



**UNIVERSIDADE CATÓLICA PORTUGUESA**

Instituto de Estudos Políticos

**Considering restrictions on political liberty and individual autonomy:**

A critical analysis of Libertarian Paternalism and the effect of behavioural  
sciences on policy

Author: Maria Archer

Student Number: 100514006

**Master Thesis in Political Science and International Relations**

Dissertation Advisor: Professor Hugo Chelo

Lisboa, 2018



## ***ACKNOWLEDGEMENTS***

First, I'd like to offer my thanks to Professor Hugo Chelo who was patient enough to help me with this long endeavour. Without his input this dissertation would fall to ground.

I'm very thankful to Professor Ivone Moreira who allowed me to invade her office from time to time and who was always available to help me with any questions or requests I might have had.

Finally, I'd like to show my gratitude to all my family who has never ceased to support me in this, seemingly, endless journey. Here, a special thanks to Barry for always being available to read the English text.



## **ABSTRACT:**

*For the past decade, a new form of paternalism has been emerging on the policy stage. Unlike ‘traditional paternalism’, which sought to make people conform to religious or moralistic notions of goodness, this form of ‘new-paternalism’ seeks to make people better off by their own judgement. For the better part of fifty years, behavioural sciences have been challenging assumptions of human behaviour and rationality. As such, the neo-paternalists seem to use behavioural economics and psychology’s insights to justify and legitimize their paternalistic interventions.*

*Behavioural Law and Economics are now involved in the process of policymaking, contributing to the relatively new field of Behavioural Public Policy. Libertarian Paternalism has become one of the most famous forms of this so called ‘new-paternalism’. Its advocates claim their proposal of soft paternalism is libertarian, in the sense that it does not restrict freedom of choice and action, it does not coerce nor force, but rather it “nudges” citizens on their subjectively preferred path. The purpose of this dissertation is to analyse these claims by assessing the ‘new’ paternalistic assumptions, by studying some of their proposed initiatives and the possible implications that those might have on political liberty and decisional autonomy focussing on some liberal ideas from John Stuart Mill and Isaiah Berlin.*

Dissertation Word Count (excluding footnotes): 30 923

Dissertation Word Count (with footnotes): 35 361

Abstract Word Count (208)



## ***TABLE OF CONTENTS***

<b>Introduction</b>	<b>1</b>
<b>Chapter 1: On Liberty and Autonomy</b>	<b>8</b>
1.1 John Stuart Mill's ' <i>On Liberty</i> '	<b>9</b>
1.2 Isaiah Berlin's ' <i>Two Concepts of Liberty</i> '	<b>16</b>
1.3 Liberty for the Libertarian Paternalists	<b>21</b>
<b>Chapter 2: Behavioural Sciences and the Origins of Behavioural Public Policy</b>	<b>26</b>
2.1 The Emergence of Behavioural Sciences	<b>26</b>
2.1.1 Assumption of Rationality in Economics	<b>27</b>
2.1.2 Bounded Rationality	<b>29</b>
2.1.3 The Emergence of Behavioural Economics	<b>32</b>
2.1.4 Dual Process Theory	<b>33</b>
2.1.5 Behavioural Economists on Heuristics and Cognitive Biases	<b>35</b>
A. Representativeness, Availability and Anchoring	<b>37</b>
B. Loss Aversion and Framing	<b>39</b>
C. Status Quo and Present Bias	<b>41</b>
2.2 Behavioural Welfare Economics and the Call for Paternalism	<b>43</b>
2.3 Libertarian Paternalism	<b>49</b>
2.3.1 Applications of Choice Architecture	<b>51</b>
A. Defaults	<b>54</b>
B. Expect Error	<b>57</b>
C. Informational Interventions that Target our System 1	<b>57</b>
<b>Chapter 3: Libertarian Paternalism in Practice</b>	<b>60</b>
3.1 In ' <i>Nudge: Improving Decisions about Health, Wealth and Happiness</i> '	<b>60</b>
3.1.1 The Cafeteria	<b>61</b>
3.1.2 Retirement Plans and The Save More Tomorrow Scheme	<b>65</b>

3.2 The British Movement	70
3.3 The Behavioural Insights Team in Practice	74
3.3.1 Organ Donation	74
3.3.2 Employment	76
3.3.3 Education	79
<b>Chapter 4: Libertarian Paternalism and Liberty</b>	<b>84</b>
4.1 Existing problems with Libertarian Paternalism (in practice and in theory)	84
4.1.1 Interpreting the Empirical Evidence	84
4.1.2 Notes on Rationality	87
4.1.3 Libertarian Paternalists' definition of Rationality	89
4.1.4 On Preferences	90
4.1.5 The Libertarian Paternalists' Two-Selves Picture	93
4.2 Problems arising with Libertarian Paternalism in relation to political liberty and decisional autonomy	96
4.2.1 Nominal Freedom or Real Freedom	96
4.2.2 The Importance of Experience and Learning	99
4.2.3 Transparency and Salience	102
4.2.4 The Slippery Slope of Libertarian Paternalism	104
4.3 Libertarian Paternalism. Is it Libertarian? Is it Paternalistic? Or is it, simply put, Welfarist?	107
<b>Conclusions</b>	<b>110</b>
<b>Bibliography</b>	<b>122</b>

## ***INTRODUCTION***

Recently, policymaking within different democratic governments has been influenced by the rise of new forms of paternalism. Unlike traditional paternalism, which looked to make people conform to religious or moralistic notions of goodness, this new-paternalism seeks to make people better off by “*their own judgements*”. Behavioural sciences have been challenging assumptions of rationality by trying to better understand and analyse human behaviour and decision-making. Its research and evidence have led to the proposal of these new emerging paternalistic interventions<sup>1</sup>.

Findings from behavioural psychology and the emerging field of behavioural economics seem to justify, for some, the institution of soft forms of paternalism. Libertarian Paternalism has become the most famous of these proposals, gathering more and more support within different parts of the world. Their proposal is said to be libertarian in the sense that it does not restrict freedom of choice and action, it does not coerce nor force, but rather “nudges” citizens on their “subjectively” preferred path.

At first glance, one might see the term Libertarian Paternalism as sounding like an oxymoron, but its advocates do not perceive it as such<sup>2</sup>. Instead they contend that it is possible to be paternalistic without being coercive and without restricting freedom of choice. It is my purpose to challenge this claim. It is because they use such a narrow meaning of negative liberty – the mere preservation of choice and not the capacity to make a decision<sup>3</sup> - that Libertarian Paternalism can become an oxymoron<sup>4</sup>.

---

<sup>1</sup> Mitchell, G. (2005). *Libertarian Paternalism is an Oxymoron*. FSU College of Law, Public Law Research Paper No. 136; see also Ginsburg, D. H.; Wright, J. D. (2012). Behavioural Law and Economics, Its Origins, Fatal Flaws and Implications for Liberty. *Northwestern University Law Review*, Vol. 106, No. 3: 1033-1090.; Whitman, D. G. (2010). “The Rise of New Paternalism”. *CATO Unbound: A Journal of Debate*, April Issue.

<sup>2</sup> Sunstein, C. R.; Thaler, R. H. (2003a). Libertarian Paternalism is not an Oxymoron. *The University of Chicago Law Review*, Vol. 76, No. 4: 1159-1202.

<sup>3</sup> Ginsburg and Wright (2012)

<sup>4</sup> Oxymoron here is understood as a simple contradiction. The term is used in reference to Sunstein and Thaler’s first paper introducing their thesis in 2003: *Libertarian Paternalism is not an Oxymoron*.

## Why the relevance of the topic?

In a contemporary context discussing Libertarian Paternalism is not merely a theoretical exercise. For almost a decade, libertarian paternalistic techniques have been used in the political domain. With the results from these policies in hand, more governments are now showing interest in the application of this proposal, resulting in different behavioural public policy units being consistently integrated into government departments, many adopting (not exclusively) policy interventions proposed by libertarian paternalists<sup>5</sup>.

In 2015 the World Bank Development Report retained a large interest in behavioural public policy and initiated the Global Insights Team. The next year the European Commission released a report that analysed the application of behavioural public policies within various European countries<sup>6</sup>. But what are the underlying assumptions behind these types of policies? How do public officials look at citizens under light of the emerging behavioural evidence?

“[W]e are woefully muddled information processors who often stumble along ill-chosen shortcuts to reach bad conclusions. In their book *Nudge*, Thaler and Sunstein jokingly compare us with Homer Simpson, a character prone to bumbling stupidity, in order to justify governmental paternalism that protects us from ourselves”<sup>7</sup>.

As we shall see, it is due to the emergent behavioural evidence that libertarian paternalists tend to look at individuals as two selves<sup>8</sup>: a rational-self, limited by the

---

<sup>5</sup> Oliver, A. (2013a). From Nudging to Budging: Using Behavioural Economics to Inform Public Sector Policy. *Journal of Social Policy*, Vol. 44 No. 4:685-700.

<sup>6</sup> Lourenço, J.; Ciriolo, E.; Almeida, S.; Troussard, X. (2016). *Behavioural Insights Applied to Policy: Report 2016*. Brussels: European Union.

<sup>7</sup> Gigerenzer, G. (2016) Taking Heuristics Seriously in A. Samson (Ed.) *Behavioural Economics Guide 2016*.

<sup>8</sup> Rebonato, R. (2012). *Taking Liberties: A Critical Examination of Libertarian Paternalism*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.

cognitive biases of an unreflective self. The new-paternalists' "implicit model of a human decision maker is that of an inner rational agent interacting with the world through an error-prone psychological shell"<sup>9</sup>. For the libertarian paternalists our rational-selves detain our "true" preferences and, in turn, these preferences are being constrained because our less than rational-self acts according to automatic, non-deliberate preferences. It is with this belief that Libertarian Paternalism justifies itself.

### **Working Definitions**

#### ***Libertarian Paternalism:***

Libertarian paternalists assert that they have found a way of combining liberal theory with paternalistic theory, they offer Libertarian Paternalism as "a real Third Way"<sup>10</sup>, the middle ground between extreme opposites. These ideas have originally been developed by law professor Cass R. Sunstein and the economist Richard Thaler.

In *Why Nudge: The Politics of Libertarian Paternalism* (2014), Sunstein explains and defends his thesis to the libertarian by creating a kind of "dialogue" between his position stance and that of John Stuart Mill<sup>11</sup>. He explains that, for Mill, individuals should be left alone to do as they wish unless their actions cause harm to others, and that people are in a much better position to act in self-interest than any outsider. It follows, that public officials lack information to overrule freedom of choice on grounds of promoting one's own well-being.

There are other arguments in favour of Mill's liberty, however, for Sunstein, this is the most important one, for he claims this principle as empirically false. In Sunstein's

---

<sup>9</sup> Sugden, R. (2017). Do people really want to be nudged towards healthy lifestyles? *International Review of Economics*, Vol.64, No. 2:113-123. p. 116

<sup>10</sup> Sunstein, C. R.; Thaler, R. H. (2009). *Nudge: Improving Decisions about Health, Wealth and Happiness*. New York: Penguin Books. p. 253

<sup>11</sup> Sunstein, C. R. (2014a) *Why Nudge? The politics of Libertarian Paternalism*. New Haven and London: Yale University Press

view, what behavioural psychology has shown is that individuals often don't know what they truly desire and what is best for themselves. The new-paternalists, with their two-selves picture, introduce the concept of *internalities* – harm done to oneself (more precisely one's "true" self)<sup>12</sup> as a form to justify paternalism and to challenge what Sunstein calls the "Epistemic Argument of the harm principle"<sup>13</sup>.

Within the emerging fields of behavioural sciences different empirical evidence has shown that human beings, more often than not, err. Theories of *bounded rationality*, of *heuristics* and *cognitive biases* have emerged and have come to play a significant role in economic theory and public policy. For paternalists, these human cognitive "failures" are what gives legitimacy for the state to intervene in order to promote what individuals "truly" want – welfare<sup>14</sup>.

To the libertarian paternalist welfare can be achieved by use of "nudges", which they define as: "initiatives that maintain freedom of choice while also steering peoples' decisions in the right direction (as *judged by people themselves*)"<sup>15</sup>. Thus, this novel form of paternalism describes itself by "[steering] people's choices in directions that (...) will make choosers better off, as *judged by themselves*"<sup>16</sup>. To the paternalists, their proposed nudges don't coerce nor force behaviour and as such are in essence libertarian, so long as a person maintains the freedom to either ignore or reverse the nudge. We shall learn better in chapter 2, that libertarian paternalists plan to use psychological insights to influence behaviour, proposing to target people's automatic cues<sup>17</sup> and that nudge policies can be more intrusive, in the form of default rules, where a choice is

---

<sup>12</sup> Rebonato (2012)

<sup>13</sup> Sunstein (2014a)

<sup>14</sup> *Ibid.*

<sup>15</sup> *Ibid.* p. 17

<sup>16</sup> Sunstein and Thaler (2009) p. 5; emphasis in the original.

<sup>17</sup> Oliver, A. Nudging, Shoving and Budging: Behavioural Economic-Informed Policy. *Public Administration*, Vol. 94, No.4:700-714 (2015)

made for the individual, or they can be unobtrusive in the sense that they simply provide information.

Understanding the concept of nudging and some assumptions of Libertarian Paternalism, I shall use in this dissertation Rebonato's definition:

“Libertarian Paternalism is the set of interventions aimed at overcoming the unavoidable cognitive biases and decisional inadequacies of an individual by exploiting them in such a way as to influence her decisions (in an easy reversible manner) towards choices that she herself would make if she had at her disposal unlimited time and information, and the analytical abilities of a rational decision-maker (more precisely, of Homo-Economicus)”<sup>18</sup>

This statement does not mean that there are no limits to Libertarian Paternalism. At the same time it doesn't seem to exist a clear and concrete acceptability line in their literature<sup>19</sup>. According to Sunstein, paternalistic measures must submit to a cost-benefit analysis, in which the benefits of applying paternalistic policies must outweigh the costs. He goes further and implies that when considering paternalistic policies of this kind, public officials should account in their calculus for a perceived loss of autonomy by an individual<sup>20</sup>.

But, before we turn to analyse in detail Libertarian Paternalism, there are two categorical distinctions that try to explain the difference between this form of paternalism and other, less liberal, forms. I shall try and expose them briefly through the eyes of the paternalists.

---

<sup>18</sup> Rebonato (2012) p. 6

<sup>19</sup> *Ibid.*

<sup>20</sup> Sunstein (2014a)

*Means Paternalism vs. Ends Paternalism; Soft Paternalism vs. Hard Paternalism:*

*Means Paternalism* respects people's ends and only tries to change *means* employed. *Ends Paternalism*, disagrees with what people define as life objectives and tries to change these through more restrictive policies. This distinction can be very problematic, because it tends to generalize *means* and *ends* for everyone, when in actuality different people might take *means* to be *ends* and others might take *ends* to be *means*. Sunstein and Thaler position themselves as the means type paternalism<sup>21</sup>, but we shall see that in many domains their measures can be considered more akin to an ends-paternalism type<sup>22</sup>.

*Hard Paternalism* usually imposes itself through material or psychological costs on a given behaviour, it is usually achieved through fines, bans or criminal punishment, and it cannot be considered liberal in any kind. Hard paternalism is thus supposed to be more restrictive and coercive.

On the other end, *Soft Paternalism* tries to maintain freedom of choice. It does not ban, forbid, nor force penalties on behaviour. Instead, it tries to change behaviour through different methods, like education or the proposed use of "nudges", the latter which can take the form of default rules, graphic warnings, and provision of salient and edited information.

To further analyse the implications of Libertarian Paternalism, I propose to divide the dissertation in four different chapters. The first two, mean to contextualize the problem. Chapter 1 will present some of Mill's and Berlin's ideas on liberty: if, and when, there is legitimacy to infringe on liberty; what objections liberals tend to carry against paternalism; and, lastly, the libertarian paternalistic stance in relation to liberty.

---

<sup>21</sup> *Ibid.*

<sup>22</sup> Rebonato (2012)

In Chapter 2, I will address some of the findings that have been responsible to change perceptions of human rationality. This section will try to answer the question of what behavioural sciences brought about to justify paternalism? Finalising this chapter, the theoretical thesis of Libertarian Paternalism is introduced.

The following two chapters, 3 and 4, will be dedicated to examining, respectively, the practical and theoretical implications of Libertarian Paternalism on previously discussed understandings of liberty.

Chapter 3, will display practical applications of these paternalistic policy ideas, first proposed by nudge theorists and later by policymakers within the British government. These practical examples will serve to demonstrate how Libertarian Paternalism is not as non-coercive as its advocates argue, for the paternalists seem unable to foresee the unintended consequences of their policies over people's freedom to act<sup>23</sup>.

Lastly, Chapter 4 will challenge some assumptions of Libertarian Paternalism and discuss different implications for liberty by analysing how some of the initiatives put into practice can lead Libertarian Paternalism to hinder political liberty and decisional autonomy. In closing, I will briefly ponder over what Libertarian Paternalism really is, as much of their proposed policies go beyond the scope of paternalistic measures.

---

<sup>23</sup> Veetil, V. P. Libertarian Paternalism is an Oxymoron: an essay in defence of Liberty. *European Journal of Law and Economics*, Vol. 31, No. 3: 321-334 (2011)

## **CHAPTER 1:**

### **On Liberty and Decisional Autonomy**

In *Why Nudge: The Politics of Libertarian Paternalism*, Cass Sunstein presents what he calls the “Epistemic Argument of the Harm Principle”; he directly traces the latter back to John Stuart Mill’s ‘*On Liberty*’, presenting it as the fundamental reason to why people are so averse to paternalism. The Epistemic Argument for the harm principle, according to Sunstein, is the belief that each individual is much better equipped than any outsider, to act in their own self-interest and towards their own individual well-being. Under this Epistemic Argument there cannot be legitimacy for external influence on individual behaviour, especially if that behaviour does not harm others.

“If the goal is to ensure that people’s lives go well, Mill argues that the right solution is for public officials to allow people to find their own paths. Because individuals know their tastes and situations better than officials do, they are in the best position to identify their own ends and the best means of obtaining them.

Here, then, is an enduring argument, instrumental in character, on behalf of free markets and free choice in countless domains, including those in which people choose to run risks that may not turn out so well. Let us call this the Epistemic Argument for the Harm Principle”<sup>24</sup>

Even though Sunstein recognises other arguments in Mill’s defence for non-interference, he sees the Epistemic Argument as the most relevant one, and claims that “(...) behavioural findings are creating problems for the Epistemic Argument, because they show that people make a lot of mistakes, some of which can prove extremely damaging”<sup>25</sup>.

---

<sup>24</sup> Sunstein (2014a) p. 7

<sup>25</sup> *Ibid.* p. 8

In that same work of 2014, Sunstein presents Libertarian Paternalism in relation to questions of autonomy and the impracticality of what he call the “maximize choices” position, claiming that in many aspects of life people often “choose not to choose”<sup>26</sup>.

