

# Is Libertarian Paternalism an Oxymoron? A Comment on Sunstein and Thaler (2003)

Christophe Salvat

CNRS-Greqam, Aix-en-Provence  
Robinson College, Cambridge

## Summary

In this article I have attempted to discuss two points brought up by Sunstein and Thaler's new perspective on paternalism. The first is that whatever individuals do, they have to take second rank decisions, i.e. decisions that are equally attractive for them but impact others' people welfare. They then face a three alternative choice: either they do not care of the collateral consequences of their choice and act randomly, either they select the option that will make others better off, either they pick the option that will make others worst off. The authors assume they will necessarily act benevolently. I argued this statement is mistaken, and that the authors did not sufficiently distinguish the functional and the personal motivations of the planners. This leads me to define paternalism as the personal benevolence of the planners. Because they do not have any professional or moral duty towards the non-planners, they have absolutely no obligation to be benevolent. In most cases, the easiest choice is to act randomly. This incertitude about the planners' choice has dramatic consequences on non-planners: unless they know and trust the planners they cannot expect them to be paternalistic. They are subsequently less prone to blindly adopt default rules. The second point of the authors is that one can conceive a paternalistic system based on planners' benevolence that would not infringe libertarian principles. Default rules can, for instance, be combined with absolute freedom of choice. I argued that freedom of choice would not contribute to individual freedom if choices are not voluntary. The easiest they set up opting out options the less discriminating their system is, and eventually the most inefficient it will be.



## 1. Introduction

In a recent article entitled “Libertarian paternalism is not an oxymoron”<sup>1</sup> Cass Sunstein and Richard Thaler challenge the common idea that, under some conditions, paternalism does not infringe upon the individual’s freedom of choice. Sunstein and Thaler propose to dispel the myth of a coercive and morally demeaning paternalism which would be incompatible with the liberal spirit of our societies.

For Sunstein and Thaler, the general discrediting of paternalism in modern societies can be explained by a false assumption and two misconceptions they propose to contest. Firstly, it is commonly assumed that people actually behave as economic models expect them to do. The standard of individual rationality formulated by the economists has been incorrectly understood as a positive and experimental account of human nature. Behaviourist economists like Richard Thaler have demonstrated, through numerous experimental studies, how faulty this ‘*idée reçue*’ can be. The first misconception is that there are viable alternatives to paternalism. The question of paternalism is always discussed from a moral point of view (‘Is paternalism right or wrong?’) when it should be considered from a pragmatic standpoint (‘Since paternalism is unavoidable, what can we do to make it acceptable?’). The second misconception is that paternalism always involves coercion, and should therefore be rejected by those who revere freedom. Paternalism is not, however, the exclusivity of the State and most forms of paternalistic actions do not involve any coercion. In any case, the authors claim, it should always be possible to avoid coercion of public actions by securing freedom of choice. In a book they published in 2008, in which they popularize but also further develop their ideas, Sunstein and Thaler term this kind of benevolent but non-coercive action a ‘nudge’ (Sunstein & Thaler 2008).

Although I disagree with their conclusions, I believe that Sunstein and Thaler’s view of paternalism is very suggestive. Introducing behaviourism into the topic opens up a new area of reflection: paternalism is no longer a political or a moral issue, it is an element of individual deliberation. Every time a person has to make a decision that might impact on others’ lives, she has the choice of paternalism. For an economist, and *a fortiori* a utilitarian, it is a unique chance to go beyond the individuals’ bounded rationality, and to improve their welfare. For a libertarian, however, it might represent a new and uncontrollable source of infringement to freedom. The concept of libertarian paternalism that the authors develop aims at building a bridge between them. The present article offers a critical assessment of this concept and discusses its main components: inescapability and non-coercion. It also proposes to develop further

---

<sup>1</sup> The original article has been published in the *University of Chicago Law Review*, vol.70, 2003, pp.1159-1202. Three other versions have been published: Thaler R.H. and Sunstein C.R., (2003), ‘Libertarian Paternalism’, *The American Economic Review*, Vol.93, No. 2, May, pp. 175-179; Sunstein C. and Thaler R.H., (2006), “Libertarian Paternalism is not an Oxymoron”, in S.Lichtenstein, P.Slovic (eds.), *The Construction of Preference*, Cambridge University Press, pp.689-707; Sunstein C. and Thaler R.H., (2006), “Preferences, Paternalism, and Liberty” in S.Olsaretti (ed.), *Preferences and Well-Being*, Royal Institute of Philosophy Supplement, vol.59, Cambridge University Press, pp.233-264.

their analysis through two fundamental notions, benevolence and voluntariness, which are necessary to characterize paternalism. Sections 1 and 2 respectively present the two misconceptions that the authors offer to clear up. Section 3 criticises the idea that planners act paternalistically by necessity and suggests an alternative definition of paternalism. Section 4 introduces the concept of voluntariness to show that the libertarian aspect of libertarian paternalism cannot be ensured by freedom of choice.

## 2. Necessity of Paternalism

The term paternalism covers a wide variety of meanings<sup>2</sup>. Joel Feinberg proposes to clarify any potential debate by distinguishing two families of paternalism, a presumptively blameable and a presumptively non-blameable paternalism (Feinberg 1986:4). The first kind of paternalism consists of treating adults as if they were children (I), and is further subcategorised into benevolent (Ia) and nonbenevolent (Ib) paternalism. This last form of paternalism is partly responsible for the negative image of paternalism in public opinion, and can be associated, for instance, with the condescending attitude of a superior and a patronizing colleague. The second type of paternalism is more consensual. It refers to protective actions towards intellectually or physically vulnerable persons which match the caring function of parenthood (II). Although they are undoubtedly the closest form of parent-like behaviour, neither those types of benevolent actions nor the sheer demeaning conducts of Ib will be considered as paternalistic in my argument. From now on, I shall refer to paternalism as actions or decisions intended to protect someone or promote someone's welfare (Ia).

Paternalism thus characterises *prima facie* a positive intention or action of a person (or a group of persons) B towards a person (or a group of persons) A that would have been totally appropriate if A were objectively unable to make responsible decisions. The debate around the legitimacy of paternalism is, therefore, necessarily linked to the assessment of the degree of responsibility of sane adults and of the degree of voluntariness of their actions. The debate has long been focussed on political issues opposing liberal thinkers to moral Conservative and pragmatic Utilitarians. Cass Sunstein and Richard Thaler have recently made a significant step in reorientating the discussion towards individual choice. This very thought-provoking work can lead to a more specific definition of a paternalist action and, consequently, to a more detailed account of its legitimacy.

