King's Research Portal

Document Version
Peer reviewed version

[Link to publication record in King's Research Portal](#)

LIBERTARIAN PATERNALISM AND HEALTH CARE POLICY:
A DELIBERATIVE PROPOSAL

Authors: Giuseppe Schiavone, Gabriele De Anna, Matteo Mameli, Vincenzo Rebba, Giovanni Boniolo

Abstract: Cass Sunstein and Richard Thaler have been arguing for what they named libertarian paternalism (henceforth LP). Their proposal generated extensive debate as to how and whether LP might lead down a full-blown paternalistic slippery slope. LP has the indubitable merit of having hardwired the best of the empirical psychological and sociological evidence into public and private policy making. It is unclear, though, to what extent the implementation of policies so constructed could enhance the capability for the exercise of an autonomous citizenship. Sunstein and Thaler submit it that in most of the cases in which one is confronted with a set of choices, some default option must be picked out. In those cases whoever devises the features of the set of options ought to rank them according to the moral principle of non-maleficence and possibly to that of beneficence. In this paper we argue that LP can be better implemented if there is a preliminary deliberative debate among the stakeholders that elicits their preferences, and makes it possible to rationally defend them.

INTRODUCTION

In a series of publications, R. Thaler and C. Sunstein have argued in favor of what they call "libertarian paternalism" (hereafter, LP).[1] According to this approach, it is possible to reconcile respect for individuals' freedom of choice with paternalistic interventions aimed at making individuals behave in health-promoting (and, more generally, welfare-promoting) ways. This can be done by designing choice situations in ways that—given individuals' psychological tendencies, limitations and biases— are apt to generate health-promoting (and, more generally, welfare-promoting) choices. Individuals retain the freedom to choose unhealthy (or self-harming) options, thereby making this kind of paternalism *liberty-preserving*. This suggestion has prompted a long-standing debate, partly reviving the old debate about the legitimacy of paternalism in public policy.

In what follows, we first present a detailed account of LP. Hence we discuss some of the objections with which LP has been charged. We conclude that those charges are not fatal to the theory, but they point to some serious shortcomings for which we eventually suggest a 'deliberative' solution.

WHAT IS LIBERTARIAN PATERNALISM

Neutral designs and the argument from beneficence

In their book—Nudge (2008, 6)—Thaler and Sunstein start by pointing out that the assumption that people typically make choices that are in their best interest is wrong. In many circumstances, they already noted, "individuals make inferior choices, choices that they would change if they had complete information, unlimited cognitive abilities, and no lack of willpower" (Thaler and Sunstein 2003, 175). This includes circumstances where the choice impacts the health of the individual in question. Such a remark was not novel in 2008, when the book was first published. Economists and psychologists have been studying the limits to individual decision-making for quite some time and a sub-discipline has developed around this pursuit.[2] The novelty brought about by *Nudge* lies in fact in the suggestion that, given what we know about the aforementioned limitations, we ought to make wise (and good) use of that knowledge. In particular, one should exploit biases and restraints in order to make people better off, *as judged by themselves*. In this respect the authors put forward an argument from beneficence. Of course, there is a series of different ways of influencing people's choices in order to increase their welfare, i.e. LP is not the only instance of paternalism, but apparently knowledge of irrational (or partly rational) behavioral patterns allows for an uncoercive kind of paternalism.

---

[1] See Thaler and Sunstein (2003, 2008) and Sunstein and Thaler (2003).

[2] See Kahneman and Tversky (2000) and Gilovich et al. (2002).

What Sunstein and Thaler call a *choice architect*—i.e. anyone who has responsibility for organizing the context in which people make decisions—may, of course, be unaware that their (meta-)choices will have an impact on the decision of choice makers. But, if they are aware of this fact, they always have roughly five options available:[3]

> *1. Maximizing welfare, all things considered* They can design the choice situation with the aim to increase the number of decisions that promote (or are thought to promote) the health and welfare of the choice makers.

> *2. Deliberately diminishing people's welfare* They can design the choice situation with the aim to encourage decisions that go against (or are thought to go against) the health and welfare of the choice makers.

> *3. Maximizing profit* (or other self-interested values) They can design the choice situation on the basis of considerations that have nothing to do with the health and welfare of the decision makers (e.g. profit considerations).

> *4. Accommodating people's existing preferences* They can arrange the situation so as to get people to choose what they would have chosen anyway.

> *5. Randomly picking architectures* They can design the choice situation randomly.

Option 1 is the LP option. The proponents of LP argue that any other option is undesirable (Sunstein and Ullmann-Margalit 1990), hence suggesting that, despite a degree of intrusiveness in option 1, we are left with no better strategy. Indeed, aiming at producing decisions that go against the welfare and health of choice makers (option 2) seems to be patently immoral. In case a choice architect chooses option 3, she will exploit her knowledge at the expenses of people's welfare, or health, which also seems morally problematic for people employed in public institutions. Option 4 looks like the most promising, and, in Sunstein and Thaler's words, it would "be an honorable attempt to avoid intrusion" (Thaler and Sunstein 2008, 2). Their argument against option 4 goes as follows: the ease with which people change their behavior, namely changing it according to the architecture in place, makes it hard to tell the 'true' preferences apart, hence reducing the likelihood that attempts at avoiding intrusion result in any sensible arrangement. Finally, option 5, which apparently seems to provide a neutral solution, turns out to be undesirable at a slightly deeper inspection. Any random attribution will in fact make those randomized to some specific arrangement worse off with respect to the ones randomized to other—better—arrangements.