The purpose of this chapter is to analyse the Libertarian Paternalism position in relation to Mill’s ideas of liberty and Berlin’s “Two Concepts of Liberty”. But, as Sunstein dedicates some of his work to justify Libertarian Paternalism in relation to Mill, I propose before entering the discussion of liberty’s dualism presented in Berlin, to analyse briefly what John Stuart Mill, one of the forefathers of modern liberalism, had to say on the subject. It can only be fitting, for as Sunstein admits, Mill’s influence even today is still conspicuously present in most debates and works on the topic.

### **1.1 John Stuart Mill’s ‘On Liberty’**

John Stuart Mill was born in London in 1806, son to James Mill, an intellectual in his own right. This academic influence coupled with the tutelage from the likes of Jeremy Bentham, helped shape Mill’s mind and life, especially regarding his utility theory and his utilitarian ethical view.

He had, what one might consider, a strict childhood, starting to learn Latin and Greek at the mere age of three, violently overdeveloping his mind<sup>27</sup> at an early age. In his twenties he suffered a nervous breakdown, somewhat blamed on his emotion-controlling education, making him revisit all his beliefs and becoming less radical than he had been in his earlier thought.

Later, in 1859, he published *On Liberty*, which would become one of the most influential references to the liberal debate. This essay provided a view of human

---

<sup>26</sup> *Ibid.* see also Sunstein and Thaler (2009)

<sup>27</sup> Berlin, I. (2002c). John Stuart Mill and the Ends of Life (1959) in H. Hardy (Ed.) *Liberty: Incorporating Four Essays on Liberty*. Oxford University Press. p. 220

existence at both a social and political level, especially regarding *individual liberty*<sup>28</sup>. It differentiated between self-regarding actions and actions that affect other people. Its main objective, as stated by the author in his introductory remarks, was to assert a “simple” principle that should guide all dealings of society – the legitimacy that anyone had to interfere or restrict individual liberty – the *harm 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 from others (...) His own good, either physical or moral is not sufficient warrant. (...) Over himself, over his own body and mind, the individual is sovereign”<sup>29</sup>

For Mill, people should be absolutely free to question, speak, and act regarded that none of these actions hurt others. This approach directly benefits the well-being of each individual and consequently that of society. Liberty is then, a means to achieve happiness. However, Mill does not take the harm principle to be absolute. The possibility of harm to others does justify limitations on freedom, yet it does not mean that, even then, those limitations are always justified.<sup>30</sup>

Always asserting himself as a utilitarian, Mill believed that happiness was the ultimate goal for every individual, happiness which could only be attained through liberty under the base thought that the individual is “the primary locus of moral value”<sup>31</sup>. Quoting Wilhelm von Humboldt, whom he admired, he emphasized his belief and naturalistic view of humans:

---

<sup>28</sup> Marlin, R. (2013) *Propaganda and the Ethics of Persuasion* (2<sup>nd</sup> Ed). Peterborough: Broadview Press

<sup>29</sup> Mill, J. S. (1863). *On Liberty*. Boston: Tickner and Fields. p. 23

<sup>30</sup> *Ibid.*

<sup>31</sup> Wertheimer, A. (2002) *Liberty, Coercion and the Limits of the State* in R. L. Simon (Ed.) *The Blackwell Guide to Social and Political Philosophy*. Malden, Mass.; Oxford: Blackwell. p. 38.

“The end of man, or that which is prescribed by the eternal or immutable dictates of reason, and not suggested by vague and transient desires, is the highest and most harmonious development of his powers to a complete and consistent whole.”<sup>32</sup>

For him, happiness would only be complete when the self, through freedom, reached the full potential of realization – his ultimate goal. This is why individuality is one of the elements of well-being, highly important to the self-realization process, but also to the development of society. Individuality in Mill can only be reached in freedom and through diversity of experiences.

“It is not by wearing down into uniformity all that is individual in themselves, but by cultivating it and calling it forth, within the limits imposed by the rights and interests of others, that human beings become a noble and beautiful object of contemplation. (...) In proportion to the development of his individuality, each person becomes more valuable to himself, and is therefore capable of being more valuable to others”<sup>33</sup>.

The “eccentric” man does not conform himself to customs merely because he does not want to, but because he does not believe that what custom dictates is the right way for himself. He should be free, internally and externally, to choose and act in a way he understands to be right. But Mill goes further and claims that the unreasonable man, by acting differently to custom awakes in others the possibility to do so. Everyone is different from one another and unquestionable conformity only cripples society.

A person who adheres to custom just because it is customary, not even questioning themselves as to why they do it, is not in fact making any kind of choice, but only becoming mere imitators, automatons doing as they are told or programmed to

---

<sup>32</sup> Von Humboldt, W. *The Sphere and Duties of Government*. quoted in Mill (1863) p. 110

<sup>33</sup> Mill (1863) p. 121

do<sup>34</sup>. The real danger of modern times to the author is that, with the advent of industrialization and the masses, such conformity can stifle a man and contribute to a *collective mediocrity*. According to Berlin, Mill was acutely conscious of the:

“dehumanising effect of mass culture; of the destruction of genuine purposes, both individual and communal, by the treatment of men as irrational creatures to be deluded and manipulated by the media of mass advertising and mass communication – and so 'alienated' from the basic purposes of human beings by being left exposed to the play of the forces of nature interacting with human ignorance, vice, stupidity, tradition, and above all self-deception and institutional blindness”<sup>35</sup>

A man who acts and chooses with reference to his own internal motives and reasons, regardless if it is to abide to custom or against it, grows as a whole. This argument for individuality goes in the direction of decisional autonomy<sup>36</sup> where choices (actions) are made intelligently and motives are easily understood by the actor: human beings are those who make the important choices in their own lives. Mill accepts that people have natural forces external to them and sees those as part of people’s decisional autonomy. He does not mean autonomy in an absolute sense, but in the sense of direction towards self-realization and dignity<sup>37</sup>. With this reasoning we can consider Mill to “operate[s] with a “positive” notion of moral freedom as well as a “negative” notion of civil liberty”<sup>38</sup>.

However, the apparently overly-simplistic harm principle raises some difficulties that even ardent supporters of Mill find somewhat difficult to defend,

---

<sup>34</sup> *Ibid.*

<sup>35</sup> Berlin (2002c) p. 245

<sup>36</sup> Donner, W. (2010). *Autonomy, Tradition and the enforcement of Morality* in C. L. Ten (Ed.) *Mill’s on Liberty: A Critical Guide*. Cambridge: Cambridge University Press; see also Ginsburg and Wright (2012)

<sup>37</sup> Rebonato (2012)

<sup>38</sup> Skorupski, J. (1998). Introduction: The Fortunes of Liberal Naturalism in J. Skorupski (Ed.) *The Cambridge Companion to Mill*. Cambridge: Cambridge University Press. p. 4

especially when considering what harm might actually entail. Do we intend to state that only genuine physical harm falls under this category? Or at the same time, can we define “offence” as harm?<sup>39</sup> Under this light, critics claim that everything one does has the potential to be harmful to others as we live in society and not isolated<sup>40</sup>. Because Mill does not linger in working definitions for what might, or might not, mean harm to others, the principle is left wide open to interpretation. Notwithstanding, throughout his essay he introduces arguments that might help us draw a better picture of what he actually meant.

Harm is when someone’s actions cease to be self-regarding and injure other people’s interests, understood here as rights. So where is the definitional difficulty? Mill admits that a person is not an isolated being, and therefore some actions, even when done without intention of harm to others, may still injure someone’s rights.

“Whenever, in short, there is a definite damage, or a definite risk of damage, either to an individual or to the public, the case is taken out of the province of liberty, and placed in that of morality or law”<sup>41</sup>.

This becomes especially difficult regarding moral obligations towards others, which the author also considers. For Mill there are moral obligations, individuals should be accountable to themselves and act accordingly, not acting merely on a whim

---

<sup>39</sup> In Wertheimer (2002), we are briefly introduced to Joel Feinberg’s *Treatises on the Moral Limits of the Criminal Law*. According to Wertheimer, Feinberg presents three principles besides the harm principle that can be used to justify limitations on liberty; one of which is the offence principle – the state is justified in limiting someone’s liberty to prevent them from offending others – I do not believe Mill takes offence to be harm, especially because to him an opinion is of much higher value to the individual who holds it, than the feeling of offence against that same opinion (Mill, 1863). Feinberg’s two other principles are legal paternalism and legal moralism. In his essay Wertheimer actually goes further and presents other three principles.

<sup>40</sup> Berlin, I. (2002b). Two Concepts of Liberty in H. Hardy (Ed.) *Liberty: Incorporating Four Essays on Liberty*. Oxford: Oxford University Press

<sup>41</sup> Mill (1863) p. 158

of self-interest<sup>42</sup> (this thought goes hand in hand with his high regard for human beings as moral and rational). An individual can only self-develop within freedom, determined by one's own morality and reason. Only then can he achieve true happiness, the ultimate goal.

Mill admits that individuals are fallible, he understands that they can act according to lower impulses which, in turn, can guide towards self-harm, but he also believes that everyone is capable of experiencing "higher pleasures" and, as such, revealed preferences do not necessarily align with people's own well-being<sup>43</sup>. Thus, he understands that people can make judgements over other people's decisions. However, contrary to paternalists, Mill does not conceive that there is justification for interference in individual life in reference to those personal judgements.

"Mill's case for the right to self-determination revolves around the right of competent adults to assess the meaning and value of their experiences for themselves (...) any attempt to enforce from outside a particular conception of the good life undermines these essential liberal interests"<sup>44</sup>

Accounting for these different characteristics of liberty in Mill, we shall now see the three main reasons why the author finds it essential for the state to be left out of the private scope of each person's life.

First, Mill believed that individuals are better suited to act in their own self-interest than any external party, for individuals tend to act in a self-regarding manner, towards their own happiness. Secondly, and in response to the first reason, Mill argues

---

<sup>42</sup> In Mill (1863), Chapter V. He addresses the case in which people have moral obligations towards others, especially in cases of contracts, where expectations and calculations have been made. These contracts can and should be free to be overruled, but the individual is also bounded by moral obligations, especially if in ceasing the contract it directly affects injuriously third parties.

<sup>43</sup> Sugden, R. (2006). What we Desire, What We Have Reason to Desire: Mill and Sen on the Value of Opportunity, *Utilitas* Vol. 18, No. 1

<sup>44</sup> Donner (2010) p. 142

that even when people are not the most able to do some task, whatever that task may be, the state should still allow them freedom to do it, given it does not constitute harm to society. This will allow individuals to develop and learn on their own, and therefore extend their capacities to become self-determined and reach their ultimate goal.

The third reason, but nonetheless relevant, is that interfering in areas that should not fall under the state's domain only contributes to increasing the state's size. By becoming bigger, there's not only an increase to the responsibility of state affairs, but at the same time the risk of making individuals much more dependent on the state, cramping their space for development and leading to a uniform society.

It is hard to pin down the full extent of Mill's stance on liberty, especially considering that his views seem to be somewhat paradoxical within his own thoughts. He presents a principle that should guide all dealings of society but finds it hard to make his views absolute, leaving space for endless interpretations on legitimacy; he claims himself as utilitarian but asserts that utility can only be reached through individual freedom; he is a pluralist, admiring diversity in people, fearful of the role of custom in their individuality<sup>45</sup>. These are just some of the many difficulties present when interpreting Mill, yet we should not let these deter us or lead us to a dismissal of his work. As a fundamental element in liberal theories, Mill's work is indispensable.

“He broke with the pseudo-scientific model, inherited from the classical world and the age of reason, of a determined human nature, (...) with the same unaltering needs, emotions, motives, responding differently only to differences of situation and stimulus, or evolving according to some unaltering pattern. For this he substituted (not altogether consciously) the image of man as creative, incapable of self-completion, and therefore never wholly predictable: fallible; (...) a free, imperfect being, capable of determining his own destiny in circumstances

---

<sup>45</sup> I followed, some of these seemingly opposing positions in Mill in Ryan, A. (1998). Mill in a Liberal Landscape in J. Skorupski (Ed.). *The Cambridge Companion to Mill*. Cambridge: Cambridge University Press.

favourable to the development of his reason and his gifts. (...) He believed that it is neither rational thought, nor domination over nature, but freedom to choose and to experiment that distinguishes men from the rest of nature”<sup>46</sup>

## **1.2 Isaiah Berlin’s ‘Two Concepts of Liberty’**

In the introduction to his essay ‘*Two Concepts of Liberty*’ Isaiah Berlin presented us, in 1958, with a dualistic idea of liberty, where he worries about the power of philosophical ideas and the influence they might have within the political scope. Here he is recognising that ideas when turned into ideologies can have the power to destroy societies.

He contends that, considering the value of liberty, two coexisting and opposing ideas have conflicted with one another for the past centuries, influencing different movements within society. Both these concepts regard liberty from different departing points and answer different questions. One form is unforgivingly rationalistic<sup>47</sup>, looking at action restrained from within. The other form looks at men as free to develop within their own rights, where the state should only intervene in those areas of society where non-interference would result in chaos.

### ***Negative Liberty***

This concept regards liberty from a negative standpoint and it relates to the modern liberty of Constant and Mill meaning *freedom from* outside intervention.

“I am normally said to be free to the degree to which no man or body of men interferes with my activity. Political liberty in this sense is simply the area within which a man can act unobstructed by others. If I am prevented by others from doing what I could otherwise do, I am to that degree unfree; and if this area is

---

<sup>46</sup> Berlin (2002c) pp. 251-252

<sup>47</sup> Rebonato (2012)

contracted by other men beyond a certain minimum, I can be described as being coerced (...)<sup>48</sup>

Berlin does not dismiss the existence of internal constraints to our actions, but he only considers coercion to exist when there is a deliberate outside intervention that restrains action. The difference lies in the cause of our incapacity to act; internal constraints, originating in our own limitations, do not imply less political freedom just less capacity to act. But if those constraints come from others, if it is an external interference that “frustrates my wishes”, then I am coerced and therefore less free to act<sup>49</sup>.

Within this negative concept, freedom is not understood as an absolute value but rather as a fundamental part of political life. Other important values must coexist with liberty, but there ought not be a “confusion of values”. Favouring justice or security to the detriment of liberty does not make life freer but rather more just or more secure.

Coercion, even if benevolent, even if justified in the name of freedom, is not freedom. And as individuals live in society and do not conduct life isolated from one another, the state must impose some limits in order to avoid total social chaos<sup>50</sup>. The government’s role becomes to protect individual rights, through laws, and to keep safe the:

“area of personal freedom that which must on no account be violated; for if it is overstepped, the individual will find himself in an area too narrow for even that minimum development of his natural faculties which alone makes it possible to pursue, and even to conceive, the various ends which men hold good or right or sacred.”<sup>51</sup>

---

<sup>48</sup> Berlin (2002b) p. 169.

<sup>49</sup> *Ibid.* He adds: “Helvétius made this point very clearly: ‘The free man is the man who is not in rios, not imprisoned in a gaol, nor terrorised like a slave by the fear of punishment.’ It is not lack of freedom not to fly like an eagle or swim like a whale” p. 169

<sup>50</sup> *Ibid.*

<sup>51</sup> *Ibid.* p. 171

As such, negative liberty is freedom from private or institutional coercion, to protect and ward off interference, even if benevolent, as it only restricts freedom for the development of individuals, consequently disrespecting their individual modes of life. This form of liberty does not reply to the question of “who governs me?”.

For Charles Taylor, it is the fear of state control that enables a limited, and untenable, negative concept of liberty which is separated from the truths of modern psychology<sup>52</sup>. For Taylor, those who support a negative concept of liberty succumb too easily to the attractive, yet unrealistic notion, that freedom can mean the mere lack of interference in one’s actions. They deny the inner-self as relevant to men’s liberty. According to Taylor, negative liberty becomes too quantitative and he argues that this cannot be so. Individuals are as much bounded by external obstacles as by inner restrictions (feelings, thoughts, etc), these that present obstacles to act according to one’s own purpose. To Taylor, liberty cannot be conveyed whilst ignoring these inner restrictions, regardless if those are brute or import-attributing feelings.

As such, Taylor argues that the concept of negative freedom needs to be restructured. Freedom is no longer being capable to act without obstacles, but to act authentic in regard to one’s true-self, a person must respond to “judgements of significance”<sup>53</sup> and understand what is his/hers true purpose.

“freedom now involves my being able to recognize adequately my more important purposes, and my being able to overcome or at least neutralize my motivational fetters, as well as my being free of external obstacles”<sup>54</sup>

---

<sup>52</sup> Taylor, C. (1985) What is Wrong with Negative Liberty in *Philosophical Papers Vol. II, Philosophy and the Human Sciences*. (pp. 211-229). Cambridge: Cambridge University Press

<sup>53</sup> *Ibid.* p. 219

<sup>54</sup> *Ibid.* p. 228

However, Berlin doesn't conceive a quantitative version of negative liberty<sup>55</sup>, as Taylor argues. Berlin rather understands that there might be "many incommensurable kinds and degrees of freedom, and that they cannot be drawn up on in a single scale of magnitude"<sup>56</sup>. Furthermore, he recognises individuals as dynamic beings and agrees with Taylor when he states that people are dependent of inner judgements and values which are framed by their social context<sup>57</sup>.

As liberty here is considered a fundamental, but not exclusive, political right, government measures that impose on individual freedom (that should only be done in order to protect people's rights) will lead to a relationship of constant tension/monitorization between private life and government authority.

### ***Positive Freedom***

In opposition to the negative concept of freedom Berlin presents the idea of positive liberty, which responds to a concept of inner autonomy. Here, my decisions are my own in relation to my rational self:

"I wish above all, to be conscious of myself as a thinking, willing, active being, bearing responsibility for my choices and able to explain them by reference to my own ideas and purposes."<sup>58</sup>

Berlin understands that the two concepts might initially be perceived as very similar, but as they developed historically they became opposing conflicting notions of liberty, leading to very different movements. This divergence results in part to the fact that a positive concept of liberty focusses on intrinsic restrictions to freedom, presenting

---

<sup>55</sup> Castello-Branco, J. T. (2004). Isaiah Berlin: da Liberdade Negativa à Sociedade Decente in J. C. Espada and J. C. Rosas (Eds.) *Pensamento Político Contemporâneo: Uma Introdução* (pp.64-80). Chiado: Bertrand Editora

<sup>56</sup> Berlin (2002b) p. 177

<sup>57</sup> Castello-Branco (2004)

<sup>58</sup> Berlin (2002b) p. 178

an idea of two-selves within an individual. A higher-self, which acts according to reason, is constrained by a lower-self, and as such freedom responds only to rational action. “Freedom is not freedom to do what is irrational, or stupid, or wrong. To force empirical selves into the right pattern is no tyranny, but liberation.”<sup>59</sup>

Under this concept, one can argue that even when there is no external coercion over our actions (providing us with an incredible amount of negative liberty), one might still be restrained by inner obstacles, making us incapable of taking full advantage of our negative liberty. As such, Berlin regards positive liberty in terms of answering the question of “Who is master?”. Individuals want to be in control of their lives, they want to overcome their intrinsic restraints. Thus, one might talk of increasing liberty, in the positive sense, by overcoming less rational desires.