Sunstein and Thaler start by discarding the very idea that ordinary human beings voluntarily and rationally act in their own best interest. Economists have falsely assumed that "almost all people, almost all of the time, make choices that are in their best interest or at the very least are better, by their own lights, than the choices that would be made by third parties." (Sunstein & Thaler 2003: 1163). To illustrate their argument, they point out that 60% of Americans are considered either obese or overweight. It is common knowledge that obesity increases the risk of heart attacks and is at the origin of many health problems that may eventually be fatal. It would be

---

<sup>2</sup> In his article 'Paternalism' in the *Stanford Encyclopedia of Philosophy*, Dworkin for instances distinguishes hard vs soft paternalism, narrow vs broad paternalism, weak vs strong paternalism, pure vs impure paternalism and moral vs welfare paternalism. <http://plato.stanford.edu/entries/paternalism/>

preposterous, sustain the authors, to claim that almost all Americans choose their diet optimally.

Stating that individuals are unable to correctly evaluate the benefit of the decision they are about to make obviously casts doubt on the rationality of decision-making process. Over the last twenty years, behavioural economics has yet offered a great deal of evidence to support this thesis (Lichtenstein & Slovic, 2006). The lack of any past or similar reference, for instance, introduces an element of arbitrariness in the evaluation. Preferences are also claimed to be endogenous to social contexts, and consequently inconsistent from one situation to another. Framing, anchors or endowment effects show that preferences are substantially shaped by external factors which seriously undermine the assumptions of individual sovereignty and preference transitivity of the rational choice theory. The authors report that patients who are told “Of those who undergo this procedure, 90 percent are still alive after five years” are more likely to agree to the procedure than those who are told “Of those who undergo this procedure, 10 percent are dead after five years” (Sunstein & Thaler 2003:1161). This is what behaviourist economists call a framing effect. It shows that the way information is displayed to the economic agents is not neutral. The results of these experimental studies naturally question individual rationality, but one would be mistaken to entirely discard the concept of rational agents. If anything, they lead to new directions of research<sup>3</sup>.

Paternalism occurs (according to the authors) when planners take decisions that positively affect the choices of the non-planners. As shown in the obesity example, people act in a sub-optimally way if they are not appropriately guided in their choices. The role played by this third party is then determinant in whether or not they assume their paternalist function. The authors claim that “the design features of both legal and organizational rules have surprisingly powerful influences on people’s choices. [They] urge that such rules should be chosen with the explicit goal of improving the welfare of the people affected by them.” (Sunstein, Thaler, 2006: 234).

Sunstein and Thaler do not introduce any moral considerations in their analysis, and the motivation of the planners is never questioned. Paternalism is not morally good or wrong, it is a logical consequence of individuals’ bounded rationality. Individual choices are formed in such a way that they are necessarily influenced by planners’ decisions. Are the planners benevolent? To the authors, planners are benevolent *by default*. They are not virtuous or altruist but if they have to choose between two equally attractive options when only one also benefits the others, they will necessarily pick that one. Because people are not fundamentally evil, every time such a configuration occurs, they act paternalistically. Most of the time, they do it spontaneously, and people are not even aware of the guidance they have benefited from. Because these situations happen all the time, we are all and constantly subject to paternalism. To illustrate this idea Sunstein and Thaler take the example of a cafeteria

---

<sup>3</sup> The idea of coherent arbitrariness of preferences developed by Dan Ariely, George Loewenstein and Drazen Prelec, for instance, is based on the following observation: absolute valuation of goods or experience is for a large part arbitrary, but relative valuations prove to be coherent. It interestingly suggests that preferences do not exist as such until they are revealed, or more precisely, until they are constructed by their enactment. Once the preference has been implemented or ‘imprinted’, related valuations are then ordered consistently (Ariely, Loewenstein, Prelec, 2003; 2006).

director arranging a line of dishes. According to the authors, she has four alternatives (Sunstein, Thaler, 2003: 1164):

1. She could make choices that she thinks would be best for the customers, all things considered.
2. She could make choices at random.
3. She could choose those items that she thinks would make the customers as obese as possible.
4. She could give customers what she thinks they would choose by themselves.

Robert Sugden pointed out that one option is surprisingly missing from the list: the cafeteria director might simply choose to display the dishes in the most economically efficient way (Sugden 2007). The authors' reasons for not including it are twofold. First they argue that planners are not necessarily looking for profit maximisation. Even directors of private companies, like school cafeterias, are not constantly subject to market pressures (Sunstein & Thaler 2003: 1165). One could add that some planners are not concerned with profit at all whilst others are exclusively profit driven, but such a distinction would inevitably raise the issue of the status of the planner: Is the planner working for a public administration, for a highly competitive private company or for a public utility company? Sunstein and Thaler do not distinguish profit driven from non-profit driven planners. Although they hardly give any reason for it, I take their position as being justified. I shall come back to this point in section 3 and argue that a situation only qualifies as paternalistic when the planners go past their functional role and act benevolently towards the non-planners.

Secondly, they claim that “market success will come not from tracking people’s *ex ante* preferences, but from providing goods and services that turn out, in practice, to promote their welfare, all things considered. Consumers might be surprised by what they end up liking; indeed, their preferences might change as a result of consumption.”<sup>4</sup> (Sunstein & Thaler 2003:1165). For Robert Sugden, consumers may have context-dependent preferences instead of *ex ante* preferences but that does not mean they are not sovereign of their choice. Dissatisfied with the idea of second-rank preferences he finds too morally connoted, he suggests abandoning the assumption of consistent preferences (Sugden 2004). What matters is for individuals to have a broad array of options to choose from “whether or not [their] choices reveal any internally consistent set of judgments about well-being” (Sugden 2004: 1016). When consumers prefer buying the item first displayed (e.g. dessert) over the last one (e.g. fruit), their choice does not express an *ex ante* preference for dessert but a willingness “to pay more for ‘dessert early in the line’ than for ‘dessert late in the line’, to an extent that outweighs their lesser willingness to pay for ‘fruit late’ rather than ‘fruit early’.” (Sugden 2007: 17). Individuals value the diversity of the opportunities proposed and endorse the responsibility of their past choices – they either regret them or not – as well as their futures ones.

