



Francesco Guala and Luigi Mittone

A Political Justification of Nudging

CEEL Working Paper 7-13

Cognitive and Experimental Economics
Laboratory

Via Inama, 5 38100 Trento, Italy

<http://www-ceel.economia.unitn.it>
tel. +39.461.282313

A Political Justification of Nudging[°]

Francesco Guala^{*} and Luigi Mittone[§]

Abstract:

Nudge policies are typically justified from paternalistic premises: nudges are acceptable if they benefit the individuals who are nudged. A tacit assumption behind this strategy is that the biases of decision that choice architects attempt to eliminate generate costs that are paid mainly by the decision-makers. For example, in the case of intertemporal discounting, the costs of preference reversal are paid by the discounters. We argue that this assumption is unwarranted. In the real world the costs of reversal are often transferred onto other individuals. But if this is the case, the biases create externalities, and nudges are best justified from a political rather than paternalistic standpoint.

[°] Research for this paper was supported financially and logistically by the Fondazione Bruno Kessler. We are grateful to participants at the Behavioral Economics Workshop held at FBK in October 2013 for their comments and suggestions.

^{*} Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Italy.
Email: francesco.guala@unimi.it

[§] Department of Economics and Management, and CEEL, Università degli Studi di Trento, Italy.

1. Introduction

According to behavioural economists, a better understanding of the psychology of decision-making can significantly improve both economic theory and economic choices. On the first front behavioural economics has scored many points, and has already become part of the scientific mainstream. The second project in contrast has come to the fore only recently, after policy-makers have realized that policies inspired by behavioural economics may deliver results at a lower cost than traditional interventions.¹ Most of the debate has centred around so-called “Libertarian Paternalism”, an approach to social policy advocated by influential economists like Colin Camerer, George Loewenstein, Matthew Rabin, Cass Sunstein and Richard Thaler.² Sunstein and Thaler, in particular, have advertised the use of behavioural policies in *Nudge*, a widely read book that not only illustrates a number of actual and potential applications, but also tries to lay philosophical foundations for “soft” government intervention. This rise of interest in turn has stimulated the discussion about the pros and cons of behavioural policies, what they can deliver, and whether they are genuinely different (and better) than the alternatives.³

A “nudge” is a policy intervention that targets the environment in which citizens (the consumers of private and public goods) make decisions that have important implications for their wealth, health, and happiness (to evoke the sub-title of Thaler and Sunstein’s bestseller). The environment in most cases is manipulated in such a way as to preserve the full range of options that citizens already have, that is, without reducing their freedom of choice. Subtle features of the “choice architecture”, however, are arranged so as to facilitate the choice of options that are beneficial to the decision-makers. Nudge policies typically remove psychological biases that prevent people from making the right decision, or use the biases to direct behaviour towards the best option. Because choice architects aim at improving people’s lives without reducing their freedom, the policy of nudging has been marketed using the seemingly paradoxical label of “Libertarian Paternalism”.

Philosophical discussions of Libertarian Paternalism so far have mostly focused on this oxymoron. Critics have argued for example that the policies advocated by nudgers are not really paternalistic because even though they are aimed at improving the well-being of citizens, well-being is defined with respect to the preferences of the citizens themselves. Genuine paternalism imposes a (superior) view of well-being on recalcitrant subjects, while nudges simply help people achieve what they want (e.g. Hausman and Welch 2010). Another line of criticism has focused on its libertarian credentials: nudges may preserve so-called option-freedom, but surely infringe the autonomy-freedom of individuals. Since option freedom without autonomy is hardly valuable, a

¹ See for example Dolan et al (2010).

² Cf. e.g. Camerer et al. (2003), Loewenstein and Hasley (2008), Thaler and Sunstein (2003, 2008), Sunstein and Thaler (2006).