To sum up, given the inevitability of the impact of meta-choices on (first-order) choices (i.e. there is no such thing as neutral design), not using meta-choices to encourage choices that are health-promoting and welfare-promoting (which would be the case in options 2, 3, 4 and 5) always seems morally unsatisfactory. The principle of beneficence[4] together with the (generally shared) assumption that an omission which results in harm is just as blameworthy as any direct maleficent action, are enough grounds to argue that not (meta)acting is not a good option either[5] in this case. The claim that having knowledge about people's behavior imposes a moral burden on architects, then, seems like a fair point. The authors concede that there is a relevant sense in which option 1 is intrusive, and paternalistically so. Intrusiveness, though, comes in degrees, and does not entail full-blown coercion. Hence choice architects can implement the first option respecting freedom of choice. Sunstein and Thaler provide grounds for why we ought to endorse a paternalistic stance when certain epistemic conditions obtain (we know that there are architectures that would make people better off as judged by themselves), and argue furthermore that in those situations we have the means to leave individual liberty formally untouched. A policy, thus, counts as *libertarian paternalistic* if, while refraining from reducing the array of choices people have, it is selected with the goal of influencing the choices of affected parties in a way that will make those parties better off, "as judged by themselves" (Thaler and Sunstein 2008, 5). The intrusion proposed in option 1, hence, is not to be considered coercive as long as it does not

---

[3] Actually, it could be objected that there is a 6th option, that is, 'maximizing choice'. However we think that Sustein and Thaler maintained that this one has to be considered as the default option.
[4] "Principle of beneficence refers to a statement of moral obligation to act for the benefit of others. Many acts of beneficence are not obligatory, but some forms of beneficence (…) are obligatory." (Beauchamp and Childress 2006: 197).
[5] On omission, see Boniolo and De Anna (2006).

preclude some options altogether. Also, it is not to be regarded as illiberal (or non-libertarian) unless it changes significantly economic incentives.[6]

The epistemic conditions for LP

"Choice architects" make decisions about choice situations. That is, they make meta-choices. These decisions require that some epistemic conditions obtain. Specifically, people must be prone to an array of well-known biases that range from the status quo biase and default options bias (Samuelson and Zeckhauser 1988; Kahneman et al.1991; Korobkin 1998),[7] to framing effects, e.g. anchoring (Strack et al. 1988) and reversals of preference (Tversky and Kahneman 1981), to overconfidence (Thaler and Sunstein 2008, 31–33), hyperbolic discounting and self-control problems (Thaler and Sustein 2008, 40–42).[8]

The biases just mentioned do not, by any means, exhaust the number of ways in which individual decision-making can be affected by human laziness and limitations. Any situation in which some kind of cognitive bias can operate is a good candidate for the deployment of LP measures. These biases impair people's decision-making in a number of ways and in a series of actual situations. The intuition underpinning LP is that people wouldn't consent to choices and behaviors that they end up enacting due to the significant ways in which they cognitively lack and hence adjustments meant to counteract those biases are not only, at least *prima facie*, unrestrictive of individual liberty, but they also embody one's own set of preferences better than one's own behavior does. Such adjustments constitute the toolbox of LP policy-making.

OBJECTIONS TO LP

Sunstein and Thaler's proposal did not go unnoticed or unquestioned. In particular, M.J. Rizzo and D.G. Whitman have forcefully argued against the suggestion that we could and should harness public policies to knowledge from behavioral economics. Their arguments have been rehearsed several times in the literature.

Let us go through the main objections and see how they contribute to clarifying LP. The clarification will serve the purpose of introducing our critique, which will tackle one specific deficiency one can charge LP with. Such objections could be clustered around the following conceptual axes:

> *1*. Policies' interference with self-debiasing and learning processes.

> *2*. Likelihood of the slippage from the new (mild) forms of paternalism to full-blown paternalism.

> *3*. Policy-makers' inadequacy.

> *4*. (At the deepest philosophical level) LP's lack of libertarianism.

> *5*. Choice of the default rule.

Let us briefly review each of the points in turn.

Policies' interference with self-debiasing and learning processes

Different arguments have been proposed in support of LP and new paternalism in general. One of these has a root clearly economic in nature. In standard politico-economic terms, whenever individual behaviors have costs that 'spill over' onto third parties, i.e. they produce *externalities*, the government

---

[6] This latter condition is not well-defined. Any change introduced in an architecture changes somehow costs and incentives. Sunstein and Thaler themselves claim, in note, that some of their own nudges alter incentives. Thus, they loosen their own definition saying that "[n]udges count as such, and qualify as libertarian paternalism, only if any costs are low" (Thaler and Sunstein 2008, 8). This adjusted definition, though, lacks any precise reference to a cutoff that would be needed in order to define which costs count as low.
[7] According to Thaler and Sustein, the combination of loss aversion with mindless choosing may explain behavioral inertia. It follows that "if an option is designated as the 'default', it will attract a large market share" (Thaler and Sunstein 2008, 35).
[8] For a survey on hyperbolic discounting see Frederick et al. (2002).

has some entitlement to intervene. On the contrary, when individuals bear all costs and benefits of their own actions, there is little ground for government intervention.

New paternalists, however, argue that there is a special kind of "within-person externalities", dubbed *internalities*, at stake when people engage in self-harming behavior of specific kinds.[9] Basically, an internality occurs when a person underweights or ignores a consequence of her own behavior for herself. If one uses the language of multiple selves, present selves impose costs on future selves and this grounds the same kind of governmental intervention that are seemingly non-controversial for normal externalities. Adopting this particular perspective, some economists (e.g. O'Donoghue and Rabin 2003) have developed a theory of "sin taxes", i.e. taxes which counteract over-consumption by consumers with self-control problems while at the same time redistribute income to consumers with no self-control problems.