As such, under this light of positive liberty, to be free is to comply with self-control and self-determination. To act freely is to act according to one’s own reason. The belief in a universal truth, in the compatibility between all true answers in relation to reason, is for Berlin at the base thought of a positive notion of liberty<sup>60</sup>.

We should note that Berlin is not dismissing here the fact that those who defend a negative concept of liberty cannot be as pervasive as those who defend a positive notion of freedom. He is rather asserting that in modern times it is much more common and observable to use positive liberty as justification for coercive interventions in the name of freedom, i.e. to comply with individuals’ “true” liberties.

“Legal liberties are compatible with extremes of exploitation, brutality and injustice. The case for intervention, by the State or other effective agencies, to secure conditions for both positive, and at least a minimum degree of negative, liberty for individuals, is overwhelmingly strong. Liberals like Tocqueville and J.

---

<sup>59</sup> *Ibid.* p. 194

<sup>60</sup> Castello-Branco (2004)

S. Mill, and even Benjamin Constant (who prized negative liberty beyond any modern writer), were not unaware of this. The case for social legislation or planning, for the Welfare State and socialism, can be constructed with as much validity from consideration of the claims of negative liberty as from those of its positive brother (...). Each concept seems liable to perversion into the very vice which it was created to resist. But whereas liberal ultra-individualism could scarcely be said to be a rising force at present, the rhetoric of 'positive' liberty, at least in its distorted form, is in far greater evidence, and continues to play its historic role (in both capitalist and anti-capitalist societies) as a cloak for despotism in the name of a wider freedom."<sup>61</sup>

### 1.3 Liberty for the Libertarian Paternalists

With these liberal ideas in mind, Libertarian Paternalism finds space to define itself, it does not worry too much with a question of legitimacy but, like all types of paternalism, centres itself on a question of welfare and on the maximization of individuals well-being.

“Many of the concerns about paternalistic government focus on the idea of ‘legitimacy’, but in this context, at least, it is possible that the idea is a placeholder, or question begging, or perhaps even a mystification rather than a freestanding concept. (...) The real question is whether the action at issue makes peoples better or intrudes on their autonomy. (...) The question is not a bad start, but it should not be taken as rhetorical. On the contrary, it is less than helpful, because the abstract question of legitimacy mostly confuses matters. It diverts attention from what really matters: people’s lives.”<sup>62</sup>

To Sunstein intervening and creating soft paternalistic policies becomes legitimate because their goal is to improve people’s welfare. If the objective is to improve everyone’s own well-being, any type of interference (so long as it maintains

---

<sup>61</sup> Berlin, I. *Introduction* in Hardy, H. (ed.) (ed.) *Liberty: Incorporating Four Essays on Liberty*. Oxford University Press (2002a):3-54, pp. 38-39

<sup>62</sup> Sunstein (2014a) p. 88

freedom of choice and is not forcibly imposed on people) can be justified<sup>63</sup>. This is especially poignant, if we take into account that any interference of this kind is supported by empirical findings, which show patterns of behaviour. These measures are also only put into practice after a cost-benefit analyses over the net benefit of society<sup>64</sup>.

In short, this type of paternalism becomes justified by the mere fact that it improves people's welfare. In the same token, it becomes libertarian because it is not forced, allowing for alternative choice of action, and because it intends to lead us towards what the paternalists call our "true" preferences. Thus, Libertarian Paternalism does not worry itself with decisional autonomy or individualism, concepts which Mill believed necessary for the development of a good society and for the development of one's well-being<sup>65</sup>. More relevantly we shall see that this paternalistic movement, against liberal views mentioned above, does not fear the possibility of increasing government powers and it does not find necessary to draw specific limits because it implicitly trusts public officials to be benevolent, and to be guided by justifiability principles<sup>66</sup>.

Regarding their views on liberty, they seem to present themselves as a middle ground between the positive and negative concepts. Initially, they claim their initiative to be libertarian on grounds that they do not restrict or coerce behaviour, thus only presenting what they call soft measures of paternalism<sup>67</sup>. As they assert that their measures don't incapacitate alternative action, that people retain capacity of choice, they defend their libertarian position under a "negative" concept where freedom is the lack of external coercion. Further on I will show that, regarding some of their policies,

---

<sup>63</sup> Mitchell (2005); also in Ginsburg and Wright (2012); Rebonato (2012);

<sup>64</sup> Sunstein (2014a)

<sup>65</sup> Ginsburg and Wright (2012)

<sup>66</sup> Rebonato (2012); see also Sunstein and Thaler (2009); Sunstein (2014a)

<sup>67</sup> Sunstein and Thaler (2003a; 2003b; 2009); Sunstein (2014)

libertarian paternalists fail to differentiate between the mere preservation of choice and the actual process of decision-making. As some of their proposed policies decides for us, our “negative” freedom of action relies in voluntarily reversing the imposed intervention<sup>68</sup>.

On the other hand, libertarian paternalists justify their initiatives in similar fashion to a concept of positive liberty, sharing the belief that people are constrained by an empirically proven “lower-self” which is fallible, and proposing interventions that favour people’s more rational-self. They claim to interfere in accordance to what people “want to want”<sup>69</sup>. It’s with this in mind that Sunstein challenges Mill’s harm principle.

As noted in the introduction, Sunstein calls Mill’s objection to government interference - that people are better equipped than public officials to act in their own self-interest - the “Epistemic Argument of the Harm Principle”<sup>70</sup>. According to him, this argument is being challenged by different findings regarding human behaviour<sup>71</sup>. Sunstein uses this to justify Libertarian Paternalism, as this form of paternalism proposes soft measures that target *internalities* (harms that one does to oneself) it doesn’t infringe on the harm principle.

Despite sharing this two-selves picture, they cannot be considered as acting within a “positive” notion of liberty because their interventions are not intended to make individuals freer to act on their own, i.e. in favour of their “higher-selves”. Libertarian paternalists do not propose to overcome individuals’ fallibilities in favour of freedom, but instead they propose to exploit those fallibilities in favour of welfare.

---

<sup>68</sup> Rebonato (2012); also, Ginsburg and Wright (2012)

<sup>69</sup> Sunstein, C. R. (2016). *The World According to Star Wars*. HarperCollins

<sup>70</sup> Sunstein (2014a)

<sup>71</sup> *Ibid.*

“Their programme is essentially one of context manipulation, not one of fostering rationality”<sup>72</sup> and because they set their goal as to achieve individual’s well-being above everything else, they call themselves paternalistic. For Sunstein, decisional autonomy as an intrinsic process of development is, in itself, an apparent inner moral heuristic.

“Let me venture a stronger, more direct and perhaps reckless response to those who invoke autonomy. On one view, what really does and should matter is welfare, for which claims about autonomy are best understood as heuristic (at least in the areas under discussion here). More precisely, autonomy is what matters to System 1, but on reflection, the real concern, vindicated by System 2, is welfare. People speak in terms of autonomy, but what they’re doing is making a rapid, intuitive judgement about welfare”<sup>73</sup>

Under this light, it seems plausible to ask a few questions: Is welfare enough of a legitimate reason for paternalism to hinder our freedom in certain areas of our life? Can our decisional autonomy really be affected by these “harmless” *nudges*, which intend to use our “irrationalities” to comply with government’s ideas of what is our own well-being? Most importantly, does Libertarian Paternalism actually respect liberal understandings? Can Libertarian Paternalism be considered truly paternalistic? It is my intention to try and answer these questions by analysing various assumptions of Libertarian Paternalism and consequent implications for individual freedom. But first I propose to contextualize and present Libertarian Paternalism under theories of behavioural sciences and, more specifically, of behavioural economics.

---

<sup>72</sup> Rebonato (2012) p. 195

<sup>73</sup> Sunstein (2014a) p. 134



## ***CHAPTER 2:***

### **Behavioural Sciences and the Origins of Behavioural Public Policy**

#### **2.1 The Emergence of Behavioural Sciences**

Towards the end of the 19<sup>th</sup> century, psychology became the science inspired to study human nature. With the development of new technology and methodologies that promised to comprehend the inner workings of man, human behaviour and processes of decision-making have become more accessible to scientists.

Social psychology in these first instances was still very speculative, but with the horrors of the First World War an idea of men's irrationality became easier to conceive, leading to the realization of the powerful influence that emotion and fear play in people's decisions. Both unconscious and subconscious thought started being empirically tested leading to the belief that people's decisions can be guided by obscure intentions. Realizing the role of social influence, imagery, metaphors, framing, etc. on people's decisions warranted the impossibility of neutrally presenting choices.

With the advancement of technology; with the implementation of laboratory and field research; and with a more thorough methodology of study on the whole, psychology became one of the most important fields to understand human nature. Towards the second half of the 20<sup>th</sup> century many new emerging cognitive, social and behavioural findings would come to have great influence in different academic fields.

It was psychology, through scientific methods, that originally studied the impact of internal and external characteristics over our actions, but it wasn't psychology alone that influenced the thesis of Libertarian Paternalism. The justification for paternalism in this instance relied mostly on evidence from a field developed much later, a field put forth by many psychologists – that of behavioural economics.

For the past four decades or so, behavioural economics has become a specialized field within broader economical subjects. It analyses real behaviour in situations of risk and uncertainty, it assesses shifts in preferences, it observes behaviour throughout time in order to find patterns and causes regarding human decisions. In short, it's a scientific field dedicated to empirically studying decision-making processes within economics.

Behavioural economics is also the field that has postulated that people are constantly side-lined by cognitive “inefficiencies”. Behavioural economists, departing from psychology, concluded that people tend to overweigh irrelevant information; that time is inconsistent; that we are prone to inertia and procrastination; that memory and saliency play tricks on us and so forth. These inconsistencies and many others have been empirically tested and are presumed to lead towards systematic errors in judgement<sup>74</sup>.

These findings play a big part in justifying Libertarian Paternalism, and as such a contextualization of what has been discussed in the realm of behavioural economics seems to be in order. Bearing in mind that these are over-simplifications of much broader subjects, I hope it will be sufficient to understand some of the ideas that have come to influence libertarian paternalists.

### ***2.1.1 Assumptions of Rationality in Economics***

In the 18<sup>th</sup> century, Daniel Bernoulli devised an economic theory under assumptions of human rational choice. Even today, expected utility theory<sup>75</sup> is possibly the most influential model of rational choice in economics. Bernoulli asserted that rational beings would always try to maximise their expected utility, i.e. a subjective

---

<sup>74</sup> Sunstein and Thaler (2009)

<sup>75</sup> Expected utility theory here is presented as in Kahneman, D. (2012). *Thinking Fast and Slow*. London: Penguin Books; Oliver, A. (2017). *The Origins of Behavioural Public Policy*. New York, Cambridge University Press; and Baddeley, M. (2017). *Behavioural Economics: A Very Short Introduction*. Oxford: Oxford University Press

value of their own – understood by economists as happiness, well-being or satisfaction<sup>76</sup>. He assumed that people’s wealth was what made them more or less happy and as such, utility rose in a decreasing manner (declining marginal utility). Further in time, economics tried to devise ways of measuring cardinal utilities, i.e. people’s strength of preferences which reflect the value people place on a given good. Welfare economics would develop in a way to maximise society’s utility.

In the beginning of the twentieth century, with Alfred Marshall and the rise of neoclassical economics, more mathematical rigour was introduced to theories of utility maximization. By this point, an ordinal measure of value dominated welfare, indicators of the economical organisation of society were based on improvements and deteriorations<sup>77</sup>.

Towards the 1950’s, mathematics and logic become central to the development of neoclassical economics, where the “the quest increasingly focussed upon developing neat models of internal consistency that were divorced from the human experience”<sup>78</sup>. With the implementation of mathematical and logical rigour to economics, a more specified “neo-Bernoullian” theory of subjective expected utility rose.

Under descriptive predictions, expected utility theory assumes unquestionable rationality in the process of decision-making. Man acts as best as possible through the use of logic, reason, and a detailed analysis over all relevant information<sup>79</sup>. Choices are made in way of achieving one’s preferences, which tend to be stagnant, unchangeable, and consistent through time. In this manner, people act according to well-formed beliefs even under situations of risk and uncertainty.

---

<sup>76</sup> Baddeley (2017)

<sup>77</sup> Oliver (2017)

<sup>78</sup> *Ibid.* p.6

<sup>79</sup> Baddeley (2017)

Individual rationality has become the main foundation of mainstream economics, and expected utility theory its most dominant model. However, as some argue, normative accounts of rational choice should not be confused with descriptive accounts<sup>80</sup>. Under a normative light, expected utility theory means to show how humans ought to act and decide if people are indeed utility maximisers and fully rational. But, as these rational models yield observably inaccurate predictions, its logical assumptions must be put into question under a descriptive light.

“Although people ought to adhere to the axioms of rationality if they are expected utility maximisers, they often do not do so”<sup>81</sup>. To surmise, when human beings behave rationally one can predict their behaviour, by contrast when they become irrational by rational model standards, they become unpredictable.

Not dismissing any normative value to expected utility, especially in market-based transactions, we cannot consider it a descriptive theory, that is a theory that intends to show how human beings *do* decide<sup>82</sup>. As such, initial developments in the field of economics during the 1950’s started to challenge this rational choice model by trying to explain what was considered irrational behaviour.

### ***2.1.2 Bounded Rationality***

Even though some scientists in the 1950’s were challenging rational choice theory (a smart maximization of self-interest<sup>83</sup>) by empirically testing its formal axioms, others were questioning mainstream economics from other departing views. Herbert A. Simon, a psychologist and computer scientist, dedicated his life’s work to building a

---

<sup>80</sup> *Ibid.* also Oliver (2017); Kahneman (2012); Sunstein, C. (1997). *A Behavioural Analysis of Law*. University of Chicago Law Review; Coase-Sandor Institute for Law & Economics Working Paper No. 46.

<sup>81</sup> Oliver (2017) p. 15

<sup>82</sup> *Ibid.*

<sup>83</sup> Sen A. K. *The Idea of Justice*; quoted in Rebonato (2012) p. 54

theory of human behaviour. By the 50's he broke away from the traditional mainstream economic stance and started to introduce his highly influential concept of *bounded rationality*, which went directly against the neoclassical theory of *global rationality*.<sup>84</sup>

For Simon the problem arose from an evident detachment of economic theory from reality. Until then, economic models failed in conceiving a realist and practical view of human beings, and in his seminal paper of 1955 - *A Behavioural Model of Rational Choice* – he refuted the neoclassical position, proposing that the theory of the “economic man” needed revision. Economic models seemed to fail specifically due to their *lack of realism*<sup>85</sup> and so he set his task to:

“(...) replace the global rationality of economic man with a kind of rational behaviour that is compatible with the access to information and computational capacities that are actually possessed by organisms, including man, in the kinds of environment in which such organisms exist.”<sup>86</sup>

What Simon here calls environment is assumed to exist, in part, within the decision-maker and not just outside his biology. It is external and internal. Regarding decision-making we are bounded by our psychological capacities – inner constraints – and by the external contexts where we make our decisions – external constraints. To Simon these external behavioural variables will still be limited by our physiology and psychology. It is from this viewpoint that he moves to theorise that individuals’ inner constraints don’t allow for a full rationality as understood by neoclassic economic standards. Lack of knowledge, of information, of access, of time, etc. makes it difficult

---

<sup>84</sup> Barros, G. (2010). Herbert A. Simon and the concept of rationality: Boundaries and procedures. *Brazilian Journal of Political Economy*, Vol. 30, No. 3: 455-472.

<sup>85</sup> *Ibid.*

<sup>86</sup> Simon, H. A. (1955). Behavioral model of rational choice. *The Quarterly Journal of Economics*, Vol. 69, No. 1: 99-118. p. 99

to fully optimize any type of decision, leading individuals to “*satisfice*”, that is, to compromise and decide on something that is good enough<sup>87</sup>.

Bounded rationality assumes that individuals have goals (preferences) and at the same time constraints<sup>88</sup>. It is here that a divergence takes place from mainstream economics and a different path is set where behavioural economics<sup>89</sup> starts to develop as an alternative to neoclassical theory:

“Global rationality, the rationality of neoclassical theory, assumes that the decision maker has a comprehensive, consistent utility function, knows all the alternatives that are available for choice, can compute the expected value of utility associated to each alternative, and chooses the alternative that maximizes expected utility. Bounded rationality, a rationality that is consistent with our knowledge of actual human choice behaviour, assumes that the decision maker must search for alternatives, has egregiously incomplete and inaccurate knowledge about the consequences of actions, and chooses actions that are expected to be satisfactory (attain target while satisfactory constraints)”<sup>90</sup>

Bounded rationality here stands as a negative concept in relation to global rationality and so, it serves to illustrate and criticize the neoclassical position. At the same time, it’s perhaps too broad a term, oversimplifying what’s wrong with the former tradition without fully proposing alternative options. Some might even argue that it’s this characteristic of low specificity that explains the popularity of the concept.<sup>91</sup>

---

<sup>87</sup> Barros (2010); see also Oliver (2017) and Baddeley (2017)

<sup>88</sup> Simon, H. A. (2009). *An Empirically Based Microeconomics*. New York: Cambridge University Press.

<sup>89</sup> There are two schools of thought within behavioural economics that depart from much of the same evidence but diverge in their conclusions. The concept of *satisficing* exists merely in the bounded rationality school, asserting that despite our limits in rationality we can still make ‘good’ choices. The other school believes that these limits are what lead to systematic errors in behaviour and consequently to less than optimal results. We shall see below that the distinction seems to go amiss in the libertarian paternalist literature.

<sup>90</sup> Simon (2009) p. 17

<sup>91</sup> Barros (2010)

Regardless, the concept has given way to an alternative train of thought that behavioural economists have put forward, one where people are not assumed to be fully rational but with inherent limits to their decision-making processes. Behavioural economics “extends economic principles by allowing that our decisions are affected by social and psychological influences, as well as a rational calculation of benefits and costs”<sup>92</sup>.

Different theories of bounded rationality as alternatives to the orthodox view have been proposed: *Ecological Rationality*, *Practical Rationality*, *Selective Rationality*, all have been hypothesised as opposing/complementary traditional theories to the *Homo economicus*. As such, the general consensus over what bounded rationality actually is seems to be insufficient and possibly too diverse, nonetheless it is a fundamental part of contemporary economics.

