rho(x=1|\phi)-\rho(x=-1|\phi)]\geq 0 \text{ and } [\rho(y=x|\phi)-\rho(y\neq x|\phi)]\geq 0, \text{ or} \tag{A5}$$
$$[\rho(x=1|\phi)-\rho(x=-1|\phi)]\leq 0 \text{ and } [\rho(y=x|\phi)-\rho(y\neq x|\phi)]\leq 0.$$

Let $L\equiv\rho(x=1,y=1|\phi)-\rho(x=-1,y=1|\phi)$ and $K\equiv\rho(x=1,y=-1|\phi)-\rho(x=-1,y=-1|\phi)$. Then, (A5) corresponds to

$$L\geq\max\{-K,K\}\geq 0, \text{ or} \tag{A6}$$
$$L\leq\min\{-K,K\}\leq 0,$$

which implies the single condition $|L|\geq|K|$. Substituting back in,

$$\big|\rho(x=1,y=1|\phi)-\rho(x=-1,y=1|\phi)\big|\geq\big|\rho(x=1,y=-1|\phi)-\rho(x=-1,y=-1|\phi)\big|. \tag{A7}$$

Applying Bayes' Rule delivers the result. □

Appendix B

B Proof of Proposition 5

PROOF. The value functions for $t=3$ are all developed in the text, with the single non-trivial case, $V(3, U, 0)$, elaborated in (17). □

Decisions in period 2

There are three nodes in the decision tree at $t=2$, one each associated with $V(2, B, 1)$, $V(2, U, 1)$ and $V(2, U, -1)$.

There is no meaningful policy decision for P at the boosted-state node, the value of which is

$$V(2,B,1)=-(1+\delta)\eta+\delta^2 L. \tag{B1}$$

Next, consider the decision when the value function is $V(2, U, 1)$; here, the value of boosting is:

$$V(2,U,1)=-(1+\delta)\eta+\delta^2 L, \tag{B2}$$

in which case, $V(2, U, 1) = V(2, B, 1)$. Alternatively, P can nudge. The continuation value then depends on Condition (19): $L\geq\frac{2\eta}{\delta(1-\alpha)}$. Specifically,

$$V(2,U,1)=\eta+\delta[\alpha V(3,U,2)+(1-\alpha)V(3,U,0)]$$
$$=\begin{cases}[1+\delta(2\alpha-1)]\eta+\delta^2 L, & \text{if } L\geq\dfrac{2\eta}{\delta(1-\alpha)}\\ (1+\delta)\eta+\delta^2(2\alpha-\alpha^2)L, & \text{otherwise}\end{cases}. \tag{B3}$$

Working through the algebra, if Condition (19) holds, i.e., $L\geq\frac{2\eta}{\delta(1-\alpha)}$, then P boosts if $-2(1+\alpha\delta)\eta\geq 0$. Since this is impossible, P does not boost. Suppose Condition (19) fails. Then, again omitting the algebraic steps, P boosts if

$$L\geq\left(\frac{2\eta}{\delta(1-\alpha)}\right)\left(\frac{1+\delta}{\delta(1-\alpha)}\right).$$

Since the second term on the right-hand side is greater than 1, this condition is ruled out by the premise that $L < \frac{2\eta}{\delta(1-\alpha)}$. Therefore, we conclude that P never boosts at $V(2, U, 1)$; hence, its value is as described by (B3).

Now, consider $V(2, U, -1)$. If a boost is implemented, we immediately have $V(2, U, -1) = V(2, B, 1)$. If a nudge is implemented instead,

$$V(2, U, -1) = \eta + \delta[\alpha V(3, U, 0) + (1-\alpha)V(3, U, -2)]$$
$$= \begin{cases} [1+\delta(1-2\alpha)]\eta + \delta^2\alpha L, & \text{if } L \geq \dfrac{2\eta}{\delta(1-\alpha)} \\ (1+\delta)\eta + \delta^2\alpha^2 L, & \text{otherwise} \end{cases}. \tag{B4}$$

If Condition (19) holds, then P boosts if

$$L \geq \left(\frac{2\eta}{\delta(1-\alpha)}\right)\left(\frac{1+\delta(1-\alpha)}{\delta}\right). \tag{B5}$$

The second term on the right-hand side is greater than 1. Therefore, this condition requires that L be strictly larger than required by Condition (19). Suppose Condition (19) fails. Then, P boosts if

$$L \geq \left(\frac{2\eta}{\delta(1-\alpha)}\right)\left(\frac{1+\delta}{\delta(1+\alpha)}\right). \tag{B6}$$

Once again, the term on the right-hand side is greater than 1. Therefore, by the premise that Condition (19) fails, this condition cannot be satisfied: P will not boost in this case.