As pointed out by Whitman (2006) and Sugden (2008), internalities are about avoiding the more serious harm between the one currently inflicted to future selves and the one that would be inflicted to present selves in case paternalist policies would be put in place. The seriousness of the harm one wants to avoid is to be evaluated in terms of how inefficiently one situation allocates resources. In Sugden's words (2008, 232):

> […] before we can conclude that a paternalistic third party could do better, we need to know (or rather, the third party needs to know) which preferences reflect well-being and which do not. Is the person's well-being greater if he saves or if he spends? If he insures or if he doesn't? If he diets or if he indulges? […]

The main insight coming from this shift in perspective is that there is the chance that "some, though probably not all, of the present self's future costs will already have been internalized through *intrapersonal bargains*" (Whitman 2006, 12. Emphasis added). If this is the case, i.e. that internalities coming from present selves' behaviors are already accounted for via intrapersonal bargaining, it makes little sense to nudge people into making decisions that go to the exclusive advantage of their future selves. This would result in the underconsumption of some goods, or in the reduction of future selves' welfare in case present selves can defer costs to future selves (e.g. resorting to credit). On a somewhat similar ground, Sugden argues that "a person can have a clear understanding of his interests, sufficient for an appraisal of the market, even if his preferences are incoherent" (Sugden 2008, 239), and concludes that (247):

> We may have good reason to support many of the kinds of regulation that soft paternalists advocate—for example, standardized labeling of food products, standardized tariff structures to facilitate price comparisons, or mandatory cooling-off periods before sales of financial products are confirmed. But those reasons depend on specific and contestable hypotheses about consumers' information-processing capacities, or their susceptibility to psychological cues which induce (what are construed as) distortions of judgment.

We believe that the issue raised by Whitman and Sugden is a serious one, that relates back to the one main problem of LP, namely that it fails to provide reasons why some specific set of individual preferences is to be considered preferable to others. Further on, we will delve more in depth in the epistemic unreasonableness that underlies the assumption that there is an a priori way to adjudicate between competing orders of preferences which grounds the claim that one order is more 'real' than some other. Then we will sketch one way to fix this problem that does not seem to be unavoidably entrenched in LP.

Slippery slopes

A slippery slope argument (SSA) is any line of reasoning that objects to practice X on the grounds that X may *naturally progress* to Y, which everyone finds objectionable already. The naturalness of the progression is what makes a SSA plausible. B. Williams famously set the one condition that makes SSAs worth pursuing. Between the practices X and Y, there must be ways of either *reasonably* or *effectively* fetch a distinction (Williams 1995). Without such a distinction, no relevant SSA can be put forth. X and Y, hence, are two distinct practices, but, so the argument goes,

---

[9] The term "internality" was introduced by Herrnstein et al. (1993).

the distinction might not be reasonably defended or it might not, "as a matter of social or psychological fact" (Williams 1995, 214) be effectively grounded, i.e. if X is admitted, then Y will. In essence, the claim that X will lead to Y entails, according to Williams, that X and Y be theoretically, or, as a matter of fact, distinguishable. Thus, one has to show that the two practices involved in the argument are distinguishable before she can proceed further to the claim that X will likely lead to Y.

Proceeding from this remark, Rizzo and Whitman (2009) suggest that LP-like and full-fledged paternalist policies are distinguishable but lie along the same continuum that goes from non-coercive to coercive means of welfare-enhancing political control. "The libertarian paternalist insists on preserving choice, whereas the non-libertarian paternalist is willing to foreclose choice. But in all cases, a real question is the cost of exercising choice, and here *there is a continuum* rather than a sharp dichotomy" (Sunstein and Thaler 2003, 1185, our emphasis). The very existence of this gradient is a sensible starting point for a thorough reflection upon how slippery the slope of LP actually is, and towards which end it goes. The distinction between LP and other—non-libertarian—forms of paternalism exists, but where exactly a line can be traced is debatable. Thus there is room for SSAs that ought to make clear whether and how a distinction between the two, in specific cases can reasonably or effectively be defended.

Rizzo and Whitman further claim that the slope, the gradient between LP and non-LP, is deemed to be pushing policies towards the paternalist, rather than the libertarian end of the spectrum. First, they make one, seemingly plain, but especially interesting remark, grounding their concern that policy-making is not always as promptly revisable as it is sometimes presented to be. Slippery slopes tend to involve multiple actors. Despite the uncontroversial nature of this observation, the fact, the authors claim, goes often unnoticed. The multiplicity of the actors can be of two main kinds: we name these two kinds of multiplicity vertical (or synchronic) and horizontal (or diachronic). The former refers to the difference between policy-makers and the people who are affected by their policies: the two sets are rarely, if ever, coextensive. The latter multiplicity refers to one central feature of most liberal democratic systems: policy-makers change over time. This entails that whenever slippery-slopes skeptics argue that one can do what is right today while refraining from doing what is wrong tomorrow, they will have to concede that, most likely, they will not be in the position to refrain from doing anything tomorrow. This makes the increased likelihood of undesired consequences of a non-worrisome policy at least slightly more concerning. Again, though, this is merely a good case for the existence of a slope. As it stands, it gives no clue about the direction towards which one is likely to slip.