---

<sup>4</sup> Planners may consider appropriate to promote values the non-planners do not currently like on the ground that they will eventually learn to like them. But to what extent can we ‘benevolently’ promote features of welfare some individuals actually do not enjoy (Arneson 1999)? Are planners legitimate in imposing or changing individuals’ preferences as they like? Would planners be legitimate in switching people’s preferences for pop music to opera, TV to theatre, table wine to fine wine? This question will not be discussed further in this paper as I do not consider the moral aspects of paternalism.

Interestingly, Sunstein and Thaler changed their mind about the possibility of a fifth option. In *Nudge*, they eventually add the following alternative: ‘5. Maximize profits, period’ but argued that the planner, now named Carolyn, would not choose this option at the expense of the children’s health. (Sunstein & Thaler 2008:2). This reasoning, I believe, contains two mistakes. By featuring the planner as a school cafeteria director, they considerably narrow down the relevance of their example. What may be true for a school cafeteria is not necessarily true for a regular cafeteria or a supermarket. But, more importantly, they conflate the *personal* benevolence of Carolyn (she loves children) with her *functional* role as a planner (providing balanced meals to children).

In the example of the cafeteria, the director has only three options. Option 4 represents the anti-paternalist option. It relies on two assumptions: the first one is that consumers have *ex ante* preferences and the second one is that planners are able to identify them. Behaviourist economists dismiss the first one. Even if one disagrees with them, it is reasonable to admit that planners have no means of knowing what these preferences would be. Consequently, this option can not be logically considered. Three options then remain. The second and third options are immediately discarded by the authors. Rejecting option 2 seems very reasonable. One does not expect a good manager to adopt a random criterion of organization. Option 3 seems at first sight similarly unattractive. It is difficult to imagine a manager possessed by the desire to make her clients fat<sup>5</sup>.

The first option, the paternalist one, is thus presented, *by default*, as the only possible one. All of the non-paternalist alternatives, conclude the authors, are not viable. From this example, Sunstein and Thaler infer that “a form of paternalism cannot be avoided, and that the alternatives to paternalism (such as choosing options to make people worse off) are unattractive” (Sunstein & Thaler 2003: 1159). I shall argue in section 3 that if alternatives to paternalism may well be unattractive, they are nonetheless viable.

### 3. Preference for Default Rules

I would like to turn now to the libertarian feature of Sunstein and Thaler’s paternalism which advocates a total freedom of choice for individuals. The authors refute the common view according to which libertarians could not accept paternalism, or that alternatively paternalists should necessarily dispel libertarianism. Though this seems to be a contradiction in terms, Sunstein and Thaler intend to demonstrate that it is not (Sunstein & Thaler 2003:1160). They argue that, although planners can steer people’s choices or even enrol them in default rules, individuals are nonetheless free to make their own choices.

The benefits of libertarian paternalism are illustrated by two cases of change in setting default rules, the case of 401 (k) employee savings plans (Choi, Laibson, Madrian, & Metrick, 2002; Madrian & Shea, 2001) and the authors’ own experience of the University of Chicago’s car park payment system. In the first case, employers decide

---

<sup>5</sup> In *Nudge*, this option is abandoned and replaced by a most unlikely ‘Maximize the sales of the items from the suppliers that are willing to offer the largest bribes’ (but Carolyn, the director, is “honourable and honest, so she does not give this option any thought” Sunstein & Thaler 2008:3).

that instead of letting their employees join a savings plan, the American 401 (k) plan, they would automatically enrol them unless they specifically stated they did not wish to be included. When eligible, employees usually receive a 401 (k) plan form that they must complete in order to join. In this case, however, employees only receive a statement in which they are told that, unless they opt-out, they are automatically enrolled in the plan. In both studies, a significant increase of the joining rates has been observed (from 49 percent to 86 percent in Madrian & Shea 2001).

In the second case, a change in the tax law made it possible for employees to pay for employer-provided parking on a pre-tax basis. The University of Chicago presumed that, although being in the employees' interest, many of them would not take the trouble to fill in and send back the form and decided to enrol all employees automatically. Those who preferred to pay with after-tax dollars were, however, free to opt out. Had the university chosen a different default rule (like an 'opt in' strategy), the authors reckon that "many employees, especially faculty members (and probably including the present authors), would still have that form buried somewhere in their offices and would be paying substantially more for parking on an after-tax basis" (Sunstein & Thaler 2003:1171).

Enrolling individuals by default rules presents obvious benefits over active choices. In some cases, in particular when decisions are technical or unusual, agents experience a strong aversion to decide and therefore procrastinate. Recent developments in behavioural economics demonstrated that minor (and sometimes irrelevant) reasons are sufficient for an individual making a decision in one way rather than in the other<sup>6</sup>. A minor effort, such as filling in a form, is often a sufficient reason not to opt in. "By contrast, a thoughtfully chosen default rule, steering them in sensible directions, is a blessing." (Sunstein & Thaler, 2003:1199). In this case, the default rule is praised as a general preference not to have to decide in a certain number of matters.

This general preference for default rules is reinforced by the belief that, in addition to being bothersome, autonomous decisions are more likely to be less efficient than the 'thoughtfully chosen default rule'. Whether individuals are rational or not, they are always interested in increasing the rationality of their behaviour (Harsanyi 1982: 42). This behavioural trend has recently been empirically observed by Christopher Hsee, Jiao Zhang, Frank Yu and Yiheng Xi who showed that, when it comes to making a decision, individuals tend to give more weight to what they call the 'rationalistic' or objective factors than to their own subjective values<sup>7</sup>. In many cases, default rules are

---

<sup>6</sup> Reason-based choice, as it has been termed, considers choice as a 'reasoned' or weighted selection of options (March 1994, Prelec & Herrstein 1991; Simonson 1989, Simonson & Nowlis 2000). This approach usually describes cases where none of the proposed options correspond to the individual's initial preferences or situations in which all options are equally attractive (Hafir, Simonson & Tversky 1993).