Decision in period 1

As always, the value of the boost option is straightforward:

$$V(1, U, 0) = -(1+\delta+\delta^2)\eta + \delta^3 L. \tag{B7}$$

There are now three cases to consider, as implied by the previous findings:

Case 1: Condition (B6) is satisfied.
Case 2: Condition (B6) is not satisfied and Condition (19) is satisfied.
Case 3: Neither Condition (B6) nor Condition (19) is satisfied.

In Case 1, following some algebraic manipulation, it can be shown that P boosts if

$$-2[1+\alpha\delta+(1-\alpha-\alpha^2)\delta^2] \geq 0.$$

Since the quantity on the left-hand side is strictly negative, P never boosts in this case. In Case 2, it can be shown that P boosts if

$$L \geq \frac{2(1+\delta+\delta^2)+\delta^2(1-2\alpha)}{\delta(1-\alpha)^2}\eta$$
$$= \left[\frac{2(1+\delta)(1+r)}{\delta^2(1-\alpha)(1-\alpha)} + \frac{2}{1-\alpha}\right]\eta.$$

Noting that

$$\frac{2(1+\delta)(1+r)}{\delta^2(1-\alpha)(1-\alpha)} > \left(\frac{2\eta}{\delta(1-\alpha)}\right)\left(\frac{1+\delta}{\delta(1+\alpha)}\right),$$

the condition for boosting fails by implication in the premise that Condition (B6) fails. This leaves Case 3. Here, P boosts if

$$L \geq \frac{2(1+\delta+\delta^2)}{\delta^3(1-\alpha^2(3-2\alpha))}\eta, \\ = \frac{2(1+\delta+\delta^2)(1+r)^3}{1-\alpha^2(3-2\alpha)}\eta. \tag{B8}$$

Subtracting Condition (19) from Condition (B8) yields

$$\left[\frac{2\left((1+r)^3+(1+r)^2+(1+r)\right)}{1-\alpha^2(3-2\alpha)} - \frac{2(1+r)}{1-\alpha}\right]\eta.$$

Solving this equation for the minimum of the η coefficient over $\alpha \in [0, 0.5]$, $r \in [0, 1]$ yields a solution of $\sim$19, which is strictly positive. Therefore, Condition (B8) cannot be satisfied under the premise of Case 3. We conclude: P never boosts in $t = 1$.

Max Planck Institute for Human Development
University of Toronto

References

Allcott, H. (2011). 'Social norms and energy conservation', *Journal of Public Economics*, vol. 95(9), pp. 1082–95.

Anonymous. (2006). 'Soft paternalism: the state is looking after you', *The Economist Online*, April 6th ed.

Armstrong, B. and Spaniol, J. (2017). 'Experienced probabilities increase understanding of diagnostic test results in younger and older adults', *Medical Decision Making*, vol. 37(6), pp. 670–9.

Benabou, R. and Tirole, J. (2003). 'Intrinsic and extrinsic motivation', *The Review of Economic Studies*, vol. 70(3), pp. 489–520.

Berg, N. and Gigerenzer, G. (2010). 'As-if behavioral economics: Neoclassical economics in disguise?', *History of Economic Ideas*, vol. 18(1), pp. 133–66.

Berkowitz, T., Schaeffer, M.W., Maloney, E.A., Peterson, L., Gregor, C., Levine, S.C. and Beilock, S.L. (2015). 'Math at home adds up to achievement in school', *Science*, vol. 350(6257), pp. 196–8.

Besley, T., Montalvo, J.G. and Reynal-Querol, M. (2011). 'Do educated leaders matter?', Economic Journal, vol. 121(554), pp. F205–27.

Besley, T. and Reynal-Querol, M. (2011). 'Do democracies select more educated leaders?', *American Political Science Review*, vol. 105(3), pp. 552–66.

Bovens, L. (2009). 'The ethics of nudge', in (T. Grüne-Yanoff and S. O. Hansson, eds.), *Preference Change*, pp. 207–19, Dordrecht: Springer.