Rizzo and Whitman do address specifically what they call the 'expansive tendencies' of LP and provide a list of the several ways in which LP policies could lead to more and more intrusive interventions. Going through all of the cases would amount to a different paper. Suffice it to say that there is a main thrust shared by all the criticisms put forth: Scientific beliefs, available facts, political reasons and justifications can and do change over time and in the process of change they are subject to many sorts of unforeseen (and sometimes unforeseeable) distortions that, this they claim, have the tendency of allowing for stronger interventions. We believe that, although the existence of conditions for the expansive tendency just sketched is well-argued, when it comes to the empirical support that the actual expansion has, the arguments proposed are lacking. In order to support our claim we provide one empirical objection to one of their points.

Rizzo and Whitman (2009) make reference to a paper by Besharov (2002). In this article, the author shows that if one tries to ameliorate cognitive biases that interact with each other, she has to consider the chance that piecemeal corrections might be detrimental to the agent. He backs up his argument with the theory of the second-best. Briefly,

> "the General Theorem of the second-best optimum states that if there is introduced into a general equilibrium system a constraint which prevents the attainment of one of the Paretian conditions, the other Paretian conditions, although still attainable, are, in general, no longer desirable" (Lipsey and Lancaster 1956, 11)

Taking this theorem at face value, issues arise in case one attempts at improving people's decision-making via the amelioration of some, *but not all*, of the biases. In other words, it might be the case that co-existing biases offset one another, and hence fiddling with only some of them might tip the balance towards a lower level of welfare. Besharov's argument is pretty technical. Rizzo and Whitman, though,

go slightly further than he does and claim that given the likelihood of complex interactions between biases, there is a significant chance that LP policies that aim at correcting one bias, will end up trying to correct other biases in order to prevent the aforementioned welfare-reducing consequences. This use of the argument however, suffers from two main defects: (1) it offers no reason why perfectly rational policy-makers would decide to go for piecemeal corrections, given that they know there exist second-best-like problems; (2) it overlooks the essentially 'exploitative' nature of LP. Biases are not meant by LPs to be paternalistically corrected or amended, but rather to be paternalistically ridden, so to say. Whereas in the former case there might be space for unexpected consequences of the kind Besharov envisages, in the latter, we tend to believe there is not. In fact, LP advocates try and make use of two pieces of relevant knowledge: (a) the knowledge that some people, when a certain set of conditions (i.e. a choice architecture), be it C_1, is given, tend to do X, which is, in some sense,[10] undesirable; (b) the knowledge that when a different set of conditions, say C_2, surrounds a similar kind of decision, then people tend to do/choose/pick Y, which, in some sense, they hold to be desirable. LPs, apparently, do not want to make people aware of the fact that their behaviors are so and so contextually determined (which would be tantamount to correcting or debiasing the behaviour). They rather want to create sets of conditions which allow Y to happen. One thing must be conceded to LP: it has been factually very successful in setting the features that make up C_2-like architectures.[11] In order to show that the conditions Rizzo and Whitman envisage obtain, they will have to provide evidence that it has happened, at least sometimes, somewhere, that *exploiting* (and not amending) one or more biases has resulted in welfare-reducing consequences due to some unfortunate second-best choice.

Further, opponents of LP have to provide historical proof that LP policies, within democratic societies ended up prompting more intrusive paternalistic measures.[12] We maintain that there currently is no empirical evidence supporting either of the claims.

Policy-makers' inadequacy

Another crucial issue is related to the choice of the nudgers (government officials or other "choice architects"). Society has the legitimate expectation that nudgers will not nudge people in directions that benefit some particular interest group (the group of nudgers included) instead of the whole polity. Furthermore nudgers are expected, when deciding concerning meta-choices, to be both informationally and cognitively unbiased. Thaler and Sunstein recognize that "government officials, elected or otherwise, are often captured by private-sector interests whose representatives are seeking to nudge people in directions that will specifically promote their selfish goals" (Thaler and Sunstein 2008, 243) and to overcome this problem they suggest monitoring and increasing transparency (e.g. requiring government officials to put all their votes, earmarks, and contributions from lobbyists on their websites). Quite surprisingly, however, the problem of "bad" nudgers is somewhat underrated and very quickly dismissed by Thaler and Sunstein. The authors focus almost exclusively on transparency and accountability in this respect, subscribing to the motto according to which "sunlight is the best of disinfectants" (Thaler and Sunstein 2008, 243). We think instead that the possibility that self-interested nudgers hijack nudges requires careful consideration. Specifically, we maintain that mechanisms to ascertain and monitor the competence and expertise (besides the moral integrity) of the nudgers are required and must be differentiated according to the different domains of LP policies. The deliberative solution we propose in the last part of our contribution aims precisely at addressing these concerns. Before proceeding to this final part, let us point to the one genuine shortcoming of LP.

What preferences are 'real'?

---

[10] We will get to this point further on.

[11] Default options for organ donations are just the most astounding success story (Johnson and Goldstein 2003).

[12] A good example of the exploitative nature behind LP can be found in Dan Ariely's Predictably Irrational (2009). Although Ariely is not to be counted among LP advocates, the body of knowledge he is a contributor to, is the same to which LPs make appeal. Ariely describes what he calls the decoy effect: when people are presented with alternatives, they generally fancy having comparables. This results in astounding outcomes. When presented with three options A, B, (differently featured) and—A (practically identical to A, but with some slight deficiencies) people tend to choose A more than if they were only presented with A and B. This happens also when people have to choose a partner. Ariely ends up saying that "if you compared with a "–you," the decoy friend will do a lot to make you look better, not just in comparison with the decoy but also in general, and in comparison with all the other people around. It may sound irrational (and I can't guarantee this), but the chances are good that you will get some extra attention" (Ariely 2009, 15). This trivial example suggests that whoever is in the position to use the relevant knowledge, he/she should do so to his/her own advantage, rather than trying to debias themselves or help others be debiased.