<sup>7</sup> They asked two groups of students to buy a stereo. Two equally expensive models are proposed, Sony A and Sony B, offering different qualities in terms of sound richness and power. The first group is told that Sony A has a greater sound richness, which corresponds to their preference, but that Sony B is more powerful, an attribute presented as objective. The second group is told that Sony A has objectively a richer sound but that they prefer the more powerful sound of Sony B, an attribute presented as totally subjective. The result of this experiment was that students who believed power an objective attribute decided to buy the more powerful model in spite of their preference for sound richness, and that the students who believed sound richness to be an objective attribute decided to buy model A instead of enjoying model B as theory would predict (Hsee, Zhang, Yu, Xi, 2003).

actually superior to personal choices. Sunstein and Thaler thus explore in their book the example of the Swedish privatization of social security in 2000. Swedish citizens initially faced 456 different available funds to compare and choose from. A strong advertising campaign encouraged them to actively choose one fund rather than picking the default portfolio which was eventually adopted by a third of them. As it happens, the default portfolio, sensibly designed by financial experts, turned out to be the most efficient, or rather the less disastrous one. Unused to buying financial assets, most Swedish citizens turned towards Swedish companies; on average the actively chosen portfolios were composed with 96.2% of shares, half of them for Swedish companies, and in particular, stock in technology and health care. During the financial bubble, these assets hugely increased in value: the most popular fund, Robur Aktiefond Contura increased by 534.2% between 1995 and 2000, giving the impression that they were solid investments. While the average portfolio of the actively chosen funds lost 39.6% in the first three years, the default fund, more diversified, ‘only’ lost 29.9% (Sunstein & Thaler 2008:151).

Associated with the complete freedom to opt out, default rules are regarded by the authors as a perfectly acceptable form of paternalism. Libertarian paternalism, as they call it, is construed as “a relatively weak and nonintrusive type of paternalism, because choices are not blocked or fenced off.” (Sunstein & Thaler 2003: 1162). In practice, however, default rules are sticky and guaranteeing the conditions of free choice is difficult. Once a person has been enrolled in a default rule, he or she tends to overestimate its benefits. This phenomenon, called the endowment effect (Loewenstein & Adler 1995), has been observed in particular by Cass Sunstein with employees who value rights only because they have been granted them in the first place (Sunstein 2002; Sunstein & Thaler 2003:1174). If opting in is a deterrent for many peoples, opting out from a default rule is even more dissuasive. To respect freedom of choice, a libertarian planner must set up a default rule that will be easy and costless for individuals to opt out. Ideally, it should not be more demanding than a ‘one-click’ procedure (Sunstein & Thaler 2008: 249).

#### 4. Benevolence and libertarian paternalism

As previously mentioned, the first misconception that underpins anti-paternalism is “that there are viable alternatives to paternalism” (Sunstein & Thaler 2003:1164). Sunstein and Thaler claim that planners have no other choice than to act benevolently. I would like to challenge this view. I want to argue here that being benevolent is entirely the choice of the planners, and that paternalism is only one possibility amongst others. Although the proportion of planners being malevolent is probably limited, the very possibility that they may be malevolent has dramatic consequences. Prior to embarking on this discussion, it is essential to stress the importance of the distinction introduced by Sunstein and Thaler between planners (they are called choice architects in *Nudge* but I shall keep the term ‘planner’ here) and non-planners. Herein lies the critical contribution of their work, although I believe they did not fully exploit it.



A planner is a person (or a group of persons) whose function dictates (Queen, judge, school teacher, director of cafeteria...) that they make decisions which will *necessarily* affect – for better or for worse – the life of a certain number of other persons she is not directly related to. To a certain extent, we are all planners and all subject to the decisions of numerous planners that we are not even aware of. The traditional distinction between public and private, or between representatives and citizens, thus becomes irrelevant. The law is only one device amongst others through which one can promote people's interests. Paternalism is a matter of individual choice rather than of a public policy. In spite of our common bounded rationality, each of us benefits in his respective function (and because of it) from a specialised and more objective knowledge of a limited number of subjects that others lacks, but which could be very valuable to them. The planner has the choice of whether or not to use this knowledge to make rules in their favour. If they do, as the authors expect them to, they are said to be 'paternalistic'. I contend, however, that a positive act by the planners towards the non-planners is not a sufficient condition to characterise a paternalistic action.

Sunstein and Thaler fail to take full advantage of their idea, and, in particular, do not distinguish the choice that a planner has to make *as a planner* from the choice she has to make *as a person*. A planner should not be considered as a particular person who makes decisions based on who she is (fat/slim, happy/sad, male/female, married/single, straight/gay, black/white) but as a person holding a functional role, and who is expected to make decisions on this basis alone. Her benevolent or malevolent feelings towards her pupils, defendants or customers are irrelevant when it comes to making a professional decision. But her skills as a planner alone do not determine the decision-making process, and more often than not, she will have to make personal choices in addition to professional ones. If a doctor prescribes medications to save a patient's life, she does not act paternalistically but according to her functional role. Now suppose that the patient is distraught by the announcement of having a potentially deadly disease and that the doctor, feeling sorry for him, walks him back home or gives him a hug. In that case, the doctor definitively acts paternalistically. Whether or not a planner has the goal of promoting other people's welfare or not is therefore not relevant to our discussion. Sunstein and Thaler were thus right not to distinguish planners who were profit-driven from those who were not. What really characterizes paternalism is not the positive impact of the planners' functional decisions but the sheer benevolence with which they act beyond their professional or institutional assignment.

Suppose that the cafeteria director decides to order 15 dishes from the more to the less profitable ones and that she finds out that three sets of 5 dishes are equally profitable. She knows that some of them are healthier than others. As a planner, she is indifferent to the inner order of each set. (I am assuming here that promoting consumers' health is not part of her job.) How will she arrange her display? She has only three alternatives. The first possibility is that, in addition to being *functionally* indifferent to the way each set is displayed, she is *personally* unconcerned by her consumers' health. In that case, she orders the dishes at random. The second possibility is that, despite being functionally indifferent to equally profitable dishes, she cares about the consumers' health and makes it a point of honour to provide them with healthy food. She accordingly orders the dishes in such a way that the consumers will preferably buy what is best for them. This is a case of paternalism and the option selected by

Sunstein and Thaler. In the third and final possibility ('She could choose those items that she thinks would make the customers as obese as possible'), she is motivated by malevolent or envious feelings. Within each equally profitable set of dishes, she will then order fat and sweet dishes first.