### ***2.1.3 The Emergence of Behavioural Economics***

As noted, neoclassical economic theory seems to believe that people act in their own interest after a cost-benefit analysis over the information and evidence they have. Individuals are a reflection of their preferences and mostly rational.

Under this light, economic models have been designed to account for supply and demand, different market behaviours, speculation, prices, etc. Market failures seem to occur only due to lack of institutional support or other intrinsic factors, rarely because economic individuals are fallible. Behavioural economists defend that those assumptions depart from a mere theoretical standpoint, without proper practical or

---

<sup>92</sup> Baddeley (2017) p. 1

realistic views. They propose, then, to look additionally at the behaviour of individual decision makers who constitute a market.<sup>93</sup>

Recent cognitive and social psychological studies have presented a large body of evidence challenging standard economic assumptions. We have seen in the previous section how Simon looked at economic rationality and challenged it from a generalized point of view. As the concept of bounded rationality became too broad to explain market failures, or problems with rational behaviour, others tried to study specific failures of rationality through more observable and testable evidence, borrowing from psychological methods.

Daniel Kahneman and Amos Tversky, considered to be the prominent fathers of behavioural economics, set out in the 1970's to try and comprehend complexities in human decisions. They both had realised that their own answers to certain postulates inferred in systematic and predictable errors. Under this pretence they conceived experimental studies where they were to analyse reoccurring mistakes within individuals' decision-making processes, those which are caused by *heuristics* (rules of thumb) and *cognitive bias* (blind spots)<sup>94</sup>.

Before analysing some of Kahneman's and Tversky's specific work, let us discuss briefly Dual Process Theory as it is highly relevant to interpreting some of these findings.

#### ***2.1.4 Dual Process Theory***

Although the distinction between a fast/intuitive way of thinking and that of a deliberate/reflective one is present in early psychology, it is during the 1980's and 1990's that Dual Process Theory empirically tests the idea of a binary system of

---

<sup>93</sup> *Ibid.*

<sup>94</sup> Kahneman (2012)

thought. It proposes the existence of a System 1, which is automatic, almost intuitive and immediate. This System has throughout our lives learnt from different scenarios, applying rules of thumb to life's' different situations. It is not irrational, but due to System 1 people are led to make mistakes, wrong considerations and choose without due deliberation, which in turn contributes to less than optimal results<sup>95</sup>.

This so-called automatic System 1 triggers impulsive and fast reactions to any little push in our lives, it is not about irrationality, but adaptability to everyday life – a matter of evolution – and when we don't slow down and reconsider what it's suggesting, especially in the case of unfamiliar situations such as risk and uncertainty, mistakes tend to be made.

As System 1 can be considered the most child-like part of our brain (and we shall see it as the intended target of Libertarian Paternalism<sup>96</sup>) System 2 is its adult counterpart. On the one hand it is reflective and effortful, responsible for slowing down our processing brains and allowing for a better analysis of any given problem. On the other hand, it's responsible for monitoring and controlling actions and thoughts suggested by System 1, it's what allows for inner self-control. System 2 is thus expressed by thoughtful behaviour as well as for the suppression of thoughts and actions it doesn't deem proper. In some literature System 2 becomes Adam Smith's "impartial spectator"<sup>97</sup> whereas System 1 is more akin to Keynes's "animal spirits"<sup>98</sup>.

“[System 2] allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration”<sup>99</sup>

---

<sup>95</sup> Sunstein and Thaler (2003a; 2003b; 2009)

<sup>96</sup> For Rebonato (2012) it is this idea of Dual Process Theory that allows the libertarian paternalists to depict individuals with two separate selves.

<sup>97</sup> Oliver (2017)

<sup>98</sup> Baddeley (2017)

<sup>99</sup> Kahneman (2012) p. 21

There is some interaction between these two systems, System 1 is the first to analyse any situation and the first to respond. It works as a monitor and calls on System 2 only when it finds necessary, when unable to process the task at hand. Some argue that this is especially true because the reflective system is a limited resource and can be expended, a process that is called *ego-depletion*<sup>100</sup>.

Having briefly explained Dual Process Theory in our decision-making processes I shall turn now to Kahneman and Tversky's original work. However it should be noted beforehand that it was this theory of Dual Process and the extensive research done in behavioural economics that went on to influence advocates of Libertarian Paternalism.

### ***2.1.5 Behavioural Economists on Heuristics and Cognitive Bias***

When making decisions we are constantly subjected to inner heuristics and cognitive biases, as a result “errors of intuitive thought are often difficult to prevent”<sup>101</sup>. The following questions, which ask people to reply with the first answer that comes to mind (usually suggested by the automatic system) are part of a wide Cognitive Reflection Test done in hundreds of experiments.

---

A bat and a ball cost 1.10 €. The bat costs one euro more than the ball. How much does the ball cost?

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

---

<sup>100</sup> In Kahneman (2012); *Ego-Depletion* is a term originally proposed by Roy Baumeister whom, alongside his team, has implied that our mental efforts run on glucose. His research has led to the conclusion that glucose levels in the body are directly correlated with deterioration of thought performance (*ego-depletion*), that is, when glucose levels are high *ego-depletion* decreases; see also Garvey, J. (2016). *The Persuaders: The Hidden Forces that are Trying to Change Your Mind*. London: Icon Books.

<sup>101</sup> Kahneman (2012) p. 28

In a lake there is a patch of lily pads. Every day the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half the lake?

---

More than half the respondents to these questions, usually done to university students in academic studies, answer wrongly<sup>102</sup>. Obviously, these questions have been purposely chosen because they tend to induce in error by targeting our more automatic brain cells. It is only when the reflective system comes to aid, that people rethink their intuitive answers and understand that a ball costs 5 cents, that it takes 5 minutes for 100 machines to make 100 widgets, and that it takes 47 days for the patch of lily pads to cover half the lake.

In an article published in 1974 entitled *Judgement under Uncertainty: heuristics and biases*, Kahneman and Tversky argued that we assess probabilities as we do physical quantities, such as distance or size. We use rules that System 1<sup>103</sup> suggests in order to move away from unnecessary complexities. These rules are valid for most cases, but in unfamiliar situations, as in cases of risk and uncertainty, they often lead to errors. The authors were starting to demonstrate failures that occur systematically in human judgement. They hypothesized that these judgements are highly influenced by the role of belief and they concluded that:

“people rely in a number of heuristic principles which reduced the complex tasks of assessing probabilities and predicting values to simpler judgemental operations. In general, these heuristics are quite useful, but sometimes they lead to sever and systematic errors”<sup>104</sup>.

---

<sup>102</sup> *Ibid.* see also Garvey (2016)

<sup>103</sup> Even though Evan, Stanovich and West developed Dual Process Theory in the 1980's/90's; Kahneman (2012) recognises and assumes this theory as complementary to his and Tversky's work.

<sup>104</sup> Kahneman, D., Tversky, A. (1974). *Judgements under Uncertainty: Heuristics and Biases*. Science, Vol. 185 in Kahneman (2012: Appendix A) p. 419.

Kahneman and Tversky's research program has been responsible for finding many of these systematic errors, connecting them with System 1 of Dual Process Theory. Other behavioural scientists expanded their programme and found evidence of many more biases. In turn, the study and discovery of these factors of behaviour have deeply influenced Sunstein and Thaler in their proposal of Libertarian Paternalism, and as such it seems relevant to touch on some of these findings to later understand many of libertarian paternalists' assumptions and how they are proposing to use insights from behavioural economics to put Libertarian Paternalism into practice.

#### ***A. Representativeness, Availability and Anchoring***

In that same article mentioned above, Kahneman and Tversky identified three heuristics in the assessment of probabilities and statistics. The *representativeness heuristic* is the way in which people tend to adjust their judgements to stereotypical ideas<sup>105</sup>. Consider the two following descriptions, knowing that both have been present in studies over probabilities.

---

Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.

Linda is a thirty-one years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and also participated in antinuclear demonstrations.

---

---

<sup>105</sup> *Ibid.*

When presented with these questions people are asked: what is Steve more likely to be? A farmer, a salesman, a pilot, a librarian or a doctor? What about Linda, is she more likely to be a bank teller, an insurance salesperson, or a “bank teller and active in the feminist movement”?

Statistically speaking Steve is much more likely to be a salesman or a farmer as there are much more people in those professions than in any other options presented. Yet, respondents often answer that Steve is probably a librarian. More surprisingly, a majority of respondents believe that “bank teller and active in the feminist movement” is probably the right answer in relation to Linda, because apparently both descriptions are coherent with our stereotypes of a librarian and then the feminist movement.

Kahneman and Tversky conclude that to analyse probability people often consider subjective descriptions as the most representative input. According to Kahneman when regarding the Linda conundrum, we are primed (respond to subconscious cues) by irrelevant information to commit the *conjunction fallacy* - to judge a conjunction of two events more likely than just one of those same events<sup>106</sup>.

In their research the authors continued to analyse different choices under situations of uncertainty, realising that people tend to judge frequency according to the examples that come most easily to our mind. They called this rule of automatic response the *availability heuristic* and assumed that different factors contribute to it.

When asked about frequency, individuals automatically access their subjective memory, which is highly influenced by saliency, familiarity and time. Therefore, the picture we draw of the world is highly dependent on the prevalence and emotional intensity of messages and information that can be provided by personal experience or

---

<sup>106</sup> Kahneman (2012)

external sources<sup>107</sup>. Availability can be the cause that makes us overlook or overestimate certain dangers, the frequency of an event or even different beliefs we hold about reality.

Upon demonstrating *the Anchoring Effect* both these authors concluded that when you have a number in your mind, regardless of where it comes from, it will always influence and guide a possible estimate. There are two different types of anchors, on one hand they are involved in a process of adjustment – you guide yourself by accessing information you already have. On the other hand, anchors can be of a priming nature<sup>108</sup> - acting as mere suggestions, thus becoming impossible to avoid.

What research has shown is that no matter the source or the information you have, it seems to be impossible to stray very far from the initial anchor, we are easily guided by random and non-random numbers which challenge our assumptions about value, speculation, prices and about the power of influence in negotiation<sup>109</sup>. Consequently, we are highly dependent on the context and in how numbers are presented to us<sup>110</sup>.

### ***B. Loss Aversion and Framing***

Human beings have genuine *loss aversion*, i.e. the possibility of losing something is felt much more ardently than the possibility of winning, in comparison, that same particular thing. “What this means is that people do not assign specific values

---

<sup>107</sup> *Ibid.*; see also Garvey (2016); Sunstein and Thaler (2008): Regarding the availability heuristics these authors provide the example of terrorism since 9/11/2001, which in the west has been seen like a constant threat. Nations spend time, effort and public resources to tackle this issue, which results in that people tend to overestimate the frequency of terrorism on basis of memory, recurrence and salience when it so happens that even in Israel people die less of terrorism than of heart-attack.

<sup>108</sup> Kahneman and Tversky (1974) present a study where two groups observed two rigged games of a roulette wheel, both groups were then asked to guess the percentage of African nations in the UN. Results showed that the roulette wheel experiment actually influenced responses, despite everyone knowing that such a random number is irrelevant in such assessment.

<sup>109</sup> Garvey (2016)

<sup>110</sup> Kahneman (2012); see also Garvey (2016); Sunstein and Thaler (2009)

to objects. When they have to give something up, they are hurt more than they are pleased if they acquire the very same thing”<sup>111</sup>. Loss aversion explains the *endowment effect* which observes that people place a much greater value on a particular good when they own it than they did before they owned it<sup>112</sup>.

Similarly, when it comes to framing, Kahneman and Tversky realized that people’s reactions and preferences are highly influenced by the way in which contexts and ideas are proposed to them. They tried to demonstrate how life is affected everyday by how information is presented, how language can have a deep impact on decisions, and how different words with virtually the same meaning can lead to different answers. “Preferences between the same objective outcomes reverse with different formulations”<sup>113</sup>.

The idea of *framing* is not exclusive to economic realities; in media, advertisement and politics, frames are of the utmost importance and have a deep impact on public opinion<sup>114</sup>.

By studying the power of frames experts can realise how to propose ideas or policies in a better light, resulting in possible better results<sup>115</sup>. Libertarian paternalists, for example, call attention to this and recognise the use of framing as possibly perverse, in the sense that it can be used to shift general opinion in certain directions. At the same time these paternalists admit that frames can be used in a manner to “nudge” towards

---

<sup>111</sup> Sunstein and Thaler (2009) p. 36

<sup>112</sup> Oliver (2017)

<sup>113</sup> Kahneman (2012) p. 368

<sup>114</sup> Garvey (2016)

<sup>115</sup> The power of framing metaphors goes well beyond behavioural economics theory. For a brief exposition on the matter see Garvey (2016), Chapter 5. See Thibodeau and Boroditsky (2011) regarding framing within policy-design. These authors concluded that when asked to tackle certain policy issues, responses shift depending on how the issue is framed. For a brief exposition on framing and language within politics see Bai, M. (17-7-2005) “The Framing Wars” *The New York Times*.

individuals' own well-being by, for example, providing framed information that targets our System 1<sup>116</sup>.

Notwithstanding, it was concepts of loss aversion and framing that allowed Kahneman and Tversky to directly challenge economic rational choice models. Their research proposed preference inconsistencies, risk-averse and risk-seeking characteristics within the same individual, plus a decreasing utility associated, not to amounts of wealth, but to strong emotional responses regarding losses.

### ***C. Status-quo and Present Biases***

Two highly relevant cognitive biases in the behavioural literature that directly influence Libertarian Paternalism in its policy application is the *status quo bias* and *present bias*. In reference to the first, it is implied that people generally tend to stay with their current situation no matter what. This means that people tend to suffer from *inertia* as they do not want to change their *status quo* in favour of an uncertain future. As we shall see further on, it is this idea of inertia that will allow libertarian paternalists to propose interventions that implicitly design a *status quo* as the “default choice”.

In this same regard, *present bias* or what others call, *hyperbolic discounting* shows that people attach a particularly heavy weight to the present time, valuing the immediate moment to the detriment of the future. Individuals tend to show dynamic inconsistency throughout time and consequently decisions in the present moment are not aligned with people's more deliberative preferences, i.e. not aligned with that which will give them the most utility<sup>117</sup>.

---

<sup>116</sup> Sunstein and Thaler (2009)

<sup>117</sup> Oliver (2017)

This idea of hyperbolic discounting allows behavioural economists to divide the individual into two semi-autonomous selves<sup>118</sup> somewhat associated with Dual Process Theory. A present-self that responds to immediate moments imposing costs on a future-self with different, perhaps more rational, preferences<sup>119</sup>. The idea of two-selves in time, present in behavioural economics, leads to the assumption that we have negative *internalities* – harms that impose on one’s future-self. It is with this in mind that libertarian paternalists will highlight individuals’ tendencies for lack of self-control (*akrasia*<sup>120</sup>).

The exposed work above seems sufficient to demonstrate some of the underlying assumptions of behavioural economics and which, combined with Herbert Simon’s bounded rationality, allow to see individuals with inner constraints that directly impact decisions against their own well-being<sup>121</sup>. Dual Process Theory has also deeply impacted many behavioural fields, and under its wing Sunstein and Thaler distinguish two selves, the “adult” and the “child”, and assume that the real world is shared between *Humans* and *Econs*, between *Homo sapiens* and *Homo economicus* respectively. They argue that the majority are *Humans* (rationally bounded individuals that act in favour of their System 1 and therefore need intervention in many areas to be able to reach better outcomes).

“Those who reject paternalism often claim that human beings do a terrific job of making choices, if not terrific, certainly better than anyone else would do (especially if that someone else works for the government) (...). If you look at economic textbooks, you will learn that homo economicus like Albert Einstein, store as much memory as IBM’s Big Blue, and exercise the willpower of Mahatma Gandhi. (...) Real people have trouble with long division if they don’t have a

---

<sup>118</sup> Rebonato (2012)

<sup>119</sup> Sunstein and Thaler (2009)

<sup>120</sup> Sugden (2017)

<sup>121</sup> *Ibid.*; see also Sunstein and Thaler (2003a, 2003b); Sunstein (2014a)

calculator, sometimes forget their spouse's birthday, and have a hangover on New Year's Day. They are not homo economicus; they are homo sapiens. To keep our Latin usage to a minimum we will hereafter refer to these imaginary and real species as Econs and Humans"<sup>122</sup>

For libertarian paternalists intended purposes, System 2 (responsible for our rationality) is separate from System 1<sup>123</sup> (responsible for our lack of self-control, time inconsistency, and other alleged errors in rational behaviour). These paternalists assume it as their responsibility to create policies that engage our System 1 in order to act with *as if* rationality, i.e. in order to make us behave towards our subjective well-being<sup>124</sup>.

## 2.2 Behavioural Welfare Economics and the call for Paternalism

Paternalism has many forms and possible definitions, here it will be understood as libertarian paternalists define it: a policy that intends "to steer people's choices in directions (...) that will make choosers better off, *as judged by themselves*"<sup>125</sup>. The metaphor of paternalism comes from a parent who interferes with the choices of a child in order to promote that child's greater good, for they do not believe that their children know best or are acting accordingly. But as harder forms of paternalism demand and force people's choices, Libertarian Paternalism is proposing to guide and "steer" towards good behaviour. The emphasis on "*judged by themselves*" is highly relevant, because it shows the authors' belief in knowing and being able to attain, objectively, what people's "true" preferences are.

---

<sup>122</sup> Sunstein and Thaler (2009) p.7

<sup>123</sup> Rebonato (2012)

<sup>124</sup> *Ibid.*

<sup>125</sup> Sunstein and Thaler (2009) p. 5; emphasis in the original

Under this definition paternalism is driven by the *intention* to guide citizens<sup>126</sup>, who don't act in their best interest, towards their ultimate well-being. As such, it is not enough to judge people's individual choices to be paternalistic, but it is rather imperative to intervene in accordance to those judgements<sup>127</sup>.

In neoclassic economic theory, paternalism is not a very common viewpoint, usually associated with unnecessary regulations and bureaucracy that encroaches on free-markets, consumer sovereignty and self-interest. Notwithstanding, in 2003 two different papers<sup>128</sup>, relatively similar in content, were published by different academics in the fields of economic law and behavioural economics. Both proposed conceptual frameworks on how insights taken from behavioural sciences could and should be applied to policy. They justified themselves by presenting different findings from this new emerging behavioural field.

The two articles are similar in content in the sense that they both propose soft forms of paternalism, specifically targeting change in consumer/citizen behaviour. But also, because they take welfare as a normative objective and at the same time make modifications to modern welfare economics by seemingly detaching the achievement of well-being from individual choice<sup>129</sup>. These two separate ideas have been defined as proposals of a “new” paternalism, and they are both adamant that behavioural findings call for paternalism<sup>130</sup>.