The claim that LP is a liberty-preserving way of enhancing people's welfare, *as judged by themselves* relies on a series of assumptions. One of these assumptions is particularly debatable: That there are epistemically reliable tools to judge the attainment of one's welfare. This assumption requires in turn that one is capable of listing people's 'real' preferences.[13] LP theorists readily admit that this is a problematic task. Sunstein and Thaler argue that "if the arrangement of the alternatives has a significant effect on the selections the customer makes then their true preferences do not formally exist" (2003, 1164). As Thaler and Sunstein show, it is the case that the arrangement of alternatives does influence choices in a number of contexts. Hence, the use of revealed preferences, i.e. sets of preferences inferred from people's behavior, is to be rejected. Still, new paternalists need some account of welfare that grounds the claim that one's welfare is increased by his/her own standards when LP policies are enacted. Instead of attempting to determine people's actual preferences, Thaler and Sunstein try and solve the issue claiming that one's 'real' preferences are those he would entertain under ideal and non-biased conditions. Grüne-Yanoff (2012) argues that this attempt could amount to two different proposals: (1) it could prompt an effort to provide actual people with actually better information, offering time and training to reflect upon their decision and watch how people behave or (2) it could refer to the counterfactual reconstruction of people's preferences had they "complete information, unlimited cognitive abilities, and no lack of self-control" (Sunstein and Thaler 2003, 1162). LP opts for (2), further arguing that, so conceived, one's 'real' preferences would be his/her future self's preferences. However, LP's attempt at reconstruction suffers from at least two theoretical issues. As Grüne-Yanoff puts it (2012):

> *(a)* It is unclear what the three counterfactual conditions would amount to: What is complete information? What does "unlimited cognitive abilities" precisely mean? And how much willpower should one have in order not to lack self-control? In fact, "standards of complete information are as pluralistic as values are, and hence do not constitute a universal criterion for judging preferences" (2012, 642)

> *(b)* It is unlikely that any account provided for these three conditions will result in a robustly predictive tool for the change in individual preferences brought about by the addition of particular quantities of information, cognitive ability and willpower.

Points (a) and (b) firmly back up our claim that LP suffers from a serious epistemic deficiency: LP is unable to articulate the reasons why specific sets of preferences are picked (over others). On the one hand the evidence that supports LP claims undermines the existence of revealed preferences. On the other, counterfactually reconstructed preferences are deemed to make unwarranted assumptions relative to information, cognitive abilities and willpower. Hence, the counterfactual reconstructive attempt they envisage is not grounded and the claim that long-term selves' preferences are to be preferred needs better arguments.

We maintain, and in the next section we argue, that, as compared to LP strategies, non-reconstructive and deliberative means for the assessment of people's preferences fare much better in two essential respects: (1) they require that people (or some subset within the groups of stakeholders) be *actually* asked about their preferences, and (2) they do not hold true any assumption concerning intra-subjective bargains.

Epistemic and pragmatic reasons for deliberation in health care policies

We have shown that the only problem that severely affects LP is its incapability to articulate the reasons that ground the choice of one intrapersonal set of preferences (that belonging to future selves) over some other. We hereby propose a deliberative solution to this troubling issue. Specifically, we provide epistemic and pragmatic reasons why our (deliberative) solution is superior to that LP deploys (i.e. counterfactual reconstruction). A similar solution to a narrower—but related—problem (that of obesity), has been put forward by Anand and Gray (2009). We will expand the scope of their argument to health policies, broadly understood.[14]

Epistemic superiority

In the following paragraph we mean to show that endorsing a deliberative stance concerning health policy-making affords some epistemic gains and is hence a solution epistemically superior to LP. This superiority derives from one, central feature that deliberation as a tool, obtains: it allows for citizens to probe their own positions' internal consistency. Deliberative methods, while not discounting the local knowledge only accessible to people *there and then* (which is instead the case for LP's reconstructed preferences) make it possible, via the rational defense of one's preferences over choice architectures, the discovery of the rationality behind one's set of values. This is a genuine epistemic result, unattainable via LP reconstructive efforts.

We must stress that our contention here is merely that deliberative processes afford a genuine epistemic result with respect to preferences over choice architectures, i.e. preferences over meta-choices. Here we do not discuss the possibility that reasons giving and heeding might result in a better knowledge of people's preferences over choices (which, as we have seen, is a controversial claim both in case one refers to revealed and to reconstructed preferences). Whereas liberalism imposes extreme caution in the domain of personal choices (at least of those personal choices that do not impose burdens on others) choice environments (at least those choice environments that do not significantly restrict the array of options available to anyone) look like a sensible target for public deliberation. This entails that it is reasonable to deliberate about choice architectures, even though deliberation about one's preferences over choices would be insensible. Such a consideration is even more important because currently there seems to be an unjustified lack of democratic control within the sphere of meta-choices. "Critically, although consumers choose*particular items*, conditional on the choice environment, they have no choice over the choice environment itself" (Anand and Gray 2009, 184). Ironically, whereas people's specific preferences (i.e. preferences over choices) are held dear in the face of their alleged irrelevance as means of political justification, people are politically very badly positioned to defend their preferences over meta-choices.

Summing up, we have argued that, exposing the internal consistency of people's positions, an ideal conversation would help constructing (or better unveiling) tenable (meaning reasonable) personal preferences over meta-choices.

> A well-conducted deliberative process is never unsuccessful, because the rational discussion by itself leads participants to reflect upon the tenability of their own and their opponents' positions. In short, there is always a deliberative result beyond any desired consensus. (Boniolo 2012, 20).