³ The literature is already too large to be reviewed here, but see for example Glaeser (2006), Berg and Gigerenzer (2007), Amir and Lobel (2008), Sugden (2008), Bovens (2009), Hausman and Welch (2010), Selinger and Wythe (2010), Brennan and Brooks (2011), Grune-Yanoff (2012), Rebonato (2012), Schnellenbach (2012), Qizilbash (2012), Wilkinson (2012), Haybron and Alexandrova (2013).

genuine libertarian should be unhappy with the policy of nudging (Hausman and Welch 2010, Grune-Yanoff 2012).

In this paper we follow a different line of reasoning. Although we are sympathetic with the policies advocated by behavioural economists, we agree with the critics that nudgers so far have not done a very good job at justifying their policies. Part of the problem, we think, is their attempt to justify nudges in paternalistic terms. A tacit assumption behind this strategy is that the biases of decision that choice architects attempt to eliminate generate costs that are paid mainly by the decision-makers. We argue that this assumption is problematic. In the real world the costs of biases are often transferred onto other individuals. But if this is the case, then the biases create externalities and nudges are justifiable from a *political* rather than paternalistic standpoint.

The paper is organized as follows: section 2 provides a quick overview of standard justifications of government intervention in the economic realm. This section is largely descriptive and paves the way for the behavioural economics account of market failures, in section 3. Section 4 illustrates a specific example, focusing on temporal discounting. The main thesis of this paper is formulated and defended in sections 5 and 6, where we argue that biases like temporal discounting can create externalities and therefore nudge policies are better justified from non-paternalistic premises.

2. Market failures and policy intervention

We begin with some general remarks on the role of government intervention and regulation. Although most readers are probably familiar with these remarks, we repeat them because they will become important later in the paper. (Impatient readers may want to skip this section and jump directly to section 3.)

The justification of government intervention is a central topic in contemporary debates of political economy. When, where, and how much interference with individual choice is legitimate? On the far right of the political spectrum, libertarians support the idea of a legal system exclusively concerned with non-economic matters, and a minimal apparatus of state regulation of economic matters. Libertarian justifications range from the purely political – to defend individual freedoms – to purely economic ones. The latter rely especially on theorems proving that, *in the appropriate circumstances*, free markets deliver allocations of goods that are efficient in Pareto's sense. The proviso is important, for neoclassical economics has devoted significant resources to the study of "market failures".

The theory of market failures is concerned with those circumstances in which a competitive market does not or cannot take place because of some particular feature of the commodities to be exchanged. The classic cases studied by economists include the existence of natural monopolies; the transaction costs caused by asymmetric information; externalities generated by the impossibility to attribute property rights; and the existence of non-excludable, non-rival public

goods.⁴ The neoclassical theory – through the fundamental theorems of welfare economics – suggests that state intervention can be extended in order to correct these imperfections, with the exclusive aim to restore efficiency.

It is important to stress that the neoclassical theory of market failures has little to do with Keynesian critiques of the market as an efficient allocation mechanism. According to Keynes (1936) markets can fail to assure full employment of the available resources even when they operate in ideal conditions, unencumbered by information asymmetries, externalities, or monopolies. Moreover, the theory of market failures is also independent from normative approaches that criticize markets for their distributional consequences. Neoclassical theory in fact tries to steer away from politically contentious considerations regarding justice and inequality – the use of the Pareto principle is aimed precisely at avoiding controversies about distributional issues.

In practice, however, the “political economic” justification for the welfare state systems of contemporary Western countries is a blend of arguments taken from neoclassical economics and from Keynesian theory, with a pinch of normative advice (we ought to redistribute wealth so as to reduce vertical inequality). This is the background that must be kept in mind to frame the debate about Libertarian Paternalism.

Behavioural economics and Libertarian Paternalism introduce a new set of considerations, in particular the idea that market failures may be caused by biases of individual decision-making. The idea that competitive markets are efficient allocation mechanisms in fact presupposes the existence of rational agents. A neoclassical competitive market is a sort of super-organism that derives its overall efficiency properties from the efficiency of each part. Strictly speaking, the market may deliver efficient allocations in spite of sub-optimal agents, if agents’ interactions reduce the overall effect of individual mistakes. But this is true only if the majority of the agents adopt intrinsically well-behaved strategies, or if the behavioural errors are distributed randomly, so that they compensate one another. What if the mistakes are widespread and systematic? If several agents choose the same (wrong) behaviour, the markets generate an inefficient allocation of resources. Behavioural economists claim that this is a realistic prospect, because many types of economic choice – inter-temporal choice, choice under uncertainty and risk, strategic decisions, etc. – are systematically affected by cognitive biases.