The authors hold that the cafeteria director will *necessarily* pick the benevolent option. I wish here to challenge this view and demonstrate that paternalism is *not* unavoidable. The second option proposed by the authors ('She could make choices at random') is a viable choice since it does not undercut her functional status nor strip her of all credibility. When she makes this choice, she has already acted *as a planner*. She has fulfilled her professional responsibility. Ordering the dishes at random is not only a possible but a worthwhile choice since it is the simplest and the quickest one. The third option ('She could choose those items that she thinks would make the customers as obese as possible') is, for the same reason, equally possible. Once she has fulfilled her function as director of the cafeteria, she has no more reason *a priori* to be benevolent than to be malevolent towards her clients<sup>8</sup>. Imagine, for instance, that the cafeteria director is unhappily overweight. A subjective approach to welfare, such as the one developed by Pigou on relative incomes (Pigou 2002), would conclude that her welfare does not depend on her absolute weight, but on her deviation from the mean or more simply from social pressure<sup>9</sup>. One can easily imagine that thinking that her clients will get fatter makes her feel less miserable. This gives her a real incentive to highlight fat dishes when organizing her display<sup>10</sup>.

Option 3 is therefore also a viable alternative to paternalism. From a libertarian vantage point, one cannot deny this opportunity to the planner without infringing upon her self-ownership. Self-ownership excludes the possibility of individuals not having the choice not to be benevolent<sup>11</sup>. This notably implies that, as a free person, I don't have any moral obligation to others and that I cannot be forced to benevolence either directly or through a redistributive system (Nozick 1974). An immediate objection to this assertion is, of course, that such malevolent behaviour is exceptional and need not, therefore, be taken into account. I would argue, firstly, that malevolent behaviour is considered exceptional because it is considered as virtually nonexistent by economists and not the other way round. So far, altruism has received much more attention from economists than envy (Kirchsteiger 1995; Mui 1995). But, more importantly, stating that an action is 'benevolent' or 'malevolent' is a risky

---

<sup>8</sup> The example of the cafeteria is a bit confusing since it is possible that the cafeteria has a public health mission, e.g. schools cafeteria.

<sup>9</sup> A recent survey on well-being and income in South Africa (Kingdon and Knight 2007) shows that at equal income, a White feels relatively deprived (being a 'poor' member of the rich race) when a Black feels satisfied or very satisfied (being a 'rich' member of the poor race). Should the policy maker help in priority those who feel deprived over those who are satisfied?

<sup>10</sup> It does not mean however that she will deliberately try to harm the consumers. Planners, and more generally human beings, often subconsciously act malevolently. If asked, they would probably genuinely deny having any evil intentions.

<sup>11</sup> Sunstein and Thaler naturally acknowledge this libertarian principle. They however argue that if benevolence cannot be forced, it can be artificially prompted while freedom of choice is respected. The different systems of organs donations adopted by European and American countries illustrate that possibility. When Americans need to actively express their consent to be removed organs when they die, some European countries such as France presuppose that French implicitly consent to give organs unless explicitly stated. The authors reckon that the French setting of the default rule does not infringe on individuals' freedom of choice and that a change in the American policy could save thousands of lives each year. They call it *libertarian benevolence*.

declaration to make; cases exist where a single action can have multiple opposing consequences on differing populations. The director of the cafeteria may, for instance, decide to promote gluten-free dishes out of sympathy for those who are allergic to it, at the expense of all those who are not. On the other hand, she could also decide to promote iron-supplemented food because it improves most people's diet even though it is highly contraindicated for those who suffer from hemochromatosis<sup>12</sup>. The addition of iron, unbeknownst to the consumers, has much more negative consequences for the minority affected by this condition compared to the small benefit it gives to the immense majority of people.

Default rules, however benevolent they may be, propose a unique ideal of welfare which is unlikely to meet the diversity of situations. Should the initial intention of the planner be primarily considered to characterise the 'benevolence' of her action? If so, a xenophobic measure could be seen as benevolent (and hence paternalistic) as it promotes the interests of the planner's community despite the consequences to other communities. A utilitarian may argue, however, that in such cases, a benevolent action is one that takes into consideration the general balance of benefits and disadvantages produced by the action and considering each individual as equal. But the evaluation of the advantages and disadvantages conjointly produced remains an intricate business<sup>13</sup>. In all the cases where individuals have different interests but are subject to the same decision of a planner, it is difficult to assign a 'good' or a 'bad' meaning to the planner's decision.

But, apart from the fact that utilitarianism is difficult to combine with the authors' libertarianism (section 4), this kind of response misses the fundamental point the authors failed to notice: individuals allow themselves to be enrolled in default rules only when they *trust* the planners. Even in the very implausible case where all planners are inspired by good intentions, non-planners will not all benefit from the plan; some of them may even be harmed by it. Since the default rules that systematically enrol everyone do not clearly identify those who will benefit from it from those who won't, non-planners have no choice but to *trust* the planner. In many cases, this trust is entirely justified, as illustrated by the Swedish default portfolio. In some others, however, complying with the planners is not recommended. Consider now the case when a non-planner does not know the planner and vice-versa (a first-time consumer in the cafeteria for instance). Even if this person is naturally trusting, she cannot exclude the possibility of an unintentional harm that could result from complying to the plan or the default rule. Nor can she exclude the possibility that the planner may have acted randomly<sup>14</sup>. Should she blindly comply with the default rule? This would be irrational. Default rules are accepted by individuals firstly because it prevents them from procrastination and secondly because they supposedly embody a higher rationality. If people have to systematically assess the default rules before adopting them, they lose all interest. Introducing scepticism about the benefits of the

---

<sup>12</sup> Hemochromatosis is a hereditary disease characterised by an excessive quantity of iron accumulated in the body and which disrupts normal functions of organs such as liver, heart and pancreas. Iron supplemented food, like cereals or 'healthy' biscuits, worsens this condition.

<sup>13</sup> To set up default rules, planners can either use a technical process, the economists' Cost Benefit Analysis (CBA), or alternatively the more intuitive rules of the thumb. CBA is, however, difficult to implement in cases where preferences are incoherent (Sugden 2007).