Furthermore, deliberation delivers such an epistemic advantage on top of the solution to the lack of democratic control over the selection of choices environments.

To this last, pragmatic, issue we now turn.

Before beginning let us remind that deliberation is the procedure through which people, starting from different initial positions, try and reach a shared decision via a debate based on rational arguments and counter-arguments.[15]

Pragmatic superiority

We have seen that revealed preferences do not formally exist and that LP's reconstructive efforts have proven unwarranted. Henceforth we will argue that, given the epistemic difficulty just summarized, deliberation is a viable procedure to politically legitimize health policies.

Isaiah Berlin famously fetched a distinction between positive liberty (roughly the freedom to lead some prescribed form of life) and negative liberty (roughly the freedom not to have interferences) (Berlin 1969). Under a positive conception of liberty, the ideals of freedom and equality in freedom are fully realized in the kind of political participation that deliberation demands. Specifically, if freedom is self-realization and self-realization is conceived as entailing the pursuit of joint projects,[16] then deliberation is clearly a good way (though not the only one) to promote democratic ideals of liberty. On

---

[15] For a discussion on deliberative democracy and its story, see Boniolo (2012).

[16] This is what an advocate of a version of positive freedom like John Dewey would claim. Whatever pursuit one might engage is, it is valuable because it is the realization of a project-for-us and not because it achieves some self-interested goal. "[D]emocracy […]is primarily a mode of associated living, of conjoint communicated experience" (Dewey 1934, 87).

the other hand, if we endorse a negative stance on the definition of liberty, one might wonder whether and to what extent deliberation and participation can help achieving the ideal standards of freedom and equality in freedom. Kitcher argues deliberation does so in a way that is unachievable through standard tools for political participation, i.e. majority rules and polls. Deliberation promotes the ideals of freedom and equality in freedom, preventing what he dubs *unidentifiable* oppression (see Kitcher 2001, 2011).

As Kitcher points out, historically speaking, classic democratic tools, like majority rules, have served the purpose of annihilating identifiable forms of oppression, in order to extend negative freedom to oppressed groups. Identifiable forms of oppression are the tangible forms of oppression by means of which some specific groups in society are prevented from enjoying negative freedom. Nonetheless, when it comes to contrasting covert, i.e. *unidentifiable*, forms of oppression, subtle limitations on liberty that are not felt, and are difficult to trace back to specific sources, one cannot rely solely on punctual solutions, nor on individual initiatives that would restore equality in freedom. Instead, continuous spread participation in the process of democratic deliberation is a way to keep in check democracy itself, making sure that no hidden source of oppression is at work anywhere in society. Hence, also under a negative conception of liberty, deliberation offers an edge over standard forms of democratic participation. Deliberation applied to LP-justified policies would constitute a substantial attempt at democratically legitimize those policies, thereby contrasting the unidentifiable forms of oppression to which democracy itself is liable, and offering an effective remedy to the issue of 'bad' nudgers highlighted above.

LP policies are clearly—at the very least potentially—subject to the exercise of covert forms of oppression. Even though LP logically entails no identifiable oppression against any particular liberty or group (in fact, the very core of the idea of LP is that of not impinging upon anyone's negative freedom more than it already is being impinged upon), the institution of policies informed only by expert policy-makers looks politically suspicious for two main reasons: (1) the proper choice of the experts is unregulated and, even if it were, (2) the choice itself seems to stand in the face of the ideal of equality in freedom which ought to be promoted by democratic institutions.[17] One risk that might be elicited by the involvement of lay citizens in the process of deliberating about meta-choices is that their preferences concerning architectures of choices are uninformed. However, through an appropriate process of tutoring of people's preferences,[18] the deliberation could both build upon the best knowledge available and be sensitive to people's actual preferences, i.e. the decision will not assume anything about people's conception of welfare. In this respect, the "as judged by themselves" clause put forth by LP can be reinterpreted as making reference to the actual judgment of a properly sampled group of stakeholders rather than to the counterfactual reconstruction of every individual set of preferences.

Thus, we suggest, deliberative tools can help constructing choice architectures that are better informed to people's preferences and more politically legitimate at once. They are better informed because the process of tutoring of preferences described above is a way to ensure that preferences voiced directly by citizens do not arise out of ignorance, misconception or biases. They are more legitimate because they are procedurally so, i.e. they are achieved via a procedure that promotes the ideals of freedom and equality in freedom.[19] Moreover, the revisability[20] of deliberative decisions keeps those in constant democratic check.

Examples of deliberatively justified LP policies

In this section we explore some examples of LP policies, mostly focused on behavioral change, and show how deliberation could play a significant role in the political process of policy-making.
A host of policies have been proposed in the last years that allegedly fall under the heading of LP. We believe that only few of them would qualify as genuine LP policies, either because they are, essentially, information campaigns under another name, or because they entail some sort of change in the architecture of economic incentives. Some of these proposals we discuss here.

---

[17] This is also the reason why private firms should not be subject to the constraints of deliberation. Private firms do not seem to owe any particular toll to the foundations of democracy.
[18] Which, as we argued above, does not entail any washing-away of previously existing dispositions.
[19] For the correlation between deliberation and ligitimisation, see Boniolo (2012).
[20] Concerning the revisability of any deliberative conclusion, see Gutman and Thompson (2004, 133).