Notice an important difference between behavioural economics and standard neoclassical market failures: the source of the problem here is not an architectural flaw of the market (the existence of asymmetric information, externalities, etc.) but a widespread failure of rationality. Individual agents fail to behave according to their “true” preferences. The interventions prescribed by

⁴ Economists usually distinguish between public goods and externalities on the basis of the non-rivalry criterion. While goods with externalities are merely non-excludable (once produced, their consumption cannot be restricted), public goods are non-excludable and non-rival (consumption by one individual does not prevent consumption by another individual).

behavioural economists therefore are aimed at restoring the smooth operation of markets, by replacing irrational agents with well-behaved ones. The nudge approach, in this specific sense, may be considered an ally of the neoclassical approach.

3. Bounded rationality and “genuine” preferences

An irrational choice, according to behavioural economists, is a choice that does not reflect a consistent structure of preferences. This idea is fairly recent. The work of contemporary behavioural economists is in many ways a descendent of the theory of bounded rationality developed by Herbert Simon in the 1950s. But according to Simon (1955), the bounded decision maker is not irrational; on the contrary she is consistent but forced by her cognitive limitations to consider only a subset of the available alternatives. Simon uses the analogy of a physical constraint: consider a bird that can only fly at up to 70 kilometres per hour; the set of alternatives available to escape from a predator does not include flight at 200 km/h. Analogously, a decision maker cannot compare more than a certain number of alternatives. She must define a boundary for the comparative process and then choose within this boundary. Another way to put it is that Simon’s decision maker does not suffer from internal inconsistency of preference, and any apparent behavioural incoherence that we observe is due to a change in the alternatives that are included in her decision (sub)set.

Contemporary behavioural economists, in contrast, make stronger claims:

Our major emphasis is on the fact that in many domains, people lack clear, stable, or well-ordered preferences. What they choose is strongly influenced by details of the context in which they make their choice, for example default rules, framing effects (that is, the wording of possible options), and starting points. These contextual influences render the very meaning of the term ‘preferences’ unclear. If social planners are asked to respect preferences, or if they are told that respect for preferences promotes well-being, they often will be unable to know what they should do. (Sunstein and Thaler 2006: 233)

While Simon’s agent has a clear preference ranking but is bound to consider only part of the option space, Sunstein and Thaler’s agent cannot be said to possess a preference at all. This is a deeper failure of rationality that calls for more drastic solutions.

What is the link between choice inconsistency and the lack of genuine preferences? In his recent book on *Preference, Value, Choice and Welfare*, Dan Hausman (2011) defines preferences as *total comparative evaluations* of the consequences of actions. Intuitively a preference is a judgment that, all things considered, x is better than y . There are three important elements in this definition: first, a preference is an *evaluation*. Second, it is *comparative*, not absolute. It must involve at least two items that are compared (in economics, a preference is expressed via a

ranking). Third, it is a *total* evaluation: nothing that matters must be excluded.⁵ The “total evaluation” requirement is a consequence of the principle of independence, which is crucial for the stability of the preference relation.⁶ But the axiom of independence has been shown to be empirically violated in a number of experiments, which have fuelled behavioural economists’ scepticism concerning preferences.