<sup>14</sup> When the United-States tried to implement a private social security system such as the one established in Sweden, the default portfolio had been randomly selected.

default rules (or the benevolence of the planners) basically ruins the system of libertarian paternalism.

## 5. Voluntariness and Freedom of Choice

I turn now to the libertarian feature of Sunstein and Thaler's paternalism that they describe as "a relatively weak and nonintrusive type of paternalism" (Sunstein and Thaler 2003:1162). Libertarian paternalists, like any other libertarian theorists, defend absolute freedom of choice. Mandatory rules should be the exception. But unlike traditional theories of libertarianism, based on self ownership, libertarian paternalism *exclusively* relies on freedom of choice. It is rooted in a very weak notion of voluntariness that makes it closer to hard paternalism.

According to Mill's first principle, "the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant." (Mill, 1989: 13). Mill's seemingly unambiguous claim, however, fostered different views on the exact meaning of the expression 'against his will'. Is Mill referring to an immediate *desire* or to a deliberate *will*? And, if Mill really means *will*, as Gerald Dworkin suggests it, could this *will* be irrational<sup>15</sup>? For example, could someone autonomously and deliberately want to alienate his freedom? The degree of *voluntariness* granted to individuals' actions determines whether it falls into the category of a soft or a hard type of paternalism.

Soft or weak paternalism typically describes situations of involuntary acts or misinformed decisions that must be interfered with for the person's own good. Typically, a soft paternalist can legitimately prevent a person from using the saltcellar if she has not been informed that it contains arsenic instead of salt. Similarly, a person under the influence of drug or alcohol can be temporarily prevented from making decisions that he/she may bitterly regret the next day (Feinberg 1986). On the other hand, if I am appropriately informed of the risks I incur from crossing a dangerous bridge, smoking, or duelling, no external interference is ever justifiable to stop me performing or carrying out such an action<sup>16</sup>. Soft paternalism re-establishes what some regard as an essential feature of liberalism: the right to act irrationally.

---

<sup>15</sup> In his first article on paternalism, Dworkin refutes this possibility and misleadingly associates voluntariness, autonomy and rationality. He changes his mind in *Paternalism, Second Thoughts*. A paternalist action, thus reconsiders Dworkin, characterises a usurpation of decision making (Dworkin 1988: 123).

<sup>16</sup> In some cases, however, it's not necessarily true. To understand this, it is important to introduce Feinberg's distinction between *reasonable* and *rational*. Person A can judge person B's action *unreasonable* either because she believes person B's risk assessment incorrect or because she would not act herself on such a risk level. Knowing that the bridge has a probability of 60% to collapse, for instance, B decides nevertheless to cross it. A disapproves B she believes *unreasonable* but has no right to stop her. Now, suppose a different situation where A knows for sure that the bridge will fall down if she attempts to cross it but is still resolute to cross it to visit a friend. B's action is then judged *irrational*, and B insane or mentally impaired and A has the moral duty to stop her. The reason of this is that one cannot simultaneously want to cross a bridge and die on it. In this case, but in this case only, a coercive action on a voluntary and purely self-regarding action is considered as legitimate (paternalism type II).

Arneson argues against Dworkin that mistakes and wrongdoings cannot be considered as involuntary even if they seem objectively irrational. He considers paternalism illegitimate if individual choice matches the description of what he calls a weak notion of voluntariness. A person acts voluntarily “if and only if his choice of the act (a) would not be abandoned if he were apprized of all the act’s unforeseeable consequences, (b) does not proceed from an emotional state so troubled as to preclude the full use of the reasoning faculty, and (c) does not occur under conditions of external coercion or compulsion.” (Arneson, 1980: 482). It has been argued that this liberal form of paternalism could equally be understood as a soft paternalism as well as a soft anti-paternalism (Feinberg 1986). Soft paternalism indeed only interferes with impairing and external factors to the will. Technically speaking, it does not limit freedom.

Hard paternalism, on the contrary, is by nature coercive. To Gerald Dworkin, for instance, paternalism covers “the interference with a person’s liberty of action justified by reasons referring exclusively to the welfare, good, happiness, needs, interests or values of the person coerced” (Dworkin 1972:65). It legitimizes interference from voluntary and well-informed actions. Among all the possible examples, Dworkin mentions laws requiring motorcyclists to wear safety helmets, laws regulating sexual conduct, such as homosexuality between consenting adults in private, laws forbidding people from swimming at a public beach when lifeguards are not on duty and laws compelling people to save a fraction of their income in retirement plans. Although some of these examples could not be considered as paternalistic today, either because they involve a third element (the economic cost borne by the society) or because they are no longer considered self-destructive or wrong (homosexuality), these examples are representative of an essential feature of hard paternalism: the source of risk assessment. Whilst supporters of soft paternalism consider risk assessment as the risk taker’s responsibility, risk is externally assessed by the paternalistic agent<sup>17</sup>.

Sunstein and Thaler’s libertarian paternalism borrows features from these two opposite approaches. Its spirit is libertarian: individual choice is always respected, however self-destructive it may be. The authors are in favour of governmental rules that guide rather than compel whenever this is possible. Sunstein and Thaler thus defend a privatization of social security, retirement plans, education (for those countries which have not yet done so) and even of marriages. Bans, such as bans from riding motorcycles without a helmet, should be limited as much as possible (Sunstein and Thaler 2008:232-33). Every time a hindrance on individual freedom of choice can be avoided through a simple ‘nudge’, the authors support it by encouraging initiative – which could be either public or private – over regulation.

---

<sup>17</sup> In cases like the obligation to wear seat belt or helmets when driving a car or a motorcycle, personal and public risk assessments converge. Most of the people do not feel coerced as they voluntarily fasten their seat belt. In forcing their children to wear helmets when cycling, or in wearing one themselves while it is not compulsory, parents implicitly recognise the voluntary feature of this rule. These rules can be both supported by partisans of soft and hard paternalism. But in some other cases, however, the difference between a person’s subjective risk assessment and the paternalistic authority prompts coercion. In the case of the swimming ban cited by Dworkin, good swimmers’ liberty is curtailed because of the potential risk incurred by bad and irresponsible swimmers.