In a discussion paper issued by the Cabinet Office Behavioural Insights Team in 2010 (Cabinet Office Behavioural Insights Team 2010) that addresses the concern for the lack of organ donations in a healthcare system—the NHS—desperately in need for such contributions (Department of Health 2008). The paper addresses such topic from a somewhat LP perspective. They set up their suggestion saying that there is a discrepancy between the actual number of people who are enlisted in the NHS Organ Donor (roughly the 27 % of the overall eligible population) and what emerges from surveys (65 % of people would be willing to donate and 90 % are in favor of donation). This tension, they claim, means that people, for some reason, do not manage to live up to the standards that they set for themselves. Hence, they suggest that policy-makers consider the opportunity of prompting the choice of whether or not they would be willing to donate. This choice should be prompted upon applying for a driving license. Evidence that such a strategy has an impressive effect can be drawn from the several US states, e.g. Illinois, Texas and California (Abadie and Gay 2006).

This policy, trialed in England in 2011, seems to pass the test of LP. First, the measure is not coercive in any relevant sense. Secondly, it does not involve the introduction of any non-negligible economic incentive. Third, it exploits one of the biases we mentioned above, namely the so-called power of defaults. And finally it nudges people towards decisions that would enhance their welfare, as judged by themselves (or at least as inferred from their judgments in surveys).

As we have argued, however, the reasons that ground the decision to nudge people towards the satisfaction of their counterfactually reconstructed preferences are not compelling. In this specific case these reasons are even less convincing: preferences expressed in surveys are speech-acts, whose "statement […] does not reveal a serious willingness to incur the opportunity cost[…]. It is evidence simply of [one's] willingness to incur the costs of the *statement* to attain *its* benefits" (Rizzo and Whitman, 2008, 19–20). One could still defend the prompted choice policy, arguing that no decision, after all, is being suggested other than that of taking a decision. This is a fair reply, but we maintain that people have some entitlement in the decision of when to decide concerning subjects that affect them personally. Along this line, our suggested deliberative turn entails that people be involved in the process of taking what we called the meta-choice. Practically, this means that a properly sampled group of stakeholders (in this case British citizens representative of the overall population) ought to convene in a deliberative platform. They should hence be provided with the figures concerning organ donations and the available evidence about strategies for behavioral change. After this phase of tutoring of preferences that relies on state-of-the-art scientific findings, people are asked to argue what they believe would be the best way to go. This deliberative process of reasons giving and heeding is likely to yield outcomes, i.e. policies, similar to the one endorsed by the Cabinet Office. Differently from the suggestion issued by the Cabinet Office, though, this proposal would have an edge as to its political legitimacy, in that it would be the result of a procedurally fair democratic deliberation.

Shopping trolley re-design. Costs associated with overweight and obesity in the UK have been estimated at £7 billion a year (Government Office for Science 2007). This is due to bad eating habits. In England 65 % of adults eat less fruit and vegetables than they ought to (Bates et al. 2010). In order to tackle this issue, the Cabinet Office (2010) suggested following the lead of the New Mexico State University College.[21] C. Payne found that putting a line of yellow tape with a sign designating the part of the trolley to be devoted to fruits and vegetables increased the amount of fruits and vegetables purchased without affecting the overall profit for the retailer. It is unclear how far a policy inspired by such evidence would qualify as LP. In fact, the only ground for the claim that people are being made better off, as judged by themselves, by such a policy, lies in the assumption that people know they should be eating five portions of fruits and vegetables a day.[22] Nonetheless, the policy would indeed be non-coercive and it would entail no change in the architecture of economic incentives.

In this instance, the ideal of deliberation would demand that, again, relevant stakeholders convene and be given certified and understandable information concerning obesity, overweight, their economic and social costs, and the effectiveness of likely strategies for the intervention. This phase of tutoring would be followed by arguments in favor and against the endorsement of specific policies, finally reaching some form of consensus around the means to be deployed to tackle the issue. Two features of the

---

[21] The evidence cited concerning the experiment conducted by Collin Payne at the New Mexico State University College of Business is said to be contained in personal communication between Payne and the Cabinet Office (2010).

[22] This, of course, unless one assumes that everyone prefers being healthy over eating junk food.

deliberation could be controversial in this case: (1) the extent to which the deliberation manages to be inclusive of different social and cultural backgrounds and (1) the way in which the stakeholders participating will be meant to be representative of the overall population. As to (1), the deliberation will demand the provision of collaboratively drafted informational material in plain language and the facilitation of effective participation of subsets of the population unlikely to be familiar with the argumentative style that is proper of ideal conversations. As to (2) the case of children affected by policies poses a challenge to the deliberative approach just sketched. We suggest that the hosting institution makes sure the advocacy of interest groups inherently unable to advocate themselves (and hence not immediately sample-able like children or future generations) does not remain unspoken, possibly presenting their fair claims alongside other standard points of view within the informational material and granting them due space during the deliberation itself.

CONCLUSION

We have shown that LP is affected by one major issue. This issue is epistemic in nature. New paternalists do not provide grounds for the claim that long-term selves' preferences ought to be regarded as better representing—i.e. better than short-term selves' preferences—alleged 'real' individual preferences. This epistemic problem cannot be overcome by means of deliberation concerning meta-choices. However, deliberation can serve the two-fold purpose of unveiling the internal consistency of people's explicit preferences over meta-choices and constructing the political legitimacy that LP lacks, due to its somewhat elitist assumption.[23] Hence we call for the construction of a community of properly chosen deliberators. This community would behave according to the principles of deliberative democracy, outlined by Gutman and Thompson (2004), and according to the method indicated by Boniolo (2012). We do not make assumptions about the outcomes of actual resolutions achieved via this inclusive deliberative procedure, but it will likely be the case that, if, as LP advocates claim, people hold dear their future selves' preferences and are capable of understanding the arguments supporting some architectures of choices over others, then consensus will be reached around arrangements similar to the ones new paternalists have already proposed and experimented.