---

<sup>126</sup> Sugden (2017)

<sup>127</sup> Sugden, R. (2008) Why Incoherent Choices Do Not Justify Paternalism. *Constitutional Political Economy*, Vol. 19: 226-248

<sup>128</sup> In 2003 three different papers on this matter were published, as two of these are companion articles written by Sunstein and Thaler (2003a; 2003b), I have considered both them as one in the text.

<sup>129</sup> Oliver, A. (2013b). Introduction in A. Oliver (Ed.) *Behavioural Public Policy*. Cambridge, UK; New York: Cambridge University Press. Also, in Oliver (2017).

<sup>130</sup> Glaeser, E. L. (2006). Psychology and Paternalism. *University of Chicago Law Review*, Vol. 73, No. 1: 133-156. I should also note that this call for paternalism justified by behavioural sciences, does not in any way, limit itself to soft types of paternalism. Other authors have found space to justify forms of hard paternalism under light of behavioural insights; see for example Conly, S. (2013). *Against Autonomy: Justifying Coercive Paternalism*. Cambridge: Cambridge University Press.

“Virtually every scholar who has written on the application of psychological research on judgment and choice to law has concluded that cognitive psychology supports institutional constraint on individual choice.”<sup>131</sup>

The paternalists argue “for self-conscious efforts, by private and public institutions, to steer people’s choices in directions that will improve their own welfare.”<sup>132</sup> Offering measures that improve behaviour on boundedly rational individuals, without ever forcing or coercing.

In Thaler and Sunstein their form of paternalism is said to be libertarian because it preserves freedom of choice, but as its’ main goal is to influence people’s choices in order to make them better off, it becomes descriptively paternalistic<sup>133</sup>.

In Camerer *et al.*, they term their approach as *Asymmetric Paternalism*. The authors admit this form of paternalism as a conservative notion on their part for fear of promoting, “at this stage”<sup>134</sup>, harder forms of paternalism. They become cautious and conservative under three assumptions: (1) They recognise that, even though there’s evidence showing the limits of human rationality, those limits are not universal. Some individuals can and do act according to standard concepts of rationality. (2) They admit that behavioural economics is a relatively new science and therefore its findings should be dealt with caution and should not be treated as unquestionable. (3) They share the common knowledge that people value freedom of choice and they fear that some of the proposed policies might be understood to encroach on people’s freedom<sup>135</sup>. In this sense

---

<sup>131</sup> Rachlinski, J. *The Uncertain Psychological Case for Paternalism*; quoted in Mitchell (2005) pp. 4-5

<sup>132</sup> Sunstein and Thaler (2003a) p. 1162

<sup>133</sup> *Ibid.*

<sup>134</sup> Camerer, C.; Issacharoff, S., Lowenstein, G, O’Donoghue, T., Rabin, M. (2003). Regulation for Conservatives: Behavioural Economics and the Case for “Asymmetric Paternalism”. *University of Pennsylvania Law Review*, Vol. 151: 1211-1254, p. 1214

<sup>135</sup> Interestingly, in Camerer *et al.* (2003) p. 1214 the authors request to further analyse, empirically, the perception of freedom of choice: “We feel that how people perceive limits on their free choice should

it differentiates from Thaler and Sunstein's work, who don't seem to look so carefully at the behavioural evidence, but also because Camerer *et al.* admittedly assume that the preservation of choice and soft paternalism is not enough to claim their proposal as liberal as it implies costs on those who are considered rational decision-makers and because in some regards their policies restrict voluntary actions.

Despite their differences, both these works seem to share the view that the debate about paternalism should shift focus from a legitimacy standpoint to an empirical one. If empirical evidence is showing that fully-rational beings are indeed rare in society and that individuals do not always act according to their own self-interest, thus challenging the foundation for free-markets and individual liberty. Then, they claim, the debate between the two positions (paternalist vs. anti-paternalist) becomes counterproductive<sup>136</sup>. Where one is discussing emerging empirical evidence, the other is arguing on bases of belief.

The new paternalists hope to reach a compromise between these two positions – the middle ground. Policies to be effective should contend emerging empirical evidence that show how human beings do act, yet those policies should not encroach on individual freedom of choice.

For their part Sunstein and Thaler assume that the problem with the anti-paternalistic stance is that it's based on one false assumption and two misconceptions. The false assumption is that people usually make choices in their own best-interest, or at least better choices than any external person might impose on them. The authors argue that this is a testable and measurable assumption and what behavioural evidence has shown, so far, is that it is also obviously false.<sup>137</sup> In claiming this, the key liberal

---

itself be subject to behavioural research, rather than be treated as an axiom of resistance in the exploration of paternalism”.

<sup>136</sup> Camerer *et al.* (2003), Sunstein (2014a)

<sup>137</sup> Sunstein and Thaler (2003a; 2003b)

assumption in Mill, the so called “Epistemic Argument of the harm principle” - individuals are the best judges and protectors of their own well-being – falls to ground.<sup>138</sup>

The two above mentioned misconceptions are: first, the belief that there are possible alternatives to paternalism and secondly that paternalism must always involve coercion. Against the first misconception, Sunstein and Thaler argue that there’s always some sort of paternalism at hand, it is inevitable: individual choices are constantly influenced by external intervention. This idea is somewhat of a conundrum, as the libertarian paternalists seem to be implying that anti-paternalists believe any influence over human behaviour to be paternalistic<sup>139</sup>. Under light of what we have seen before, where paternalism is defined by the intention to change behaviour this idea becomes somewhat of a moot point.

Regardless, in response to the second misconception – the belief that paternalism always involves coercion – Sunstein and Thaler argue that outside interference, which influences people’s decisions in favour of their well-being, can be done without necessarily blocking individual choice. It is under this belief that they name their paternalism as libertarian.

Having this in mind, the theory of Libertarian Paternalism seems to contend various suppositions: First, it refuses an idea of rationality in decision-making and therefore treats individuals for what they truly are: boundedly rational. In doing this, libertarian paternalists claim that what individuals choose on their own can be, not only harmful for society, but also, and mainly, harmful for themselves<sup>140</sup>.

---

<sup>138</sup> Sunstein (2014a)

<sup>139</sup> Veetil (2011)

<sup>140</sup> *Ibid.* see also Mitchell (2005); see also Rebonnato (2012); Sunstein and Thaler (2009)

Secondly, they assume, under pretence of behavioural sciences, that preferences are endogenous, not stable or well-define but rather context-dependent<sup>141</sup>. This means that preferences can be influenced by any “frame” in the setting and as a result of our bounded rationality and preference endogeneity, people’s choices are not reliable to promote one’s best interest unless “choice architects” nudge individuals to select the welfare-advancing option.<sup>142</sup>

Thirdly, because of this context-dependency, the power of influence over people’s behaviour by those who control the design of contexts is unavoidable, and as such, a form of paternalism becomes the only viable option; because what truly matters is welfare<sup>143</sup>. That is, if the benefits of Libertarian Paternalism are larger than the harms done, if after a cost-benefit analyses it is concluded that individuals’ well-being is improved<sup>144</sup>, then legitimacy lies on how to produce situations that promote people’s welfare. “[I]t is legitimate for choice architects to try to influence people’s behaviour in order to make their lives longer, healthier, and better”<sup>145</sup> because ultimately, Sunstein argues, this is what individuals really want. Here he defines welfare from the individuals’ point of view to mean: “whatever choosers would think make their lives go well”<sup>146</sup>

“Once it is understood that some organizational decisions are inevitable, 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, we can abandon the less interesting question of whether to be paternalistic or not, and turn to the more

---

<sup>141</sup> Sugden (2008; 2017; 2018)

<sup>142</sup> *Ibid.*

<sup>143</sup> Sunstein (2014a)

<sup>144</sup> Even though it is not as evident in Sunstein and Thaler, in Camerer *et al.* (2003) they admit that “asymmetrically paternalistic policies are not likely to yield Pareto improvements, wherein everyone benefits” p. 1221

<sup>145</sup> Thaler and Sunstein (2009) p. 5

<sup>146</sup> Sunstein (2014a) p. 73

constructive question of how to choose among the possible choice-influencing options”<sup>147</sup>

Lastly, the role of Libertarian Paternalism becomes to figure out, through the input of behavioural sciences, when and where people tend to systematically err. A policymaker needs to understand how a person truly behaves so steps can be taken to prevent harmful behaviour. To Sunstein and Thaler the form in which government ought to help *humans* is by using what they call “nudges” and by principles of “choice architecture”, to which I now turn.

### 2.3 Libertarian Paternalism

Although both these articles were written around the same time and had fundamentally the same proposals in mind, though displaying some differences in their conceptual frameworks, it was only in 2008 with the publication of *Nudge: Improving Decisions about Health, Wealth and Happiness* that the thesis of an alternative “new” paternalism became widely known.

Within this book, they proposed several *nudging* interventions. Nudges being the fundamental application of Libertarian Paternalism, understood as tools of influence applied through private or public policies, targeting our less conscious reactions.

It should be noted that these nudges are not new tools of influence, industries such as communications, sales, marketing, etc. are experts in nudges.<sup>148</sup> The libertarian paternalists are merely acknowledging the power of these techniques in steering behaviour and are inviting governments to design policies that, without being coercive, try to prevent or ameliorate negative *internalities* – harms imposed on oneself in contrast to the concept of externalities. Nudges are supposed to be an interference that

---

<sup>147</sup> Sunstein and Thaler (2003a) p. 1166

<sup>148</sup> Garvey (2016)

“alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives”<sup>149</sup> by specifically targeting our System 1 of Dual Process Theory discussed before.

Whether the predictability of these changes in behaviour is certain or that nudging, by adding psychology to the matter, is just a little bit more than external persuasion, the reality is that this way of changing behaviour appeals to governments because of its alleged low-cost applicability, of its simplicity and apparent guarantee of success (based on *ex ante* / *ex post* research).

Having in mind the conceptual framework of Libertarian Paternalism noted above and its application process by use of nudges, I return to Rebonato’s definition of Libertarian Paternalism.

“[It is] The set of interventions aimed at overcoming the unavoidable cognitive biases and decisional inadequacies of an individual by exploiting them in such a way as to influence her decisions (in an easy reversible manner) towards choices that she herself would make if she had at her disposal unlimited time and information, and the analytical abilities of a rational decision-maker (more precisely, of *Homo-Economicus*)”<sup>150</sup>

Under this definition, libertarian paternalists defend that: the knowledge of how *Humans* do decide and behave, against assumptions of how they ought to, if they were *Econs*, can and should be used to change the environment that people find themselves in.

Automatic decisions, those that occur without due deliberation, can be counteracted<sup>151</sup> by way of “choice architecture” and of “nudges” in order to align with

---

<sup>149</sup> Sunstein and Thaler (2009) p. 6

<sup>150</sup> Rebonato (2012) p. 6

<sup>151</sup> Sunstein (2014a) in response to Rebonato’s definition assumes exploitation as too aggressive a word and proposes “counteracting” instead. Notwithstanding, it is also true that in Sunstein and Thaler (2009) the verb “to exploit” appears in some instances when regarding the application of nudges.

individuals' deliberative preferences, i.e. preferences that, in retrospect, are what people really desire as "*judged by themselves*". Thus, they do not imply to satisfy people's revealed preferences, but rather what people "truly" prefer<sup>152</sup>.

Here libertarian paternalists demonstrate three different highly relevant beliefs; (1) that, in retrospect, they know what individuals' "true" preferences are; (2) that policies done under this thesis are non-coercive because they're reversible/optional and; (3) that Libertarian Paternalism becomes justified because it is proven that human beings don't always act in their best-interest, often harming themselves.<sup>153</sup>

"If individuals' preferences are sensitive to framing, the planner's choice of frame can affect the preferences that individuals reveal; there is no way of standing back and simply respecting preferences. Once the planner recognises this fact, she cannot avoid a decision about the direction in which to steer the individual, and the only reasonable criterion for making this decision is the planner's judgement about the individual's best interests. This is the sense in which there are no viable alternatives to paternalism. But because individuals remain free to ignore the planner's framing, this form of paternalism can (it is said) still be called 'libertarian'"<sup>154</sup>

### ***2.3.1 Applications of Choice Architecture***

Metaphorically nudges are presented as a GPS system<sup>155</sup>. A GPS indicates the best route to reach an intended place and it provides information about the best means to employ without forcing to take a recommended route. It serves as a mere suggestion, not an imposition, even though it is trying to steer someone towards a specific goal (set

---

<sup>152</sup> Sugden (2008; 2017)

<sup>153</sup> Rebonato (2012)

<sup>154</sup> Sugden (2008) p. 228

<sup>155</sup> Sunstein (2014a) p. 61

by one's self). Under this light, nudges are intended to focus on the means employed by individuals rather than the ends chosen.

As Libertarian Paternalists claim to approve only choice preserving policies, nudges cannot force, coerce or impose extensive costs on citizens. Measures that present themselves as hard forms of paternalism and therefore go against the maxim by which libertarians stand for, cannot be considered nudges. In short, they claim that their position is a means oriented soft form of paternalism<sup>156</sup>. With Libertarian Paternalism people's true ends are not judged or changed, it is rather the means employed to achieve those ends that are.

What Sunstein and Thaler call "choice architects" are any of those people that have the capacity and responsibility to create the environments/contexts in which people make decisions. These authors assert that choice architecture is inevitable and everywhere<sup>157</sup>. The private sector, government, society on the whole, all play part in designing the social background in which choices are made. As it is impossible to live without a social background, they argue that choice architecture is unavoidable and as such paternalism is the only viable option<sup>158</sup>.

To be fair, libertarian paternalists admit that often conscientious choice architects don't necessarily have people's best interests in mind, designing environments in order to serve their own purposes. However, they believe that self-consciously benevolent constructed nudges can steer people in the direction of their real preferences<sup>159</sup>.

To the extent that governments design policies under the belief that those affected by those same policies will behave in a certain way, policymakers fall under

---

<sup>156</sup> *Ibid.*

<sup>157</sup> Sunstein and Thaler (2003a; 2003b); see also Sunstein (2014a)

<sup>158</sup> *Ibid.*

<sup>159</sup> *Ibid.*

the category of choice architects. Both Sunstein and Thaler agree that in order to design efficient interventions, policymakers need to understand human behaviour. Only by understanding human psychology<sup>160</sup> choice architecture can be effective in facilitating and improving people's decisions towards their own well-being. The authors here are presuming that, when facing and recognising the inevitability of the impact their "designs" have on people, choice architects can, and will, create an environment where everyone will benefit. The "nudger" is thus a kind of benevolent king, "designing a garden maze that leads sinners to the right exit"<sup>161</sup>. As of yet, Sunstein and Thaler don't seem to provide a concrete and substantial reason to why it is unviable for the choice architect to act with other priorities in mind aside from welfare.<sup>162</sup>

Furthermore, taking behavioural concepts discussed previously such as; anchors, availability, loss aversion, framing, the status-quo bias, hyperbolic discounting, etc., libertarian paternalists propose a number of applications of choice architecture in the form of nudges. The idea is to use psychological mechanisms that economists have traditionally regarded as non-rational, but which are evidenced to influence behaviour<sup>163</sup>. By engaging these behavioural mechanisms Thaler and Sunstein propose six different principles of choice architecture, and they claim that individuals actually want to make the choices they're being nudged towards.

"(...) bounded rationality and cognitive biases lead people to make what they themselves see as serious errors, or would see as serious errors after reflection, and choice architecture should be established to help make those errors less likely or less damaging"<sup>164</sup>

---

<sup>160</sup> Sunstein and Thaler (2009)

<sup>161</sup> Poole, S. (22-9-2014). "*Not So Foolish*". Aeon

<sup>162</sup> Oliver (2013a); see also Sugden (2008); Rebonato (2012)

<sup>163</sup> Oliver (2013a)

<sup>164</sup> Sunstein (2014a), p. 157

Sunstein and Thaler in their book offer the word “nudge” as mnemonic device to remember their six principles of choice architecture, these serve as a guide to the choice architect.

iNcentives  
Understand Mappings  
Defaults  
Give Feedback  
Expect Error  
Structure Complex Choices

To analyse these six principles in more detail it must be understood that they imply two different types of policy interventions, one is mostly non-intrusive and focus on providing information and on complementing traditional incentives. In bringing behavioural insights to the provision of information libertarian paternalists propose to frame salient information in way of influencing behaviour. The other type of nudge policy proposed by libertarian paternalists can be seen as more intrusive, as it offers to make choices for us, and to exploit behavioural insights to policymakers’ own advantage, this is the case of default rules.

### ***A. Defaults***

Possibly the most prominent form of nudge policy in the literature are *default rules*. They intend to tackle a few heuristics that people seem prone to adhere to: *the status quo-bias, inertia, procrastination, the ‘yeah whatever’ heuristic*<sup>165</sup>, and so on. Robust empirical findings on human decision-making show that people tend to stick to any given situation. Either because individuals are lazy, fearful or just distracted they

---

<sup>165</sup> Sunstein and Thaler (2009)

prefer to stay with the *status quo* even when there are virtually no costs in changing to a better situation. This has been known and widely exploited by commercial and marketing businesses for their own benefit and people's own demise.<sup>166</sup>

Some argue that default rules are inevitable<sup>167</sup>, there is always an initial *status quo* to any given situation. Libertarian paternalists propose that defaults can be used to nudge, under the knowledge that people tend to “stick” with whatever initial situation is provided for them. Furthermore, policies that rely on the setting of defaults preserve freedom of choice because people are still permitted to “opt out” and “reverse the nudge”. To public or private institutions, the implications that defaults stick, especially when people ignore them, make the option of setting specific defaults very attractive<sup>168</sup>.

For his part Sunstein believes that in any given policy the choice of the default, especially those involving automatic enrolments, should be analysed in detail by experts who are sufficiently informed<sup>169</sup> and that take into consideration all possible variables. In this regard, he argues in favour of intelligently assigned default rules, having in mind that “when people trust the choice architect, automatic enrolment will be particularly sticky, but not so much when trust is low (...)”<sup>170</sup>

Sunstein differentiates between three different types of defaults: (1) impersonal default rules: default options that are set for a wide population, establishing what happens when people do nothing; (2) active choosing: where the choice is left for the individual, to Sunstein it can be either voluntary or mandated; and (3) personalized defaults: rules that are applied to specific members of a given population in order to fit the best option for those to which the default applies (this involves a process known as

---

<sup>166</sup> *Ibid.* see also in Garvey (2016)

<sup>167</sup> Sunstein and Thaler (2003a; 2003b; 2009)

<sup>168</sup> Sunstein, C. R. (2012). *Impersonal Default Rules vs. Active Choices vs. Personalized Default Rules: A Triptych*. (unpublished manuscript). Here Sunstein notes that these types of defaults should be seen with great care as it has been shown that defaults are particularly “sticky” to lower income workers.

<sup>169</sup> *Ibid.*

<sup>170</sup> *Ibid.* p. 14

narrowcasting which in the most extreme form can be applied to one specific individual).