Sunstein and Thaler hold that dogmatic libertarianism and dogmatic paternalism are incompatible. It is true to a certain extent, but most libertarians and paternalists are not dogmatic. Besides, being dogmatic is not what matters here. The degree of compatibility between libertarianism and paternalism essentially relies on the degree of voluntariness one confers on an individual's actions. A 'dogmatic' libertarian would thus be someone who considers all human actions as fully voluntary, and hence all intervention as illegitimate. A 'dogmatic' paternalist, on the contrary, would be someone who considers all human actions as potentially involuntary and hence subject to legitimate interference. Libertarians and paternalists, however, largely agree that some actions are fully voluntary whilst some others are undoubtedly involuntary. I can, for instance, consider that an action is *involuntary* - and subject to State interference - if some information has been hidden from me. Most libertarians agree that contracts can be broken by a third party when one party has been intentionally deceived by the other. This is what we previously referred to as soft paternalism. In this case, libertarian paternalism is not an oxymoron. Now, the question for us is the following: Is Sunstein and Thaler's libertarian paternalism an oxymoron?

Sunstein and Thaler's libertarian paternalism is not a form of soft paternalism. Their analysis, based on behavioural studies, shows, however, that most individual choices are sub-optimal because their valuations depend on arbitrary anchors and frames. Their actions reflect the context in which they have been carried out rather than the proper preferences of the agents who welcome an external aid to help them make the right choice. To a certain degree, most of their actions are involuntary: they want to be thin but eat fatty foods, they want to have a comfortable retirement but don't save enough money, they want to donate their organs but don't say so etc. Actions are not made involuntary because information is hidden but because individuals misuse it. Bounded rationality is at the heart of Sunstein and Thaler's paternalism. Such a weak concept of voluntariness usually characterises hard paternalism. The difference between them and Dworkin is, however, that paternalistic interference should never be coercive. Individuals should always be offered alternatives. The libertarian feature of their theory relies exclusively on freedom of choice. Yet there is an incongruity at the defending absolute freedom of choice and a weak concept of voluntariness both at the same time. By discarding self-ownership, the authors actually undermined the possibility of libertarian paternalism.

The number of opportunities a person gets is not relevant to assess her degree of freedom. Let me explain this by borrowing an example originally presented by Locke<sup>18</sup>. A prisoner is brought into a cell and is convinced that all the doors are locked when actually one has been left unlocked. Is this person free to get out? A libertarian like Nozick would answer in the affirmative. Technically, nobody physically stops him from getting out. He is not *coerced* to stay inside. A behaviourist, or for that matter a philosopher of autonomy like Dworkin, would argue, however, that since the prisoner is not aware of this opportunity, he is not any freer than if all the doors *were* actually locked. In the same way as in the cafeteria, where the order of the dishes is more significant to the consumers than their presence, the information given about the doors is more significant to the prisoner than their real state. If individuals were as rational as the economists suppose them to be (if they

---

<sup>18</sup> I borrow this example myself from Dworkin who used it to support his distinction between liberty and autonomy (Dworkin 1988:14).

were ‘Econs’ to borrow Sunstein and Thaler’s terminology), they would not be subject to framing or anchor effects. They would correctly use the information they are given to maximise their welfare and all their actions would result from deliberate and well-informed decisions. Their actions would then be regarded as fully *voluntary*.

Having the possibility of opting out from default rules is not necessarily a sign of increased freedom. Libertarians hold that one is free insofar as one has different possible alternatives. Freedom and voluntariness are conflated because, except in the case of coercion, all actions are presumed voluntary (Nozick 1974). Introducing non-voluntary actions, however, undermines the importance of the number of options. Are only significant the options that are voluntarily chosen. In some cases, a larger number of opportunities alienates rather than promotes freedom (Dworkin 1988). Libertarians thus defend the right of a person to voluntarily restrain her future freedom of choice<sup>19</sup>. Covenant marriages, for instance, are specially designed for those who, usually for religious reasons, want to make divorce extremely difficult (Sunstein & Thaler 2008:220)<sup>20</sup>. A possible justification for this kind of contract is that I want to restrain my future non-voluntary actions. Ulysses thus commands his men to bind him to the mast so he would not be tempted to jump into the sea when the sirens sing. Similarly, I may not feel coerced if all alternatives are banned except the one I would have voluntarily chosen. Olsaretti’s example of the wired city perfectly exemplifies this case: “Wendy is the inhabitant of a city fenced with electrifying wire, which she is unfree to leave. However, her city has all that anyone could ever ask for, and Wendy, who is perfectly happy with her life there, has no wish of leaving it. She voluntarily remains in her city.” (Olsaretti 2004:138).

In Sunstein and Thaler’s 401 (k) plan example, rationality imposes adherence to the default rule. Everyone wants to receive a pension but there is a great temptation to indefinitely postponing saving for it. Our preference for the present prompts us to spend when we would really have preferred to save: what is in play in pension plans is man’s weakness of will. Planners who automatically enrol everyone into a retirement plan assume that not actively joining the plan is a non-voluntary action rather than a voluntary one. Paternalism can legitimately be accepted here. And, because some individuals might voluntarily not join the plan (e.g. they know that they will die before being eligible for retirement), it is essential to set an opting-out option. But such an option must be able to discriminate between a voluntary and non-voluntary desire to pull out. One may, for instance, require a medical consultation. The mistake the authors make is, I believe, to reduce the cost of opting out to zero (a ‘one-click’ procedure) to prop up freedom of choice. The easier it is to opt out, the more ineffective the default rule becomes. In this particular example, if individuals are totally free to opt out, all those who have an excessive preference for the present (the poor, gamblers, spendthrifts, alcoholics...) will turn down their employer’s offer. In

---

<sup>19</sup> Such a possibility would be discarded by liberals such as John Stuart Mill for whom “The principle of freedom cannot require that [man] should be free not to be free. It is not freedom to be allowed to alienate his freedom.” Partisans of paternalism diverge on this question. For Feinberg, a person may voluntarily commit or even transfer to someone else his future ability to autonomously choose. On the contrary, Dworkin claims that individuals should never be allowed to definitively impair or renounce their freedom of choice.