REFERENCES

Abadie, A., and S. Gay. 2006. The impact of presumed consent legislation on cadaveric organ donation: A cross-country study. *Journal of Health Economics* 25(4): 599–620.

Anand, P., and A. Gray. 2009. Obesity as market failure: Could a 'deliberative economy' overcome the problems of paternalism? *Kyklos* 62(2): 182–190.

Ariely, D. 2009. *Predictably irrational. The hidden forces that shape our decisions. revised and expanded*. New York: HarperCollins Publishers.

Bates, B., A. Lennox, and G. Swan. 2010. *National diet and nutrition survey. Headline results from year I of the rolling programme (2008–2009)*. London: Food Standards Agency and Department of Health.

Beauchamp, T.L., and J.F. Childress. 2006. *Principles of biomedical ethics*, 6th ed. Oxford: Oxford University Press.

Berlin, I. 1969. Two Concepts of Liberty. In *An antology*, eds. Goodin RE, Pettit P (2006), *Contemporary Political Philosophy*, 2nd edition. Oxford: Blackwell Publishing Ltd.

Besharov, G. 2002. *Second-best considerations in correcting cognitive biases*(November 2002). Duke University, Economics Working paper no. 01–08.

Boniolo, G., and G. De Anna. 2006. The four faces of omission. Ontology, terminology, epistemology, and ethics. *Philosophical Explorations* 9: 276–293.

---

[23] That policies can be forged by some group of properly chosen experts. At the very least, the fact that LP advocates do not mention the procedure to be deployed for the choice of the experts is a clue that they do not mean to include 'the public' in the choice of the relevant policies.

Boniolo, G. 2012. *The art of deliberating. Democracy, deliberation and the life sciences between history and theory*. Berlin: Springer.

Cabinet Office Behavioural Insights Team. 2010. Applying behavioural insight to health. Department of Health. 2008. *Organs for transplants: A report from the organ donation taskforce*. London: Department of Health.

Dewey, J. 1934. *A Common Faith*. New Haven: Yale University Press 1997.

Frederick, S., G. Loewenstein, and T. O'Donoghue. 2002. Time discounting and time preference: A critical review. *Journal of Economic Literature* 40(2): 351–401.

Gilovich, T., D. Griffin, and D. Kahneman. 2002. *Heuristics and biases*. Cambridge: Cambridge University Press.

Government Office for Science. 2007. *Foresight—tackling obesities: Future choices—modelling future trends in obesity and the impact on health*, 2nd ed. London: Department for Innovation, Universities and Skills.

Grüne-Yanoff, T. 2012. Old wine in new casks: Libertarian paternalism still violates liberal principles. *Social Choice and Welfare* 38: 635–645.

Gutman, A., and D. Thompson. 2004. *Why deliberative democracy?*. Princeton: Princeton University Press.

Herrnstein, R.J., G.F. Loewenstein, D. Prelec, and W. Vaughan Jr. 1993. Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making* 6: 149–185.

Johnson, E.J., and D. Goldstein. 2003. Do defaults save lives? *Science* 302: 1338–1339.

Kahneman, D., J.L. Knetsch, and R.H. Thaler. 1991. Anomalies. The endowment effect, loss aversion, and status Quo Bias. *Journal of Economic Perspectives* 5: 193–206.

Kahneman, D., and A. Tversky (eds.). 2000. *Choices, values and frames*. Cambridge: Cambridge University Press.

Kitcher, P. 2001. *Science, truth, and democracy*. New York: Oxford University Press.

Kitcher, P. 2011. *Science in a democratic society*. Amherst, New York: Prometheus Books.

Korobkin, B.R. 1998. The status Quo Bias and contract default rules. *Cornell Law Review* 83: 608–687.

Lipsey, R.G., and K. Lancaster. 1956. The general theory of second best. *The Review of Economic Studies* 24(1): 11–32.

O'Donoghue, T., and M. Rabin. 2003. Studying optimal paternalism, illustrated by a model of sin taxes. *American Economic Review (Papers and Proceedings)* 93: 186–191.

Rizzo, M.J., and D.G. Whitman. 2008. *The knowledge problem of new paternalism*, 08–60. Working Paper n: Law and Economics Research Paper Series.

Rizzo, M.J., and D.G. Whitman. 2009. Little brother is watching you: New paternalism on the slippery slopes. *Arizona Law Review* 51: 685–739.

Samuelson, W., and R. Zeckhauser. 1988. Status Quo Bias in decision making. *Journal of Risk and Uncertainty* 1: 7–59.

Strack, F., L.L. Martin, and N. Schwarz. 1988. Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology* 18: 429–442.

Sugden, R. 2008. Why incoherent preferences do not justify paternalism. *Constitutional Political Economy* 19: 226–248.

Sunstein, C.S., and E. Ullmann-Margalit. 1990. Second-order decisions. *Ethics* 110: 5–31.

Sunstein, C.S., and R.H. Thaler. 2003. Libertarian paternalism is not an oxymoron.*University of Chicago Law Review* 70: 1159–1162.

Thaler, R.H., and C.S. Sunstein. 2003. Libertarian paternalism. *The American Economic Review* 93: 175–179.

Thaler, R.H., and C.S. Sunstein. 2008. *Nudge: Improving decisions about health, wealth and happiness*. New Haven: Yale University Press.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211(4481): 453–458.

Whitman, D.G. 2006. Against the new paternalism. Internalities and the economics of self-control. *Policy Analysis* 563: 1–16.

Williams, B. 1995. Which slopes are slippery? In *Making sense of humanity*, ed. Bernard Williams. Cambridge: Cambridge University Press.