Failures of independence may be caused by failures to consider attributes that matter to the decision-maker. Intrinsic inconsistencies at the level of preferences therefore may result from a mechanism that is in many ways similar to the one discussed by Simon. A bounded rational agent in fact may not only be unable to search the full space of options; she may also be unable to search the full space of *attributes* of each option. The latter is especially relevant if the options are complex, or if choices have ramifications with multiple future consequences. A limited human being in such circumstances may only be able to consider the most salient characteristics of each option, overlooking others. If salience varies across time or across circumstances, choice inconsistencies may occur.⁷

4. An example: temporal discounting

In this section we cast these general theoretical considerations in the context of a concrete example. In the course of the paper we shall often refer to a bias that plays a prominent role in the libertarian-paternalistic literature, namely *temporal discounting*. Thaler and Sunstein devote many pages to describe programmes – like “Save More Tomorrow” – designed to nudge people toward putting more money in their pension schemes. These programmes rely on techniques such as the manipulation of default options, that influence choice without reducing substantially the freedom of decision-makers. In general, the nudges try to curtail our tendency to sacrifice future well-being in order to satisfy immediate wants. This tendency of course is not necessarily problematic, unless people come to regret their past decision to consume rather than save. But when this happens, temporal discounting leads to temporal inconsistency of preference and, possibly, inconsistent choice.

Temporal inconsistency can be explained along the lines illustrated in the previous section, using Construal Level Theory (CLT), a framework developed by Nira Libermann and Yaacov Trope.⁸ When

⁵ This is quite different from the way we talk about preferences in everyday discourse. In economic theory, it is improper to say “I prefer meat to fish, but since I am worried about my cholesterol level I’ll choose fish”. The sentence should rather be rephrased as follows: “I would prefer meat to fish, if my cholesterol level was not so high; but given that I am concerned with my cholesterol, I prefer fish to meat instead”.

⁶ Independence requires that the ranking of two options (for example, $x > y$) does not change when new options become available to the decision maker. If the addition of an option (say, z) changes the relative ordering of x and y , the agent must have overlooked some feature of x and y (for example their complementarity with z) that matters for the decision at stake.

⁷ For a general analysis of this type of choice inconsistency, see Mongin (2000).

⁸ See e.g. Trope and Libermann (2003, 2010).

a choice produces consequences in the future, the decision-maker must construct a mental representation of future events that may be influenced by her decision. The level of abstraction of this mental representation, according to CLT, depends on temporal distance: the further in time the event is, the more abstract its representation. This cognitive process can be described by positing that future events are characterized by several attributes – which, in Simon’s theory, are equivalent to the alternatives included in the decision subset. Because of her cognitive limitations, the decision maker cannot evaluate them all, but will rather select a subset that she will consider for her decision. If we accept the empirical evidence reported by CLT, the selection of the subset of attributes depends at least in part on temporal distance. As time passes, the composition of the subset changes, and choice may change too. This process may lead to a classical situation of inter-temporal incoherence.

Suppose for example that Ann is facing a saving decision early in her life. One option is to spend immediately all the extra money she is earning, buying for example holiday trips around the world. Another option is to invest the money in a pension fund. Let us imagine that, if she decides to save, her youth will be fairly boring compared to the fun she could have if she decided to spend it on holidays and other treats. But if she saves, she will be able to rest and enjoy a wealthy retirement later in life. If she spends it all, in contrast, she will face the alternative between working until a much later age or retiring with a low pension.

Ann’s decision has several different implications (attributes). To simplify, we limit the analysis to six attributes on three different levels or dimensions (fun vs. boring youth, wealthy vs. poor retirement, rest vs. toil in old age). Let us suppose that Ann’s preferences along each dimension are organised as follows:

- She prefers to have Fun rather than to have a Boring youth ($F > B$);
- She prefers to be Wealthy rather than Poor later in life ($W > P$);
- She prefers to Rest rather than Toil when she is old ($R > T$).

A rational agent should be able to form a coherent preference ranking across all the combinations of these attributes, or possible life-styles. But like most people Ann has cognitive limitations. Let us assume that she can only manage to consider two attributes at a time. We shall also assume that she suffers from the myopia of CLT: when the consequences of a decision (attributes) are far away in time, they become less salient and tend to fade. As a consequence, her choices are going to be affected by the subset of attributes that are cognitively salient.