Cadario, R. and Chandon, P. (2019). 'Which healthy eating nudges work best? A meta-analysis of field experiments', *Marketing Science: Articles in Advance*, pp. 1–22.

Carlin, B., Gervais, S. and Manso, G. (2013). 'Libertarian paternalism, information production, and financial decision making', *Review of Financial Studies*, vol. 26(9), pp. 2204–28.

Carroll, G.D., Choi, J.J., Laibson, D., Madrian, B.C. and Metrick, A. (2009). 'Optimal Defaults and Active Decisions', *The Quarterly Journal of Economics*, vol. 124(4), pp. 1639–74.

Cialdini, R.B. (2009). *Influence: Science and Practice*, 5th ed., Boston: Allyn & Bacon.

Cook, J., Lewandowsky, S. and Ecker, U.K.H. (2017). 'Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence', *PloS One*, vol. 12(5), p. e0175799.

Costa, D. and Kahn, M. (2013). 'Energy conservation "nudges" and environmentalist ideology: evidence from a randomized residential electricity field experiment', *Journal of the European Economic Association*, vol. 11(3), pp. 680–702.

Dallacker, M., Mata, J. and Hertwig, R. (2019). 'Toward simple eating rules for the land of plenty', in (R. Hertwig, T.J. Pleskac, T. Pachur and C.A. Rationality, eds.), *Taming Uncertainty*, pp. 111–30, Cambridge: MIT Press.

Fang, H. and Moscarini, G. (2005). 'Morale hazard', *Journal of Monetary Economics*, vol. 52(4), pp. 749–77.

Galasso, V. and Nannicini, T. (2011). 'Competing on good politicians', *American Political Science Review*, vol. 105(1), pp. 79–99.

Galasso, V. and Nannicini, T. (2015). 'So closed: Political selection in proportional systems', *European Journal of Political Economy*, vol. 40(3), pp. 260–73.

Gervais, S., Heaton, J. and Odean, T. (2011). 'Overconfidence, compensation contracts, and capital budgeting', *The Journal of Finance*, vol. 66(5), pp. 1735–77.

Gigerenzer, G., Hertwig, R. and Pachur, T. (2011). *Heuristics: The Foundations of Adaptive Behavior*, Oxford: Oxford University Press.

Grether, D. (1980). 'Bayes rule as a descriptive model: The representativeness heuristic', *The Quarterly Journal of Economics*, vol. 95(3), pp. 537–57.

Groseclose, T. (2011). *Left Turn: How Liberal Media Bias Distorts the American Mind*, St. Martin's Press.

Grüne-Yanoff, T. and Hertwig, R. (2016). 'Nudge versus boost: How coherent are policy and theory?', *Minds and Machines*, vol. 26(1–2), pp. 149–83.

Halpern, D. (2016). *Inside the Nudge Unit: How Small Changes can Make a Big Difference*, Random House.

Hertwig, R. (2017). 'When to consider boosting: some rules for policy-makers', *Behavioural Public Policy*, vol. 1(2), pp. 143–61.

Hertwig, R. and Grüne-Yanoff, T. (2017). 'Nudging and boosting: Steering or empowering good decisions', *Perspectives on Psychological Science*, vol. 12(6), pp. 973–86.

Hoffrage, U. and Gigerenzer, G. (1998). 'Using natural frequencies to improve diagnostic inferences', *Academic Medicine*, vol. 73(5), pp. 538–40.

Hoffrage, U., Lindsey, S., Hertwig, R. and Gigerenzer, G. (2000). 'Communicating statistical information', *Science*, vol. 290, pp. 2261–2.

House of Lords. (2011). *Behavior Change (HL Paper 179)*, House of Lords.

Johnson, E.J. and Goldstein, D. (2003). 'Do defaults save lives?'.

Kahneman, D. and Tversky, A. (1972). 'Subjective probability: A judgment of representativeness', *Cognitive Psychology*, vol. 3(3), pp. 430–54.

Kalai, E. and Lehrer, E. (1993). 'Rational Learning Leads to Nash Equilibrium', *Econometrica*, vol. 61(5 G), pp. 1019–45.

Keller, P.A., Harlam, B., Loewenstein, G. and Volpp, K.G. (2011). 'Enhanced active choice: A new method to motivate behavior change', *Journal of Consumer Psychology*, vol. 21(4), pp. 376–83.