When regarding the promise of personalized defaults, which he believes are the wave of the future and can attain large social benefits<sup>171</sup>, Sunstein proposes to use them to institute what he calls “personalized paternalism”<sup>172</sup> - a form of paternalism that is able to create different policy interventions by narrowcasting groups and thus without threatening to infringe on society’s heterogeneity.

Because personalized defaults are based in the retrieval of specific information about the people they’re trying to target (through predictive analytics), Sunstein admits the underlying threat regarding liberty. These measures can be easily used to promote choice architects’ own interests<sup>173</sup> and at the same time pose risks towards individual privacy. Furthermore, because personalized defaults will be designed regarding past behaviour, such rules do not promote individual learning, nor broaden personal experience and perspectives<sup>174</sup>.

As for the application of impersonal defaults or active choosing, different situations must be thoroughly analysed. According to Sunstein, active choosing (to voluntarily make a decision) should be promoted in familiar situations where learning is a benefit and in areas where choice architects lack pertinent information. On the other hand, in the instance of a relatively heterogeneous population impersonal defaults become the most feasible to apply, especially in contexts that are unfamiliar to people and where individual learning is not deemed relevant<sup>175</sup>.

---

<sup>171</sup> *Ibid.*

<sup>172</sup> Sunstein (2014a)

<sup>173</sup> Rebonato (2012)

<sup>174</sup> Sunstein (2014a)

<sup>175</sup> Sunstein (2012)

### ***B. Expecting Error***

When designing a policy, policymakers should expect humans to err. Different industries have realised, for some time, that to know common individual mistakes is highly relevant in designing efficient products. Only by understanding that forgetfulness, habits, automatic responses, lack of attention etc. lead people into systematic errors can a policy be designed to circumvent predictable mistakes, and consequently become more efficient.

This is not exactly new; policies, products, services have all been designed and redesigned in ways to get around constant nuisances. In the health industry, for example, making patients take the right medicine at the right time with the right frequency can be difficult<sup>176</sup>. The automatic system acknowledges that some rules are better than others<sup>177</sup>. As such, libertarian paternalists invite governments to study common human mistakes in way of designing policies that avoid them. One form to achieve this is by imposing frequent and constant practices, by for example creating habitual behaviour or providing immediate feedback.

### ***C. Informational Interventions that Target our System 1***

The other principles of choice architecture can rely on behavioural insights of saliency, of framing and disclosures. To the libertarian paternalist the provision of information has two levels of action; (1) where it educates, providing objective and all-encompassing information involving our critical reasoning; (2) where it presents

---

<sup>176</sup> Sunstein and Thaler (2009) p. 98 “Birth control pills present a special problem along these lines, because they are taken everyday for three weeks and then skipped for a week. To solve this problem and make the process automatic, the pills are typically sold in a special container that contains twenty-eight pills, each in a numbered compartment. Patients are instructed to take a pill every day, in order. The pills for days twenty-two through twenty-eight are placebos whose only goal is to facilitate compliance for Human users”

<sup>177</sup> Sunstein and Thaler (2009)

selected information, specifically targeting people's System 1 through the use of behavioural and psychological mechanisms (framing, saliency, anchoring, etc.).

To the libertarian paternalists, disclosures, leaflets and warnings can all highlight information that doesn't account for all variables, nor give us "the full picture" necessary to analyse the decision at hand. As selecting the information becomes part of the role of public officials, and as they intend to use that information to, effectively, target our less-scrutinizing selves in order to nudge us towards our own well-being, Sunstein and Thaler assume these types of nudges as paternalistic.

Regarding incentives libertarian paternalists do not dismiss the role of traditional market incentives, but rather feel that they can complement them by using behavioural insights. This can be achieved by framing salient information in order to make people see the incentive at hand or the possibility of an incentive conflict. Public officials can thus "manipulate salience"<sup>178</sup> in order to direct people's attention to the incentive at hand, thus making it more effective, or so do paternalist claim.

So far, we have seen how nudges and choice architecture are a fundamental part to the practice of Libertarian Paternalism. Both these concepts serve as tools to create low-cost and simple policies, while allegedly preserving citizens' freedom of choice by allowing the reversibility of, or possibility to ignore, those policies. By treating citizen's as they are<sup>179</sup>, behavioural insights and knowledge about human psychology can, and should, be used to create effective interventions in accordance to citizens' "true" preferences.

---

<sup>178</sup> *Ibid.*

<sup>179</sup> Dolan, P.; Hallsworth, M.; Halper, D.; King, D.; Vlaev, I. (2010); *MINDSPACE: Influencing Behaviour Through Public Policy* Cabinet Office and Institute for Government

Moreover, the underlying assumptions of Libertarian Paternalism are that human beings have bounded rationality and as such are prone to make mistakes when choosing. Under this light, they argue against a belief that individuals act in their best-interest and without harming themselves (casting doubts on Mill's harm principle and suggesting that his "Epistemic Argument" is not always right)<sup>180</sup>. The idea is to engage with psychological mechanisms that economists have traditionally regarded as non-rational, and to influence choices that, in retrospect, individuals want to make<sup>181</sup>, but that they are unable to on their own.

Nudges are then proposed in ways to alter the behaviour of *Humans* without incurring costs on *Econs*' behaviour. As such, these types of intervention target our System 1, hypothesised by Dual Process Theory, and are purposefully made to be ignored by the *Homo Economicus*<sup>182</sup>, who doesn't suffer from limits in rationality.

---

<sup>180</sup> Sunstein (2014a)

<sup>181</sup> Sugden (2017)

<sup>182</sup> Sunstein and Thaler (2009)

## **CHAPTER 3:**

### **Libertarian Paternalism in Practice**

Throughout the literature of Libertarian Paternalism, specific nudges have been exemplified for implementation. Some of these examples have already been put into practice and are used to illustrate how simple and effective nudges can be. Other examples are policy proposals by these paternalists, as the authors state that: “(...) many of the most important applications of Libertarian Paternalism are for government”<sup>183</sup>.

In this chapter my initial focus will be on specific proposals referenced in Sunstein and Thaler’s book: *Nudge*. These will show some type of interventions that the authors have in mind when proposing their thesis. Following that section, I propose to analyse three different nudging policies that have been implemented in the UK by The Behavioural Insights Teams (BIT), inspired by nudge theory and under the advisement of Richard Thaler.

#### **3.1 In ‘Nudge: Improving Decisions about health, wealth and happiness’**

There is some difficulty in understanding the scope of interventions that Libertarian Paternalism subscribes to. Some of the policy suggestions, whether private or public, fall under such disparate categories and touch on such a broad spectrum of themes that it becomes hard to define what is the ultimate objective of this thesis, and whether some proposed policies can be called paternalistic. As they claim and justify themselves under pretences of ameliorating behaviour and improving decisions, it seems pertinent to demonstrate different policy suggestions present throughout their literature.

---

<sup>183</sup> Sunstein and Thaler (2009) p. 14

At the same time some of their policies have been dismissed as they don't seem at all relevant to our discussion<sup>184</sup>. In this dissertation I will try to focus on interventions that alter choices of *relevant* behaviour, nudges that libertarian paternalists assume easy to reverse or ignore, and that intend to target our less scrutinizing and more automatic selves.

### 3.1.1 The Cafeteria<sup>185</sup>

The introduction of Sunstein and Thaler's book presents a hypothetical case in order to demonstrate the impact of choice architecture and nudging: A director of a food services for a large city school system realises the power of influence over what students choose to eat by changing the arrangement of the food display in the cafeteria.

It is interestingly odd to notice that the first case presented for the introduction of Libertarian Paternalism depicts a *children's* cafeteria. True, the authors are presenting an example that can be applied to any cafeteria, whether children's or adults'<sup>186</sup>. However, it becomes relevant to note that in their most famous book, serving almost like a kind of manifesto, they choose the former. The "judged by themselves" clause seems to lose somewhat of its meaning<sup>187</sup> as children tend to be considered not entirely responsible for their choices, at least not until they have reached the specific age of adulthood. Also, the use of this example stands out because the legal distinction between paternalism regarding children's choices and that relating to adults seems to go

---

<sup>184</sup> *Ibid.* p. 4; one of the most quoted examples by Thaler and Sunstein to illustrate how a nudge and how details of choice architecture can highly impact people's behaviour is the case of the Schiphol flies. The argument goes that authorities at the Schiphol airport, in Amsterdam, have painted a fly on urinals in men's toilet facilities. Apparently, this simple detail ended up being responsible by improving aim and reducing spillage by 80%. in

<sup>185</sup> *Ibid.* pp.1-3

<sup>186</sup> In Sunstein and Thaler (2003a; 2003b) pp. 5-6 they present a similar case but considering a private "cafeteria at some organization".

<sup>187</sup> Rebonato (2012)

amiss<sup>188</sup>. Notwithstanding, let us consider some important aspects of this suggested intervention.

According to the authors, once the realisation is reached that a given planner has an enormous impact on individual choice by just displaying and arranging the food differently, then the choice architect (the planner) will have different options of conscious action to consider:

1. Arrange the food to make students better off.
2. Choose the food order at random.
3. Try to arrange the food to get the kids to pick what they would choose on their own.
4. Maximize the sales of the items from the suppliers who are willing to offer the largest bribes.
5. Maximize profits<sup>189</sup>.

Thaler and Sunstein assume that the first option is the best, or better yet, that all other four options are far worse. They admit that it might be a paternalistic measure, perhaps a bit intrusive even, but by comparison the only viable option. Option 2 is immediately dismissed, it is completely unfeasible and raises questions to the practicality of what and how to arrange food at random. Planners and students have nothing to gain from random organizational criteria.

Also, the authors are here presenting the case of an “honourable and honest” character and as such option number 4, depicting the choice architect as relatively corrupt, must be dismissed. Once again, the presumption of a benevolent choice

---

<sup>188</sup> *Ibid.*

<sup>189</sup> In Sunstein and Thaler (2003a) they omit this option, but interestingly include 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”.

architect stands, but here there's the addendum of describing personal characteristics of the planner to justify their thesis.

Option number 5 is attractive and not at all withdrawn from reality, no one would question such practice (it is legitimate) but will only be put into practice if the cafeteria's main goal and priority is profit. And it is here that the conundrum lies, should one try to maximise profits possibly resulting in unhealthier children and "especially since she works for the school district?"<sup>190</sup> It seems that the relevance of presenting a school as an example is that it would be morally wrong to accept any other options besides number 1, as all the others options are deemed less worthy in respect to children's welfare.

Option 3 is interesting to dissect. According to Sunstein and Thaler, it's the one anti-paternalists would easily favour, but that the authors easily dismiss because it relies on *ex ante* preferences which from a behavioural standpoint do not exist – as we have seen that preferences are context-dependent<sup>191</sup>. Assuming that, not only one different arrangement, but that *any* arrangement will have an impact on people's choices, the authors have concluded that there is no natural, "neutral" way of presenting choice – and as such people's "true 'preferences' do not formally exist"<sup>192</sup>.

If this argument is right, then the problem is that real preferences cannot be truly observed, libertarian paternalists apparently are abandoning the methodology of analysing revealed preferences, the main empirical foundation for social sciences and for welfare interventions. At the same time, it seems that they're adopting a more conceptual form of preference. In a more recent work Sunstein traces back an idea from Harry Frankfurt on acting according to what "we want to want" - our second-order

---

<sup>190</sup> Sunstein and Thaler (2009) p. 3

<sup>191</sup> Sugden (2017)

<sup>192</sup> Sunstein and Thaler (2003) p.6

desires. If we act in order to satisfy our deepest values then we are acting according to our own free will<sup>193</sup>. I will further analyse this implication in chapter 4.

Recalling option 1 and accepting it as true, that the arrangement of the food display has a direct impact on people's choices of what they eat, then one can conclude as libertarian paternalists do, that even the slightest details affect people's behaviour. From this the authors seem to extrapolate that a form of paternalism becomes unavoidable (against the first misconception of libertarians).

“The first misconception is that there are viable alternatives to paternalism. In many situations, some organization or agent *must* make a choice that will affect the behavior of some other people. There is, in those situations, no alternative to a kind of paternalism – at least in the form of an intervention that affects what people choose.”<sup>194</sup>

Again, we must understand with this that libertarian paternalists seem to imply that any form of influence on behaviour is paternalism, or what they call a “kind of paternalism” raising several questions on its own. But even accepting this argument, can we consider it wishful thinking, to presume as the only viable option available the paternalistic one? One can argue that some unconsidered options are still possible, no matter how preposterous they might sound. From a liberal standpoint, nothing should stop the choice architects from acting in their own self-interest, provided harm doesn't come to others.

---

<sup>193</sup> Sunstein (2016) p.86 “(...) free will exists when we make choices that fit with our own deepest values—those that we endorse after reflection. If you decide to become a doctor, to quit smoking, or to be a bit kinder in the workplace, you are probably exercising your free will. The philosopher Harry Frankfurt makes a distinction between what we want (a cigarette, a bit of extra sleep, a visit to the Dark Side) and what we want to want (no cigarettes, a little more work, the Light). Frankfurt argues that if we act in accordance with what he calls our second-order desires, we are exercising our free will. For Frankfurt, freedom has a lot to do with the ideal of self-mastery” p. 85

<sup>194</sup> Sunstein and Thaler (2003a) p. 1164

Even when considering state interventions, one can still assume that the state has the capacity to act in its own interest. Democratic governments can indeed do this under risk of suffering the consequences from the opposition, from the electorate, from civil society, etc. For the libertarian paternalists it seems that their belief in the benevolence of planners is what makes it impossible for them to admit other viable solutions<sup>195</sup>.

### ***3.1.2 Retirement Plans and the Save More Tomorrow Scheme***

Standard economic theory assumes that regarding savings and retirement, people have the rational ability to save throughout their lives. Within Modigliani's life cycle hypothesis, for example, one is expected to optimise consumption and saving patterns in order to maintain a balanced lifestyle. "[I]ndividuals accumulate and run down assets during their expected life cycle so as to maximise their lifestyle utility function"<sup>196</sup>.

For behavioural economists there lie two fundamental problems with this theory. First a mathematical one; there exists too many variables and complexities to calculate the analytical problem of how much to save in order to have a good life after retirement, especially considering that different calculus might apply to different individuals. Even with the most sophisticated software, economists and financial advisors still have difficulty in unanimously finding a correct savings rate<sup>197</sup>.

On the other hand, there is a problem of will-power, or lack thereof. As human beings are seen to have *present bias* and *akrasia* (failure of self-control), paternalists argue that "even if the correct savings rate were known, households might lack the self-control to reduce current consumption in favour of future consumption"<sup>198</sup>.

---

<sup>195</sup> Mitchell (2005)

<sup>196</sup> as explained in Rebonato (2012) p. 65

<sup>197</sup> Sunstein and Thaler (2009)

<sup>198</sup> Benartzi, S.; Thaler, R. H. (2004). Saving More Tomorrow: Using Behavioural Economics to Increase Employee Saving. *Journal of Political Economy*, Vol. 112 No. 1:164-187. p. 165

Evidence implying that people don't act in their own best interests is extensive, especially in the case of saving towards the future. Individuals don't actively enrol in beneficial plans that allow to save for retirement, even when there are little costs associated to such enrolment. This has been shown to happen, in defined-contribution plans, where the employee signs up of his free volition deciding on a specific rate of savings. But more surprisingly, this has also been noticed to happen in defined-benefit plans where the employees only have to sign up and the responsibility of providing towards a pension fund falls to the employers<sup>199</sup>.

Under this evidence, libertarian paternalists assume that the decision to not participate in a retirement plan is a rational mistake.

“By calling [people's] low-saving behaviour a mistake, we mean that they might characterize the action the same way, just as someone who is 100 pounds overweight might agree that he or she weighs too much.”<sup>200</sup>

One way of tackling this issue is by making enrolment in a savings plan automatic, by making it the default rule. In 2001 Benartzi and Thaler presented a plan to overcome this irrational conundrum of why people don't save towards retirement. The Save More Tomorrow (SMT) scheme intends for people to commit, beforehand, portions of their future salary increases allocated towards their retirement savings. With this scheme, people are offered a plan in which their savings rate go up by 3 percentage points every time they get a pay raise and where a typical pay raise is between 3.25 and 3.50 per cent. If a person continues with this scheme for years of pay raises, their saving rates will constantly increase.

---

<sup>199</sup> *Ibid.*; see also Thaler and Sunstein (2009);

<sup>200</sup> Benartzi and Thaler (2004)

They propose to do this by way of exploiting three known cognitive bias: (1) by overcoming people's *present bias* or *hyperbolic discounting*, which assumes that people attach particular weight to the immediate present time in detriment of their future. (2) by using to their own advantage the generalized poor understanding of the difference between nominal and real income through a combination of *money illusion* and *loss aversion*; and (3) by taking advantage of *inertia* and *procrastination* in order to maintain people in their scheme once they sign up.

“ [P]resence of bounded rationality suggests that the program should be simple and should help people approximate the life cycle saving rate if they are unable to do so themselves. Hyperbolic discounting implies that opportunities to save more in the future will be considered more attractive than those in the present. Procrastination and inertia suggest that once employees are enrolled in the program, they should remain in until they opt out.

The final behavioral factor that should be considered in designing a prescriptive savings plan is loss aversion, the empirically demonstrated tendency for people to weigh losses significantly more heavily than gains”<sup>201</sup>

It is implied then that the authors are not proposing to explain or help people overcome these decisional blunders, they rather use them to their advantage by the generalized lack of financial literacy. When regarding possibilities of education in this matter libertarian paternalists agree that “unfortunately, the evidence does not suggest that education is in and of itself, an adequate solution”<sup>202</sup>. Thus, they don't consider that helping employees make better financial decisions on their own is viable, but instead prefer to “take advantage” of people's cognitive mistakes in order to increase their savings.

---

<sup>201</sup> *Ibid.* p. 169

<sup>202</sup> Sunstein and Thaler (2009) p. 121

Furthermore, given the difficult computation of the “right” saving rate even for rational selves, as admitted by the libertarian paternalists, Benartzi and Thaler assume that their defined savings rate (almost everything of their pay raise) will satisfy the employees’ preferences.

Moreover, when considering the advantages of *inertia*, the *status quo bias* and *procrastination*, one can entertain a critical view of this libertarian paternalistic proposal. As they claim these biases to be strong influences on behaviour, we can assume that when imposing a given plan on people’s lives, compliance is due to inertia itself. However, libertarian paternalists claim that the extremely high take-up rate of their choice (after their measures) is in itself an endorsement of how the choices they promote are what people truly want.