At the time of making her saving decision (t_1) only the first dimension (F or B) is salient. But the choice made at t_1 will determine which options are available later in life (at t_2):

- If Ann chooses B, then the set of options at t_2 is {BWR}, {BWT}, {BPR}, {BPT};
- If Ann chooses F, then her set of options at t_2 is {FWT}, {FPR}, {FPT}.

CLT predicts that at t_2 the attributes F and B will cease to be salient, and Ann will focus on the second and third attributes only. Let us suppose her preference ranking over these pairs of attributes is

$\{WR\} > \{WT\} > \{PR\} > \{PT\}$.

Her preferred option is to be wealthy without having to work; second-best is to toil in order to be wealthy; then to rest and be poor; worst of all is to be poor and toil. But notice that the option WR is not available, if she has chosen F when she was young. At t_2 nevertheless Ann can still compare the option FWR she had at t_1 with the best option she has now (BWT): since B and F are not salient anymore, and $WR > WT$, at t_2 she regrets the choice she has made when she was young.

Notice that Ann's choice at t_1 has narrowed her options at t_2 . Giving up R is the cost that she has to pay at t_2 in order to reverse her previous decision. Initially she could have had both W and R (at the cost of F); at t_2 she cannot. If she wants W she must pay a cost in terms of R. Psychologists' examples in the CLT literature tend to obscure this point, by focusing on seemingly reversible choices. But some choices are strictly speaking irreversible: you cannot go back in time. The option WR is unavailable to Ann after she has chosen F. What Ann can do is pay a *cost* to exchange PR for something she likes better (at t_2). But she cannot decide to have F and R simultaneously.

The same story can be told in such a way as to make it compatible with so-called Multiple Selves Theory (e.g. Ainslie 1992): in a multiple selves framework, each "self" only cares about selected attributes because these attributes (what happens *now*) are the only relevant consequences *for her*. The other dimensions concern events that do not happen to that decision-maker, but to someone else (to another self, at another time). Intertemporal choice then may give rise to problems that are formally analogous to the Condorcet Paradox: at t_1 self₁ is faced with B and F because they are the only things that matter to her, and chooses F. Then self₂ is faced with PR and WT, and chooses WT. But if self₂ had voted first on the "wealth" and "rest" dimensions, she would have chosen WR. So the order of the polls determines the outcome.⁹

Ideally, a rational agent should be able to aggregate or arbitrate among the various dimensions of attributes, or her different selves. She should take the perspective of a superior, unified self that is able to attribute more or less weight to the desires of some sub-selves or others. Equivalently in terms that are familiar to CLT, the agent should be able to build a single utility function assigning specific weights to the various dimensions of the consequences. The problem is to find a justification for this procedure. In a multiple selves framework we may see the choice architect as a third individual who forms a coalition with one of the selves. For example: she joins the older self and votes against the younger self. This is equivalent to "put the child in the adult's shoes", or to try to implement the point of view of the adult self in CLT theory. But it is not clear why one should form *this* coalition instead of another. Why not the opposite? We could join the younger

⁹ The analogy between inter-personal and intra-personal decision problems was first noticed and discussed by Kavka (1991).

self and vote the elderly out. What is the justification for forming one coalition rather than another?

The choice architect typically decides to endorse the preferences of old Ann, nudging young Ann to make decisions that will be more likeable to her future self. In a CLT perspective, one may reduce the effect of temporal distance enhancing the salience of some attributes that will become important in the future. This is equivalent to help people form a comprehensive utility function. But there are many ways to do it: in principle the choice architect might as well put elderly people in the shoes of younger selves; after all one has a lot more fun traveling, partying, spending money when she is young than when she is old. Provided she keeps fond memories, old Ann might be helped or nudged to see her life choices in such a way that she feels no regret for her earlier decisions.

Of course we are not claiming that people should be encouraged to be irresponsible when they are young. The examples are only meant to show that, seen from a multiple selves or a CLT perspective, the choice between different nudges seems arbitrary. A younger self is just as recalcitrant to be “voted out” than an older self. An old self may be nudged to empathise with a young self, or the other way around. The literature on choice architecture hides the problem under the carpet, by simply *assuming* that we ought to help youngsters to save more. But nothing in the behavioural economics or libertarian-paternalistic framework seems to justify this assumption.