Koehler, J. (1996). 'The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges', *Behavioral and Brain Sciences*, vol. 19(1), pp. 1–53.

Koriat, A. and Bjork, R.A. (2005). 'Illusions of competence in monitoring one's knowledge during study', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 31(2), p. 187.

Krueger, J.I. and Funder, D.C. (2004). 'Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition', *The Behavioral and Brain Sciences*, vol. 27(3), pp. 313–27.

Kuehnhanss, C., Heyndels, B. and Hilken, K. (2015). 'Choice in politics: Equivalency framing in economic policy decisions and the influence of expertise', *European Journal of Political Economy*, vol. 40, pp. 360–74.

Lazer, D.M.J., Baum, M.A., Benkler, Y., Berinsky, A.J., Greenhill, K.M., Menczer, F., Metzger, M.J., Nyhan, B., Pennycook, G. and Rothschild, D.and Others(2018). 'The science of fake news', *Science*, vol. 359(6380), pp. 1094–6.

Lourenco, J.S., Ciriolo, E., Almeida, S.R. and Trousssard, X. (2016). 'Behavioural insights applied to plicy: European Report 2016 (EUR 27726 EN)', European Commission Joint Research Centre, Brussels.

Marteau, T., Hollands, G. and Fletcher, P. (2012). 'Changing human behavior to prevent disease: the importance of targeting automatic processes', *Science*, vol. 337(6101), pp. 1492–5.

Nickerson, R.S. (1998). 'Confirmation bias: A ubiquitous phenomenon in many guises', *Review of General Psychology*, vol. 2(2), p. 175.

OECD. (2017). *Behavioural Insights in Public Policy: Key Messages and Summary*, Paris: The Organisation for Economic Co-operation and Development.

Pronin, E. and Schmidt, K. (2013). 'Claims and denials of bias and their implications for policy', The behavioral foundations of public policy, pp. 195–216.

Rebonato, R. (2012). *Taking liberties: A Critical Examination of Libertarian Paternalism*, Palgrave Macmillan.

Sargent, T. and Wallace, N. (1975). '"Rational" Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule', *The Journal of Political Economy*, vol. 83(2), pp. 241–54.

Schnellenbach, J. and Schubert, C. (2014). 'Behavioral Politiacal Economy', CESifo Working Paper no. 2698.

Sedlmeier, P. (1997). 'BasicBayes: A tutor system for simple Bayesian inference', *Behavior Research Methods, Instruments, & Computers*, vol. 29(3), pp. 328–36.

Sedlmeier, P. and Gigerenzer, G. (2001). 'Teaching Bayesian reasoning in less than two hours', *Journal of Experimental Psychology: General*, vol. 130(3), p. 380.

Sugden, R. (2008). 'Why incoherent preferences do not justify paternalism', *Constitutional Political Economy*, vol. 19(3), pp. 226–48.

Sunstein, C. (2014). *Why Nudge?: The Politics of Libertarian Paternalism*, Yale University Press.

Sunstein, C. and Thaler, R. (2003). 'Libertarian paternalism is not an oxymoron', The University of Chicago Law Review, Oct 1, pp. 1159–202.

Sunstein, C.R. (2013). *Simpler: The Future of Government*, New York: Simon and Schuster.

Tasic, S. (2011). 'Are Regulators Rational?', *Journal des Economistes et des Etudes Humaines*, vol. 17(1).

Thaler, R. and Sunstein, C. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven: Yale University Press.

Tversky, A. and Kahneman, D. (1974). 'Judgment under uncertainty: Heuristics and biases', *Science*, vol. 185(4157), pp. 1124–31.

van der Linden, S., Leiserowitz, A., Rosenthal, S. and Maibach, E. (2017). 'Inoculating the public against misinformation about climate change', *Global Challenges*, vol. 1(2), p. 1600008.

Walton, G.M. (2014). 'The new science of wise psychological interventions', *Current Directions in Psychological Science*, vol. 23(1), pp. 73–82.

Wegwarth, O. and Gigerenzer, G. (2013). 'Trust-your-doctor: A social heuristic in need of a proper environment', in (R. Herwig and U. Hoffrageand the ABC Research Group (institution/Organization), eds.), *Simple Heuristics in a Social World*, pp. 67–1023, New York: Oxford University Press.

World Bank. (2015). *World Development Report 2015: Mind, Society and Behaviour*, World Bank, Washington, D.C.