<sup>20</sup> Giving up the capacity to opt out from automatic enrolment is a possibility that paradoxically cannot be excluded as it would infringe the individuals’ freedom of choice. Sunstein and Thaler thus seriously advocate the possibility for medical patients to waive their right to sue their doctor for negligence in return for a lower price of care (Sunstein & Thaler 2008, chapter 14).

other words, all those who especially need the default rule will not benefit from it. In that case, libertarian paternalism is an oxymoron.

## 6. Conclusion

Because it is pragmatic and intellectually challenging, Sunstein and Thaler's concept of libertarian paternalism may appeal to a wide range of thinkers. The utility of implementing costless but efficient default rules, such as tacit consent in organ donations, deserves to be highly defended, not only from a pragmatic point of view but also from liberal (and even libertarian) theoretical perspective. In succeeding to do so, Sunstein and Thaler significantly contribute to the renewal of liberalism, which had too often been prisoner of the dichotomy 'the State versus the Individual'. Basing their approach on behavioural economics, the authors brilliantly show that the real debate about paternalism and interference in individuals' freedom is not restricted to political philosophy. In particular, the idea of dividing individual choice into two categories – the 'planners' who make decisions that they know are decisively influential to other people's choices, and the 'non-planners' who are surreptitiously led to make the best decision for themselves – is very constructive.

The authors, however, seem not to take full advantage of this distinction and reduce the planner to a person (preferably a benevolent one) instead of dividing the planner's choice into a functional part and a personal part. This further refinement would have permitted, as we have shown, (1) to narrow down the area of paternalistic actions to the sole 'personal' decisions of the planner and (2) to circumscribe the uncertainty of benevolent actions to this area. Planners, including good planners, are totally free to choose whether or not to be paternalistic. When it comes to the personal component of the planner's choice, a random decision is – from her point of view – as good as a paternalistic one. In the absence of clear information, default rules set up outside the functional role of the planner should not *a priori* been considered as paternalistic.

The argument according to which non-planners are still free to opt out if they believe that the default rule is not in their particular interest (and hence that the default rule is always worth implementing) presents another difficulty. Because the decisions of non-planners are presumed involuntary (and are supported by behavioural evidences), their freedom of choice cannot reasonably be comprehended as a matter of opportunities of choice. The easier it becomes to opt out of the default rule, the greater is the risk that the default rule will not help those who need it most. The libertarian concern of Sunstein and Thaler should be exclusively focussed on the planners' choice since they are the only ones able to make voluntary choices. Ironically, their only option is to be benevolent.



## REFERENCES

- Ariely D., Loewenstein G., and Prelec D., (2003), ‘Coherent Arbitrariness’: Stable Demand Curves Without Stable Preferences’, *The Quarterly Journal of Economics*, Vol.118, pp.73-105.
- Ariely D., Loewenstein G., and Prelec D., (2006), ‘Tom Sawyer and the Construction of Value’, *Journal of Economic Behavior & Organization*, Vol.60, pp.1-10.
- Arneson R.J. (1980), ‘Mill versus Paternalism’, *Ethics*, Vol.90, No.4, July, pp.470-489.
- Arneson R.J., (1999), ‘Human Flourishing versus Desire Satisfaction’, *Social Philosophy and Policy*, Vol.16, No.1, (Winter), pp.113-142.
- Camerer C.F., (2006), ‘Wanting, Liking, and Learning: Neuroscience and Paternalism’, *The University of Chicago Law Review*, Vol.73, No.3 (Winter), pp.87-110.
- Dworkin, G., (1972), ‘Paternalism’, *The Monist*, 56, pp.64-84.
- Dworkin G., (1988), *The Theory and Practice of Autonomy*, Cambridge University Press.
- Feinberg J., (1986), *The Moral Limits of the Criminal Law*. Vol. 3, *Harm to Self*, Oxford University Press.
- Harsanyi J.C., (1982), ‘Morality and the theory of rational behaviour’, in A. Sen and B. Williams, *Utilitarianism and Beyond*, Cambridge University Press, pp. 39-62.
- Hsee C.K., Zhang J., Yu F., and Xi Y., (2003), ‘Lay Rationalism and Inconsistency Between Predicted Experience and Decision’, *Journal of Behavioral Decision Making*, Vol.16, pp.257-272.
- Lichtenstein S. and Slovic P., (2006), *The Construction of Preference*, Cambridge University Press.
- Loewenstein G. & Adler D., (1995), ‘A bias in the prediction of tastes’, *Economic Journal*, Vol.105, pp.929-937.
- Mill J.S., (1989), *On Liberty and other writings*, Cambridge University Press.
- Mui V.-L., (1995), ‘The economics of envy’, *Journal of Economic Behavior & Organization*, Vol.26, pp.311-336.
- Nozick R., (1974), *Anarchy, State and Utopia*, Basic Books.
- Pigou A.C., (2002), *The Economics of Welfare*, Transaction Publishers.
- Sugden R., (2004), ‘The Opportunity Criterion: Consumer Sovereignty without the Assumption of Coherent Preferences’, *The American Economic Review*, Vol.94, No.4 (September), pp.1014-1033.
- Sugden R., (2006), ‘Taking unconsidered preferences seriously’, in S. Olsaretti (ed.), *Preferences and Well-being*, Royal Institute of Philosophy Supplement, pp.209-232.
- Sugden R., (2007), ‘Cost-Benefit Analysis as Market Simulation. A New Approach to the Problem of Anomalies in Environment Evaluation’, *Discussion Paper RFF DP 07-28-REV*, *The Frontiers of Environmental Economics*, 41p.
- Sunstein C.R. and Thaler R.H., (2003), ‘Libertarian Paternalism Is Not an Oxymoron’, *The University of Chicago Law Review*, Vol.70, No.4, (Autumn), pp.1159-1202.
- Sunstein C.R. and Thaler R.H., (2006), ‘Preferences, Paternalism, and Liberty’, in S. Olsaretti (ed.), *Preferences and Well-being*, Royal Institute of Philosophy Supplement, pp. 233-264.

Sunstein C.R. and Thaler R.H., (2008), *Nudge. Improving Decisions About Health, Wealth and Happiness*, Yale University Press.  
Thaler R.H. and Sunstein C.R., (2003), 'Libertarian Paternalism', *The American Economic Review*, Vol.93, No. 2, May, pp. 175-179.