5. The politics of nudging

We suspect that this conundrum is the consequence of taking the wrong perspective, right from the start. Let us step back and reflect on what nudgers have been trying to do: the supporters of nudge policies seek a solution that is in some sense a-political. Libertarian Paternalism is “paternalistic” to the extent that its policies promote the well-being of the nudged individuals. “I – says the choice architect – intervene because I care about your well-being”. But unfortunately, as we have just seen, it is not entirely clear who is “you” (multiple selves problem), or what is the “well-being” to be maximized (multiple attributes problem). If nudgers, like most economists, endorse a subjective view well-being (as equivalent to the satisfaction of people’s preferences), they end up in a dead-end alley.

Since we sympathize with the general spirit of nudge policies, we propose a simple change of perspective: we shall argue that nudges are best served by seeking a *political* justification. The idea, roughly, is that choice architects are justified to intervene to protect *other* people from the damage that may be caused by irresponsible individuals. Nudge policies are not (or not only) for the good of the nudged, but for the good of third-parties that otherwise are going to be harmed.

To reach this conclusion we have to reconsider the costs of preference reversals. All the examples examined so far presuppose that all the costs are internalised. Once Ann has chosen Fun, for example, she can only have Wealth at the expense of Rest. She pays the full cost of her myopia. But there are reasons to believe that things may not always go this way. Real people are skilful at sharing the costs of their mistakes. If somebody else can be made to pay, then myopic Ann may be able to have fun *and* wealth *and* rest. The best of all worlds.

How do you make others pay? The standard way in modern democracies is to form a coalition and legislate. This is particularly easy when the “others” cannot vote because they are not even born: irresponsible spenders can make their children and grandchildren pay for the debt accumulated by previous generations. The one who pays for Ann’s past mistakes need not always be Ann’s future self; it may also be *somebody else’s* future self.

This brings an entirely new philosophical perspective on the issue of nudging, because it opens the door to the classic arguments for policy intervention that we have reviewed in section 2. The claim that people ought to be nudged is easily defensible once we realize that saving for the future is not entirely a private matter. The problem with not saving enough is not merely that people will regret it when they will be old. It is also that we will have to deal with plenty of old people who cannot support themselves. Seen from this perspective, the focus on preferences and paternalism shifts the spotlight away from some key issues. The key political problem is that a society with a large number of elderly people who cannot support themselves is not viable. The elderlies would burden younger people with large costs, and the latter would try to resist, generating conflict and unrest.

The argument is entirely general, and applies in principle to any policy intervention that tries to correct biases of intertemporal choice. It is particularly relevant for behavioural economists, however, because nudges have been mainly defended from a paternalistic standpoint. Most cases discussed by libertarian paternalists, including smoking and dieting, are potential sources of externalities. If these behaviours were entirely private, no one would bother very much about nudging. Imagine, for the sake of the argument, a new research demonstrating that there is a strong positive correlation between frequent sexual intercourse and well-being. Would we approve of a governmental policy aimed at nudging people to have more sex? Regardless of its impact on well-being, sex really is a private matter and most people would think that governmental intrusion is unjustified in this area.

So how is saving different from sex? The main difference is that saving behaviour has important externalities. Although undoubtedly it does affect the welfare of each individual worker, saving also has non-negligible effects on others. In our own country, to take a concrete example, a couple of generations raised with unrealistic expectations of public welfare support have burdened another couple of generations with a mountain of public debt, plus the prospect of low employment and low pensions. This situation is not easily redressed, because as younger people in

Italy have learnt all too well, the votes of an ageing population can steer politicians toward policies that are detrimental for the new generations.

When we say that people do not pay enough attention to their future needs, then, we are not making a normative *moral* statement (“we must help myopic people, promote their own well-being”). We are making a normative *political* statement, assuming implicitly the point of view of society.¹⁰ And we can do it because saving decisions have important externalities. But if externalities do play a role, then the nudges that behavioural economists are proposing should not be defended solely on paternalistic grounds.