“The vast majority of those enrolled in the SMarT plan (80 percent) remained in it through the fourth pay raise, and (3) the average saving rates for SMarT program participants increased from 3.5 percent to 13.6 percent over the course of 40 months.”<sup>203</sup>

By behavioural standards, this high rate of acceptance does not prove much about the preservation of choice, as cognitive biases go both ways<sup>204</sup>. In accepting that these cognitive biases and heuristics are intrinsic to behaviour, and by aligning people with a specific, not fully transparent plan, libertarian paternalists could impose measures that people “truly” hate and yet don’t change because of their decisional blunders<sup>205</sup>.

Considering that almost everybody (80%) under this plan stayed with it for four consecutive pay raises (whereupon the contributions only stopped because they had

---

<sup>203</sup> Benartzi and Thaler (2004) p. 165

<sup>204</sup> Rebonato (2012)

<sup>205</sup> Rebonato (2012)

reached the maximum amount allowed to contribute) one can conclude, as the libertarian paternalists do, that employees, once enrolled, either saw the benefits of the plan and in retrospect stayed with it because it was what they truly wanted. Another person might reach a different conclusion, perhaps it was their decisional blunders, not their preferences, that made them stick with the SMT scheme, and maybe employees intended to change their options but somehow never got around to do it<sup>206</sup>. As the authors admitted to design the SMT in order to take advantage of the *status quo* bias and *inertia* it is entirely feasible to argue in favour of the latter.

Furthermore, the SMT scheme is designed as a response to people's tendencies to not actively save for their retirement. However, their proposal is not to actively *save* for retirement but rather, to *put aside future* pay increases. In their dual-self model of present-self and future-self, they prioritize individuals' future to the possible detriment of their present, as they seem to assume the future-self to depict "true" preferences<sup>207</sup>. They do so by exploiting the generalized poor understanding of the difference between nominal and real income.

They claim that by using future pay increases individuals will never see their incomes going down and so won't perceive any losses, taking advantage of *loss aversion* at the cost of being unable to enjoy future pay rises. They sell their programme without mentioning or alluding to the fact that in committing future pay increases to a savings plan, individuals "nominally-non-decreasing future pay cheque will have a smaller and smaller purchasing power"<sup>208</sup>. Benartzi and Thaler are not trying to overcome people's inability to focus beyond what they receive in the form of money,

---

<sup>206</sup> In Sunstein and Thaler (2009) to demonstrate status quo bias the authors relate Sunsteins' personal experience with magazine subscriptions: "For about a decade, he has continued to subscribe to magazines that he hardly ever reads. (He keeps intending to cancel those subscriptions, but somehow never gets around to do it.)" p. 38

<sup>207</sup> Rebonato (2012)

<sup>208</sup> *Ibid.*

but exploit the fact that people dismiss how their economic condition is affected, not merely by the money they receive, but by their purchasing power. As employees never see their pay check go down, they will never perceive any losses, and therefore consider the plan a win-win situation.

The SMT programme had initially a participation rate of 80 percent in environments where financial advisors met with people personally and recommended participation in the savings plan. By contrast, participation has been rather low where a person needs to decide independently and find a specific plan to participate in. One of the paternalistic proposals in response to this is to combine the SMT plan with automatic enrolment<sup>209</sup>. Libertarian paternalists also invite governments to “incorporate a Save More Tomorrow feature into a national savings plan. Workers would be automatically enrolled in a pension plan in which the amounts devoted to savings would increase with increases in wages”<sup>210</sup>

### 3.2 The British Movement

In 2010, just before David Cameron took office, he gave a talk where he expressed his belief in being at the verge of a revolution in the way of government. According to him, “with the very simple conservative thought of going with the grain of human nature, you can achieve so much more”<sup>211</sup>. As he went on to briefly explain the vast developments that have occurred in behavioural economics and psychology, he stated that to “treat [people] as they are, not as government wants them to be”<sup>212</sup> might be the key to tackle individual behaviour in a way to improve welfare with little costs

---

<sup>209</sup> Benartzi and Thaler (2004); they estimate that with automatic enrolment 90 percent of employees would participate in the programme and only 10 percent would actively opt out.

<sup>210</sup> Sunstein and Thaler (2009)

<sup>211</sup> David Cameron (Feb. 2010) “*The Next age of Government*”. TED2010 Conference

<sup>212</sup> *Ibid.*

associated. David Cameron was here referring to the introduction of behavioural sciences into the development of policymaking in the United Kingdom.

With the appointment of David Cameron as the new prime-minister at the head of the then Conservative-Liberal Democrat Coalition, the Behavioural Insights Team (BIT) was created. Informally known as the “nudge unit” BIT was, at that stage, the world’s first policy dedicated central government behavioural unit<sup>213</sup>.

It seems pertinent to highlight that since BIT was formed in 2010 different behavioural public policy units have been consistently integrated into government departments, many adopting (not exclusively) policy interventions under nudge theory. In 2015 the World Bank Development Report retained a large interest in behavioural public policy and initiated the Global Insights Team<sup>214</sup>. The following year the European Commission released a report with the following introductory remark:

“From a policy perspective, relying on unrealistic assumptions about people’s behaviour may have severe consequences. If people’s behaviour is primarily due to lack of knowledge or information, then conventional education or information campaigns could constitute an appropriate remedy. If, on the other hand, people’s behaviour reflects fundamental aspects of human nature (such as default bias, present bias, loss aversion, overconfidence etc.), a more effective approach would be to take such behavioural features into account when designing policy. Identifying the reasons underpinning people’s behaviour is therefore an essential prerequisite for effective policy making.”<sup>215</sup>

This major shift seems to have occurred with the publication of *Nudge*. However, it is wrong to assume that nudging is the only form of policies adopted by these behavioural units, it is merely contended as “one of the several behavioural techniques policymakers can use to direct people to better choices without using bans or

---

<sup>213</sup> Oliver (2017)

<sup>214</sup> Oliver (2017)

<sup>215</sup> Lourenço *et al.* (2016), p. 8

other expensive time-consuming alternatives”.<sup>216</sup> Thus, it seems pertinent to highlight that since Thaler and Sunstein’s book was published, there have been many new developments in behavioural public policy.

Interest in the implementation of BIT might have originated with Sunstein and Thaler’s work. Yet, its intellectual foundation traces back to a more recent report released in 2010 by the Cabinet Office and Institute for Government known as the “MindSpace Report”. Many of the authors in this report would eventually be part of the nudge unit alongside Richard Thaler who would join the team later on.

Following the lines of Libertarian Paternalism, they argued in favour of a soft touch to policy, under pretences of behavioural insights, intending to go with the grain of human nature and not against it<sup>217</sup>. Similarly, the MindSpace report was given its name by its different guiding principles:

**Messenger**: We are heavily influence by who communicates information.

**Incentives**: Our responses to incentives are shaped by predictable mental shortcuts such as strongly avoiding losses.

**Norms**: We are strongly influenced by what others do.

**Defaults**: We “go with the flow” of pre-set options.

**Saliency**: Our attention is drawn to what is novel and seems relevant to us.

**Priming**: Our acts are often influenced by subconscious cues.

**Affect**: Our emotional associations can powerfully shape our actions.

**Commitments**: We see to be consistent with our public promises, and reciprocate acts.

**Ego**: We act in ways that make us feel better about ourselves.

According to the authors these effects (insights from laboratory and field research in social psychology) seemed to be the most robust for tackling behaviour

---

<sup>216</sup> *Ibid.* p. 10

<sup>217</sup> Dolan *et al.* (2010)

originating, not exclusively, on automatic responses.<sup>218</sup> The Mindspace report would then be used by BIT as a “framework to aid the application of behavioural science to the policymaking process”<sup>219</sup>.

As noted, some of the authors of this report would eventually form the Team and by adapting the concept of nudges and the characteristics of the Mindspace framework, they would conduct behavioural field research in order to suggest possible policies. Ultimately, BIT became a government department<sup>220</sup> that specifically used empirical research and randomised controlled trials to assess what could be the most effective policies regarding different themes, making *ex post* suggestions to government for potential policies.

It should be noted that, even though it departed from a paternalistic standpoint, the Team’s research often focussed not only on individuals’ “own welfare” but on a general idea of social welfare too. Some argue that this was purposefully done to gather broader public support, which might have been curtailed if offering policies which focused only on affecting people’s own behaviour<sup>221</sup>. The “Mindspace” authors also highlighted the fact that, due to their policies’ nature, citizens might not be fully cognisant that their behaviours are being targeted. They concluded that this might leave governments open to charges of manipulation<sup>222</sup>.

---

<sup>218</sup> *Ibid.*; see also Dolan, P. (2010). Influencing the Financial Behaviour of individuals: the mindspace way in A. Oliver (Ed.) *Behavioural Public Policy*, Cambridge, UK, New York: Cambridge University Press

<sup>219</sup> Retrieved from BIT’s official website: <http://www.behaviouralinsights.co.uk/publications/mindspace/> (last accessed Nov. 2018)

<sup>220</sup> In 2014 the Behavioural Insights Team was no longer part of the Cabinet Office, it became a social purpose company, part of a joint venture between the UK government, Nesta (the innovation charity), and the team’s own staff.

<sup>221</sup> Oliver (2017)

<sup>222</sup> Dolan *et al.* (2010) p. 66

### 3.3 The Behavioural Insights Team in Practice

The Behavioural Insights Team in the U.K. was, in a way, intended to be the practical application of Libertarian Paternalism to government. As such, in analysing policy proposals by this team I will mainly focus on measures that have either been mentioned by libertarian paternalists, or that fall under the scope of Libertarian Paternalism as it has previously been defined.

#### 3.3.1 Organ Donation

The case of organ donation is mentioned throughout the literature, including in Sunstein and Thaler's textbook *Nudge*, who use this policy intervention as the perfect example for applying mandated choice or impersonalized defaults. How it can be considered a form of paternalism might be confusing, as it focuses on deceased donation and evidently, donation of organs does not imply benefits for one's own well-being, a matter that will be further addressed in chapter 4.

Notwithstanding, taking from the successful example of Illinois and other states in the USA, in 2011 BIT (alongside other government departments) conducted a pilot-study of prompted choice<sup>223</sup> when an individual applied online for the DVLA – Driver Vehicle Licensing Agency<sup>224</sup>. In short, upon online application to get or renew a driver's licence, a person was prompted to decide whether they wanted or not, to be an organ donor.

Since 2013 prompted choice, or mandated choice as Sunstein and Thaler call it, has been the main practice in the UK. When applying for the DVLA (whether online or physical application) the applicant must decide if they want to donate their organs.

---

<sup>223</sup> Sunstein and Thaler (2009) they call this type nudge as “mandated choice” using the Illinois example themselves, also present in Sunstein (2014a)

<sup>224</sup> Behavioural Insights Team (2010); *Applying Behavioural Insight to Health*, London: Cabinet Office p. 10

An application cannot be submitted until the decision to become an organ donor (or not) is made.

Having this in mind, mandated choice, i.e. to force one to choose, can be argued to cross the line from a soft measure of paternalism, to a somewhat harder intervention. Even though a person is not being forced to choose a particular answer (the paternalists are not forcing what to choose) and a person is still fully capable of deciding to be, or not to be an organ donor, there are still too many coercive variables in this type of measure. Aside from the more obvious fact of conditioning their choice by making hostage what the person wants (in this case the driver's licence) we can also take into account the detrimental aspect to those who might feel embarrassed in deciding, and claiming publicly, that they don't want to become donors<sup>225</sup>. Also, in forcing to choose at a particular time of one's life, there might be imposed costs on future selves, in the case of regret for example.

Later, with the same goal in mind, BIT conducted, at the time, the largest randomised controlled trial in the UK in order to compare the effectiveness of different appeals to organ donation<sup>226</sup>. Following principles of advertisement, using a high-traffic government webpage the Team displayed eight different short messages as an appeal to donate organs, each message linking to the official organ donation website. A first message displayed for the control group which only read as "Thank you. Please join the NHS donor register." while the other seven messages used behavioural insights. They formulated their messages using tools of social proof<sup>227</sup>, of reciprocity, by exploiting

---

<sup>225</sup> Oliver (2017)

<sup>226</sup> Behavioural Insights Team. (2013). *Applying Behavioural Insights to Organ Donation: preliminary results from a randomised controlled trial*. London: Cabinet Office (2013)

<sup>227</sup> Evidence has shown that people acknowledge the correct form of acting once they acknowledge and see others doing it, it can take the form of peer pressure, justify crowd and herd behaviour, etc.

*loss aversion* and by employing imagery “to see whether [they] could increase the salience of the messages using visual cues”<sup>228</sup>

The conclusions of the report went on to say that the best results drew on messages of reciprocity, loss and social proof. They estimated that in the course of one year, as many as 96.000 more people would register to donate organs when comparing to the control group message. These messages that appeal to our subconscious cues, have been included in government websites alongside a “share” your decision on social media to provide for more social proof<sup>229</sup>.

Interestingly, England has recently conducted a public consultation<sup>230</sup> that ended in March 2018, on whether to make organ donation the default rule, imposing an “opt out” option for those who don’t want to donate their organs. This measure of presumed consent will apparently be in place from 2020<sup>231</sup>. England is not the first country to consider this form of presumed consent of organ donation which is becoming more attractive to governments due to its proven efficacy. In 2015, Wales became the first nation in the UK to adopt this type of policy, but for years Austria, Belgium, Portugal, France and others have worked under a policy of presumed choice as the default which I shall analyse in more detail in the next chapter.

### ***3.3.2 Employment***

In trying to help people back into work, BIT designed a few changes to the process of applications at job centres. With the same methodology of a randomised control trial the team divided about two thousand jobseekers into two different groups,

---

<sup>228</sup> Behavioural Insights Team (2013) p. 4

<sup>229</sup> Garvey (2016)

<sup>230</sup> Boseley, S. “Jeremy Hunt launches opt-out organ donation in England” *The Guardian* (12-12-2017)

<sup>231</sup> Department of Health and Social Care. (2018). *Consultation Outcome: Introducing “opt-out” consent for organ and tissue donation in England*

the control – where application process had no change – and the trial with a different process of application.

The differences in processing between the control and the treatment group were quite simple; on the first day of application, instead of filling seemingly endless forms, the claimant was guaranteed a conversation with a job advisor, which normally would only happen within two weeks. In that conversation a plan was devised between the two about what would be done for the next two weeks.

“Encouraging claimants to focus on making specific commitments to future activities, linked to their daily routines, helped them to follow through on their job search intentions”<sup>232</sup>.

If after 8 weeks the person still hadn't gotten a job, proposals for “building psychological resilience and well-being through ‘expressive writing’ and ‘strengths identification’”<sup>233</sup> was advised. The results of the first trial showed significant effects. After 13 weeks of initial application, jobseekers in the treatment group were 20% more likely to be off benefits in comparison to the treatment group. Under this conclusion the pilot study broadened its intervention area, eventually leading to the training of over 25.000 advisors and to instituting the commitment devices to all job centres<sup>234</sup>.

This is the perfect example of a less intrusive and more libertarian form of nudge policy that uses informational interventions and simple non-monetary incentives. The government here is merely providing better institutional support under light of behavioural evidence using, for example, the effectiveness of pre-commitment

---

<sup>232</sup> Behavioural Insights Team. (2015). *Update Report 2013-2015*

<sup>233</sup> Behavioural Insights Team (2012). *New BIT trial results: helping people back to work*. London: Cabinet Office p. 7

<sup>234</sup> *Ibid.* p. 9

strategies<sup>235</sup>, of providing direct feedback, of framing and mapping choices in the direction of welfare. However, some aspects of this measure must be considered, especially regarding the possibility of government abuse of these behavioural tools which goes unconsidered within the theoretical approach of Libertarian Paternalism.

Of course, we depart from the assumption that a jobseeker truly wants to find a job and so will comply with any form of institutional support that is offered, especially if in the form of nudges which tend to offer simpler means. Notwithstanding, during the pilot-study, job advisors were apparently implying that if claimants didn't subscribe to using their tools of "expressive writing and strengths identification" they would lose their benefits<sup>236</sup>.

The implication of suffering penalties when there's no compliance with a jobseeker's plan, adds a much more restrictive and less libertarian condition to this measure. It creates difficulties for the jobseeker to ignore the nudge as easily as implied by Sunstein and Thaler in their theoretical approach.

The Behavioural Insights Team did come out publicly claiming that such conditions were not envisioned for the study and so perhaps this implication was unfounded. As it is impossible to truly know whether this adjacent condition was intended by BIT, or if it happened due to an unaccounted human variable; I turn now to analyse the fact that the psychometric test, offered to jobseekers to build "psychological resilience" and to find their "strengths" was not scientifically validated in anyway<sup>237</sup>, seemingly a mere tool to keep jobseekers occupied and to target their automatic biases.

When pressed for an official response regarding this accusation the Department of Work and Pensions came out publicly and stated that the test used was "intended to

---

<sup>235</sup> *Ibid.*

<sup>236</sup> Malik, S. (30-4-2013). "Jobseekers made to carry out bogus psychometric test". *The Guardian*

<sup>237</sup> Malik, S. (6-5-2013). "Jobseekers' Psychometric test is a Failure". *The Guardian*

help jobseekers identify their strengths” and that they had had “extremely positive feedback from both jobseekers and their advisers – it is right that we use every tool we have to help jobseekers who want to work find a job”<sup>238</sup>

Notwithstanding the good results of this trial, that an increase by 20% in the probability of jobseekers finding jobs is no small feat, we still have to consider the implications that this successful endeavour presents. The mere fact that those participating in this trial had no knowledge nor consented to being part of it raises ethical questions<sup>239</sup>, but taking into consideration the statement that it is “right” for government departments to “use every tool available to help”, even demanding time and commitment from citizen’s to respond to fake tests and threatening those who don’t comply with penalties, then we can see a fundamental detachment from the theoretical view of Libertarian Paternalism and its practical application<sup>240</sup>.

From a behavioural standpoint, as public officials are human and therefore can suffer from the same infinite heuristics and bias as other individuals, we cannot dismiss the possibility of failures in policymaking. Interpreting evidence in the wrong way, not taking into consideration possible human variables in the application of nudges, and suffering from the same cognitive biases and heuristics as other individuals can lead public officials to create unwarranted policies.<sup>241</sup>

### ***3.3.3 Education***

Since its implementation BIT has been trying to increase university applications and enrolment, especially in good-grade students from disadvantaged social backgrounds that have a relatively low rate of university application (lower than 20%).

---

<sup>238</sup> in Garvey (2016) p. 84

<sup>239</sup> Malik (6-5-2013)

<sup>240</sup> Garvey (2016)

<sup>241</sup> Glaeser (2006); see also Ginsburg and Wright (2012)

As such they devised different field experiments trying to find nudges that could change this reality.

“Attending university is a life-changing opportunity that can bring high returns to both the individual and country (...).

The government has set clear goals on widening participation in higher education: to double the proportion of people from disadvantaged backgrounds entering university in 2020 compared to 2009, and to increase the number of black and minority ethnic (BME) students going to university by 20%.”<sup>242</sup>

In one of the concluded trials, BIT postulated that, by sending a modelled letter to target students by framing the advantages and explaining why they should apply to higher education – the perfect embodiment of an informational non-intrusive nudge - they could steer students to enrol.