6. Externalities and norms

An appeal for a political justification of nudging faces some potential counter-arguments. A libertarian sceptic, for example, may argue that in an ideal free society the youngsters should be entitled to refuse welfare support for their parents and grandparents, if so they wish. Constitutional architectures should be devised to protect born and unborn citizens from the threat of organized coalitions like the ones we have envisaged above. In this ideal society nudges would be superfluous because the externalities would not arise in the first place.

Another related argument challenges the political justification as follows: suppose that we could create an institutional arrangement where the costs of myopia *cannot* be externalized. Then, one may argue that we would *still* have a moral duty to help those who have made mistakes in the past. Suppose for example that, as a matter of fact, all smokers were to die at the age of 60 without burdening the public healthcare system. Shouldn't we nudge young potential smokers towards a healthier lifestyle for moral reasons only?¹¹

Our reply is twofold. First, let us consider the prospect of setting up constitutional barriers against coalitions that want to externalise costs. Although the idea seems fair in principle, it may be very difficult to implement it in practice. A large number of low savers can easily lobby politicians and bend the rules in their favour. Constitutional defensive barriers thus must be constantly defended, which requires an investment of time and resources. Even the preservation of liberties entails costs. A policy that corrects for myopia at low cost then may look very attractive even to the libertarian constitutionalist.

Our second point is that we must be careful not to take a narrow view of the relevant costs. As the second argument correctly points out, there are social ties and moral norms that regulate insurance and cost-sharing even where legislation does not reach. People may be under strong

¹⁰ By “political” here we mean roughly “contractarian”, or that requires arbitration among the interests of several parties. We do not intend to deny that political solutions, in this specific sense, are independent from moral considerations.

¹¹ We are indebted to Luc Bovens for this example.

social pressure to support the members of older generations, quite independently of any legal prescription. This is an area where cultural differences may play a major role. Consider dieting – another example that features prominently in the libertarian paternalistic literature. Obesity is in many ways similar to irresponsible spending: although the consequences seem to be a private matter, in some situations they can constitute a social problem. However in North America the idea that medical insurance is an individual responsibility is largely (albeit not universally) accepted. In Europe, in contrast, public medical care is such an established institution that the idea of letting people die if they cannot pay their medical bills would strike most people as strongly counter-normative. Since it would be politically costly to transfer all the costs to the patients, an obesity epidemics might have significant negative externalities in Europe. In this context nudges are attractive because they lower the transaction costs of political bargaining: they prevent institutions that we value highly, such as democracy, from being subjected to excessive strain. In the worst-case scenario such institutions may even fail, so we should all appreciate the advantage of preventing political stress, if it can be done at low cost.

Our final point in this paper, then, is that the concept of externality is context-specific, and may depend heavily on local norms. The reason why we feel that people should save more is that a society that does not save enough may face catastrophic consequences in the future, and the costs will be shared at least in part with those who are saving now. The degree of this sharing will depend in part on the existence of social or moral norms that prescribe assistance towards the weakest members of society. Where such norms are strong – as in the case of medical assistance in Europe – it is reasonable to say that unhealthy eating, smoking, and similar behaviours should be treated as externalities.

But if this is the case, then paternalism is a red herring. Coercion may be licenced even if the preferences of individuals regarding their own lifestyles are stable and consistent. In the case of inconsistent preferences, the policy of nudging has the further attraction of keeping the options open to the decision-maker. But we have argued that we should not worry about preferences too much: we should stop looking for paternalistic justifications of nudge policies. The main reason we think that people ought to take care of their future is that their failure to do so will probably affect their children and the children of others. Nudge policies are attractive because they tackle these externalities in a cheaper and less intrusive ways than traditional alternatives. For all these reasons they should be welcomed and put to use whenever they help create a more viable society. The libertarian paternalistic tag may be a good selling point in the current political climate, but it fails to capture the real reasons why we feel entitled to nudge saving, dieting, and the like.

References

- Ainslie, G. (1992). *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person*. New York: Cambridge University Press.
- Amir, O. and Lobel, O. (2008). Stumble, predict, nudge: How behavioral economics informs law and policy. *Columbia Law Review*, 108, 2098-2137.
- Berg, N. and Gigerenzer, G. (2007). Psychology implies paternalism? Bounded rationality may reduce the rationale to regulate risk-taking. *Social Choice and Welfare*, 28, 337-359.
- Bovens, L. (2009). The ethics of Nudge. In T. Grune-Yanoff and S.O. Hansson (eds.) *Preference Change: Approaches from Philosophy, Economics and Psychology*. Dordrecht and New York: Springer, pp. 207-219.
- Brennan, G. and Brooks, M. (2011). On the 'cashing out' hypothesis and 'soft' and 'hard' policies. *European Journal of Political Economy*, 27, 601-610.
- Camerer, C., Issacharoff, S. Loewenstein, G., O'Donoghue, T. and Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review*, 1151, 1211-1254.
- Dolan, P., Hallsworth, M., Halpern, M King, D. and Vlaev, I. (2010). *MINDSPACE: Influencing behaviour through public policy*. London: The Cabinet Office/Institute for Government.
- Glaeser, E. (2006). Paternalism and psychology. *The University of Chicago Law Review*, 73, 133-156.
- Grune-Yanoff, T. (2012). Old wine in new casks: Libertarian paternalism still violates liberal principles. *Social Choice and Welfare*, 38, 635-645.
- Hausman, D.M. (2011) *Preference, Value, Choice and Welfare*. New York: Cambridge University Press.
- Hausman, D.M. and Welch, B. (2010). To nudge or not to nudge. *Journal of Political Philosophy*, 18, 123-136.
- Haybron, D.M. and Alexandrova, A. (2013). Paternalism in economics. In C. Coons and M. Weber (eds.) *Paternalism: Theory and Practice*. New York: Cambridge University Press.
- Kavka, G. (1991). Is individual choice less problematic than collective choice? *Economics and Philosophy*, 7, 143-165.
- Keynes, J.M. (1936). *The General Theory of Employment, Interest and Money*. London: MacMillan.

- Loewenstein, G. and Haisley, E.C. (2008). The economist as therapist: methodological ramifications of “light” paternalism. In A. Caplin and A. Schotter (eds.), *The Foundations of Positive and Normative Economics – A Handbook*. Oxford University Press, pp. 210-245.
- Mongin, P. (2000). Does optimization imply rationality? *Synthese*, 124, 73-111.
- Qizilbash, M. (2012). Informed desire and the ambitions of libertarian paternalism. *Social Choice and Welfare*, 38, 647-658.
- Rebonato, R. (2012). *Taking Liberties: A Critical Examination of Libertarian Paternalism*. Palgrave Macmillan.
- Schnellenbach, J. (2012). Nudges and norms: On the political economy of soft paternalism. *European Journal of Political Economy*, 28, 266-277.
- Selinger, E. and Whyte, K.P. (2010). Competence and trust in choice architecture. *Knowledge, Technology, Policy*, 23, 461–482.
- Sugden, R. (2008). Why incoherent preferences do not justify paternalism. *Constitutional Political Economy*, 19, 226-248.
- Sunstein, C.R. and Thaler, R.H. (2006). Preferences, paternalism, and liberty. *Royal Institute of Philosophy Supplement*, 59, 233-264.
- Thaler, R.H. and Sunstein, C.R. (2003). Libertarian paternalism. *American Economic Association Papers and Proceedings*, 93, 175-179.
- Thaler, R.H. and Sunstein, C.R. (2008). *Nudge. Improving Decisions about Health, Wealth and Happiness*. Penguin.
- Trope, Y. and Liberman, N. (2003). Temporal construal. *Psychological Review*, 110, 403-421.
- Trope, Y. and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117, 440-463.
- Wilkinson, T.M. (2012). Nudging and manipulation. *Political Studies*, 61, 341-355.