The results of this particular randomised control test weren't very successful, albeit positive. They concluded that a personal message, signed by a real person and addressed to the student directly, had a better effect in making students enrol in university<sup>243</sup>. But even considering the low rate of success, it still seems pertinent to further analyse the experiment for the implications that this policy proposal possibly carries for decisional autonomy.

As it focused on a very specific group – good students from disadvantaged backgrounds – this form of policy becomes what Sunstein calls “personalized paternalism”<sup>244</sup>. A form of paternalism that does not apply measures to all, but rather to specific, narrowed groups. This type of measure is done in ways of preserving heterogeneity. To the libertarian paternalists this form of policy design is becoming

---

<sup>242</sup> Chandle, R.; Sanders, M.; Selley, E. *Encouraging People into University: Research Report*. London: Department of Education (2017) pp. 8-9

<sup>243</sup> *Ibid.*

<sup>244</sup> Sunstein (2014a)

available due to different technological advancements and to the collection of individual information on many fronts, serving the purpose of avoiding the “one-size-fits all” approach that traditional paternalism imposes<sup>245</sup>.

In analysing this case-study one can admit that it does not impose a wide based intervention, as it stipulates a given group with a defined criterion (disadvantaged backgrounds, good-grade students). However, one can still argue that this level of narrowcasting is still wide enough to disrespect heterogeneity. As this initiative tackles everyone in the group as a unit, assuming that all involved would, not only benefit from attending university, but also that they want to attend, the policymakers are continuously dismissing the subjectivity and diversity of individual preferences.

For the sake of argument let us accept that by narrowing a demographic one can fully respect heterogeneity and individuality, it is still possible to analyse other ways in which this type of intervention imposes relevant costs to our decisional autonomy, especially if made by using more covert nudges.

Contrary to the previous measures discussed above, here we are admittedly faced with the perfect example of a libertarian nudge, where young students are entirely free to ignore the nudge and none of their options are directly hindered by it. No one is forcing these students to regard or accept the information being provided, or even to account for it in their decisional process but let us assume that they do. Let us suppose that some of these young people did enrol in university, specifically due to this nudge. Not because they made the decision on their own grounds, nor because they complemented the nudge with other pertinent information or because they did their own independent research (Libertarian Paternalism intends to target our automatic systems of

---

<sup>245</sup> *Ibid.*

thought and not necessarily provide better and more accurate information to complement information that's already available).

Then, let us imagine that the student who enrolled in university merely based on a government message, regrets<sup>246</sup> the decision and is now forced to reverse the nudge. Considering all the costs that are associated with university<sup>247</sup> it seems plausible to assume that the easy-reversibility of such a measure is not as straightforward as implied by libertarian paternalists. The consequences of such a decision have a much bigger impact on the individual than a simple nudge should impose. In assuming to know individuals' true preferences and in accepting the benevolence of planners, libertarian paternalists seem to dismiss the possible implications of such measures on individuals that are nudged towards being "worse off" "as judged by themselves"<sup>248</sup>.

If the decision were to be made by one's own desire and assessment, by their own research and consideration, one would more likely assume the responsibility over any regrets, mistakes or costs associated with that, relatively, autonomous choice<sup>249</sup>. However, if a decision is made by being, unknowingly, subjected to a nudge that uses knowledge of our impulsive psychology to influence us to make the choice that didn't fit at all, to whom the responsibility of the outcome falls on becomes less evident.

Furthermore, policymakers are supposed to design the most effective policies possible. Thus, we can presume that when testing these measures BIT is wanting them to have the best results possible in order to comply with governments objectives. As a government sets targets, a policymaker's role becomes to employ the best means to meet those (in this case to increase university enrolment). In this trial the experiment

---

<sup>246</sup> Kahneman (2012) notes the importance of regret as a variable, commonly dismissed by planners, in the decision-making process, especially regarding risky and important choices.

<sup>247</sup> Coming from disadvantaged backgrounds these students will probably have better financial help in place, either by the university or by the government, (a point that they addressed in their modelled letters). Notwithstanding, there are still a lot of costs involved with any reversal of this type of choice.

<sup>248</sup> Rebonato (2012)

<sup>249</sup> Ginsburg and Wright (2012)

wasn't as effective as intended, but the government's target to increase enrolment still exists, and so one can only assume that other attempts at designing more effective policies will be made. These future policies might put other non-intrusive nudges in practice, but the space has been left open for governments to use less salient means and involve further surreptitious interventions.

Because libertarian paternalists are inviting governments to take advantage of our behavioural shortcomings, the risk of a slippery slope that all kinds of paternalism entail<sup>250</sup> remains present, alongside the risk of government abuse over manipulative means of persuasion.<sup>251</sup>

From default-rules to the use of non-intrusive nudges that only frame information in order to impact choice. From policies that are theoretically libertarian but that in practice, due to various human and real-life variables, impose conditions not previously accounted for; we have seen in this chapter applications of Libertarian Paternalism proposed either by its theoretical fathers or by government departments.

Having in mind these practical applications and the context of Libertarian Paternalism that was discussed in chapter 2, I shall now turn to the final chapter of this dissertation where I hope to critically present some of the implications that this novel form of paternalism carries for political and individual freedom.

---

<sup>250</sup> *Ibid.*; see also Rebonato (2012); Whitman, D. G. (2010). "The Rise of New Paternalism". *CATO Unbound: A Journal of Debate*, April Issue.

<sup>251</sup> Glaeser (2006)

## **CHAPTER 4:**

### **Libertarian Paternalism and Liberty**

In this chapter I will focus on exploring the different implications of how Libertarian Paternalism can affect individual liberty and autonomy as it was understood by Mill, whose ideas were presented earlier in chapter 1. Initially I will expose some underlying beliefs of Libertarian Paternalism that have deeply influenced their proposal.

Considering these influences, we will then see some of the problems that Libertarian Paternalism entails regarding liberty. Problems highlighted by the fact that their position assigns no weight to freedom beyond the narrow focus of choice preservation as opposed to the possibility of actual decision-making<sup>252</sup>.

#### **4.1 Existing problems with Libertarian Paternalism (in practice and in theory)**

##### *4.1.1 Interpreting the Empirical Evidence*

It would be wrong to assume that there are no limitations to the retrieval of empirical evidence. Laboratory behavioural research is not without flaws, and any retrieved conclusions might lack external validity, i.e. not translate into the outside world. As it stands, much of these behavioural experiments are done in academic environments and mostly with university students, making it plausible to assume that in some instances there can lack a concrete assessment of real behaviour, as it has been observed that some biases analysed in sterile settings do not actually survive exposure to the real world<sup>253</sup>.

Even natural experiments done on the field by use of randomised control trials or conducted through the use of questionnaires can be challenged by plausible

---

<sup>252</sup> Ginsburg and Wright (2012)

<sup>253</sup> *Ibid.*; see also Baddeley (2017); Oliver (2017)

arguments<sup>254</sup>. For example, randomised control tests done in real environments can be challenged as there is no way of showing, with certainty, that the different results between those two groups came from the treatment or from the diverse uncontrollable factors that are involved in influencing behaviour<sup>255</sup>. As such, we can expect for there to be some mistakes in interventions that are based on misinterpreted evidence. Notwithstanding, it would still be foolish to disregard all that has been studied in the different behavioural fields, but even accepting all that has come to light as unquestionable, one can still argue in favour of a different conclusions when faced with the research at hand.

In chapter 2, I briefly mentioned the difference between the two schools of thought that have originated from behavioural psychology; the bounded rationality school originating with Herbert Simon, and the behavioural economics school that despite being influenced by Simon is more easily traced to Kahneman and Tversky's work.

Both these schools maintain that people do not act according to neoclassical rational choice models due to departures from Bayesian reasoning and from the axioms of expected utility theory. The behavioural economics school looks at these departures as systematic "errors" that constrain our capacity to maximize our utility, thus making us choose "erroneously". The bounded-rationality school doesn't conclude in the same vein, but rather claims that the use of heuristics can still produce good results, perhaps even almost as good as if obeying neoclassical paradigms<sup>256</sup>. Thus, the distinction between both schools hinges on the different interpretations of the quality of the

---

<sup>254</sup> Baddeley (2017)

<sup>255</sup> *Ibid.*

<sup>256</sup> Rebonato (2012)

outcomes. Outcomes which, in the case of those subscribing to Kahneman and Tversky's view, justify the call for paternalistic interventions.

Notwithstanding, there is a second distinction inherent to both schools of thought: the role of evolutionary adaptiveness. Even though behavioural economists understand the role of human adaptation in the origin of our cognitive biases<sup>257</sup>, they dismiss it as unimportant. For the bounded rationality school, this adaptiveness characteristic is of great importance. Decision-making might sometimes be fast and automatic, consequently leading to less than optimal results, but that does not necessarily mean that the choice must be dismissed. In an evolutionary sense, those automatic and intuitive processes must be reasonable, in order to have survived.

Remembering the Linda problem in chapter 2, the bounded rationality view claims that there is rational inference in replying that Linda is most likely a “bank teller *and* active in the feminist movement” because of a “relevance maxim”<sup>258</sup>. When considering what Linda is most likely to be, people assume that all the information given is relevant to the problem because when faced with those types of questions the chooser's interpretation of “probability” departs from the mere mathematical standpoint. To answer only with what is mathematically more probable renders the characters' descriptions pointless and wasteful. As the audience assumes that communicators are meeting social standards of communication, they infer that all information is relevant to their alternative understanding of probability<sup>259</sup>. In that instance, it would be more than reasonable to ask why waste time describing someone's intrinsic characteristics if those

---

<sup>257</sup> Baddeley (2017); see also Kahneman (2012); Oliver (2017)

<sup>258</sup> Gigerenzer, G.; Hertwig, R (1999). The Conjunction Fallacy Revisited: How Intelligent Inferences Look Like Reasoning Errors. *Journal of Behavioral Decision Making* Vol. 12, No. 4: 275-305 referring to Paul Grice's *maxim of relation* p. 278; see also Gigerenzer, G. (2015). On the Supposed Evidence of Libertarian Paternalism. *Review of Philosophy and Psychology*, Vol. 6 No. 3:361-383; also, Poole (2014)

<sup>259</sup> *Ibid.*

are not even relevant to the problem? By answering the least likely option all the information becomes consistent and social standards are met<sup>260</sup>.

#### ***4.1.2 Notes on Rationality***

Even though behavioural sciences are, allegedly, shedding a new light on economic assumptions, it is plausible to argue that, they are merely providing empirical proof to sustain timeless assumptions about human behaviour. Considering different literary or academic works from various times, one can assume that decision makers have always been viewed as rationally bounded, albeit with the potential to overcome those limits. Whether in fiction, poetry, philosophy, etc. rationality was not understood as the only cause that motivated action. The two-selves that cognitive psychology is now theorising about, have been present long before the birth of that modern science. Even in economics Adam Smith recognised the “impartial spectator” and he speculated about what are now being called cognitive biases<sup>261</sup>.

Fiction authors have also described humans’ “predilections to err”, with what is now being called over-confidence, the endowment effect, the availability heuristic, limited attention, etc.<sup>262</sup> The so called *status quo bias* is almost always present within the story arc of the hero’s journey – where the hero refuses the call to adventure usually

---

<sup>260</sup> *Ibid.*

<sup>261</sup> Oliver (2013a) notes that in *Theory of Moral Sentiments* Adam Smith acknowledge our hyperbolic discounting or present bias: “pleasure which we are to enjoy ten years hence interests us so little in comparison with that which we may enjoy today”; or made reference to what is now known for loss aversion: “Pain is, in almost all cases, a more pungent sensation than the opposite and correspondent pleasure. The one almost always depresses us much more below the ordinary, or what may be called the natural state of our happiness, than the other ever raises us above it.” p. 690

<sup>262</sup> Oliver (2017)

because of engagements in his life or his fears<sup>263</sup>. Even different traditional and cultural sayings seem to convey our behavioural shortcomings<sup>264</sup>.

It stands to reason that behavioural scientists are not postulating anything new at all, most certainly not a different view of human nature, but they are rather providing empirical proof to what was already, speculatively, known by “grandmothers”<sup>265</sup> everywhere, taking it as a literal depiction of how the human brain works which in turn allows them to separate the individual in two<sup>266</sup>.

As we have seen, liberals like Mill or Berlin, don’t claim that the individual is incapable of making mistakes, but rather understand that government should not interfere even when facing citizen’s self-imposed harms (*internalities*) in the form of bad choices for our future selves. Mill in *On Liberty*, accounts for some of the irrational “desires and impulses” of human behaviour. However, he accepts them as an intrinsic part of Man, of “human nature”.

“(…) To a certain extent it is admitted, that our understanding should be our own: but there is not the same willingness to admit that our desires and impulses should be our own likewise; or that to possess impulses of our own, and of any strength, is anything but a peril and a snare. Yet desires and impulses are as much a part of a perfect human being, as beliefs and restraints: and strong impulses are only perilous when not properly balanced; when one set of aims and inclinations is developed into strength, while others, which ought to co-exist with them, remain weak and inactive.”<sup>267</sup>

---

<sup>263</sup> Sunstein (2016)

<sup>264</sup> Oliver (2017) Notes that idiomatic expressions like: “a bird in the hand is better than two in the bush” accounts for loss aversion; that “never put off until tomorrow what you can do today” recognises an inclination to procrastinate and so forth.

<sup>265</sup> Garvey (2016) states bluntly that “certainly human beings have been aware that a lot of our choices are less than rational – Pascal noticed that ‘the heart has reasons that reason doesn’t know’ long before Freud got in on the act. Do you really need a psychologist to tell you that peer pressure has a strong effect on us?” p. 85

<sup>266</sup> Rebonato (2012)

<sup>267</sup> Mill (1863) p. 115

So why do libertarian paternalists view these evidenced decisional blunders, speculated for centuries, as a justification for a new form of paternalism? The answer is threefold, and relates to rooted beliefs that influence their position: first, their definition of rationality; second, their view of the two selves (apparently a literal take on Dual Process Theory)<sup>268</sup>; and lastly, their concept of second-order desires, or what Sugden more appropriately calls, our latent preferences. Paternalism becomes justified because *humans*<sup>269</sup> are unable to freely achieve what their rational selves truly want without outside intervention, to act according to their deep values.

#### ***4.1.3 Libertarian Paternalists' definition of Rationality***

When contemplating what the libertarian paternalists define as rational choice one can use their extensive literature and conclude that in the case of rationality, they subscribe to the neoclassical postulate of *Homo Economicus*<sup>270</sup>.

“Research by psychologists and economists over the past three decades has raised questions about the rationality of the judgements and decisions that individuals make. People do not exhibit rational expectations, fail to make forecasts that are consistent with Bayes’ rule, use heuristics that lead them to make systematic blunders, exhibit preference reversals (...) and make different choices depending on the wording of the problem.”<sup>271</sup>

However, the rationality of the economic man is incredibly rigid, “any deviation from the one accepted way of choosing rationally is, for neoclassical economists (...) an error. (...) Irrational decisions are ‘mistakes’”<sup>272</sup>. Thus, by purposefully using such a restrictive definition of rationality they leave the field open to claim that any decision

---

<sup>268</sup> Rebonatto (2012); see also Sunstein (2014a)

<sup>269</sup> Here I allude to the distinction made in Sunstein and Thaler (2008) between *Humans* and *Econs*.

<sup>270</sup> Rebonatto (2012)

<sup>271</sup> Sunstein and Thaler (2003b) p. 176

<sup>272</sup> Rebonatto (2012) p. 38

made against this model is a mistake, consequently the individual is left worse off and they find justification for intervention, which allows me to turn now to another paternalistic belief, individual preferences.

#### *4.1.4 On Preferences*

For the libertarian paternalist preferences are context-dependent, this means that people's decisions change when faced with seemingly irrelevant aspects of a given context<sup>273</sup> (for example the display of food in a cafeteria). This preference inconsistency justifies paternalism, to the extent that people don't have "well-formed preferences"<sup>274</sup>, however it does not justify why the error-free latent preference should be favoured instead of the revealed one<sup>275</sup>.

Libertarian paternalists assume that choice architects should plan in order to make individuals better off "as judged by themselves"<sup>276</sup>. The implication here is not only that they assume to know what people's true preferences are, but that these preferences are revealed, not in observed choice, but in the decisions that individuals would make if they were completely rational.

"Drawing on some well-established finding we show that in many cases, individuals make pretty bad decisions – decisions that they would not have made if they had paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control"<sup>277</sup>

In essence, when Sunstein and Thaler call on paternalism to steer individuals in favour of their "true" preferences, they are creating a normative criterion for the choice

---

<sup>273</sup> Sunstein (1997; 2014); Sunstein and Thaler (2003a; 2003b); Sugden (2008; 2017)

<sup>274</sup> Sunstein and Thaler (2003a)

<sup>275</sup> Mitchell (2005); see also Rebonato (2012) and Sugden (2017)

<sup>276</sup> Sunstein and Thaler (2003a,2003b, 2009); Sunstein (2014a)

<sup>277</sup> Sunstein and Thaler (2009) pp. 5-6

architect to reconstruct people's "error-free *latent preferences* and recommend policies that satisfy those preferences"<sup>278</sup>. Libertarian Paternalism designs policies that respond to what the reflective self (System 2) "would actually want"<sup>279</sup>.

From a liberal standpoint the problem with this argument lies in that the planner, an objective external observer, is legitimizing their intervention because their goal is to satisfy *hypothetical* choices of an inner rational self, who interacts with the world through an error prone "psychological shell"<sup>280</sup>. Furthermore, by stipulating that the criteria for paternalism is people's welfare "as judged by themselves"<sup>281</sup>, the authors make their task harder, as it is impossible for an outside objective view to know what people's true judgements are. With this, they are conflating rationality with the promotion of welfare. However, rationality does not require, in any way, the promotion of a perceived objective welfare, for it's still true that one person might choose rationally, i.e. according to their "true" preferences, something which the objective observer might see as an "irrational choice"<sup>282</sup>.

Going back to Mill, who also assumes that, sometimes, there is a divergence between revealed desires and well-being, that revealed preferences should not be understood as evidence as to what is best for the individual, it is interesting to see how he became such an arduous defendant of liberty where the neo-paternalists did not. Even considering the lower desires of men, he still held that:

---

<sup>278</sup> Sugden (2017) p. 116

<sup>279</sup> Sunstein and Thaler (2009)

<sup>280</sup> Sugden (2017)

<sup>281</sup> Sunstein (2014a) "With respect to the chooser, let us understand the term [welfare] to refer to whatever choosers think would make their lives go well" p. 73

<sup>282</sup> Mitchell (2005), also Sugden (2017)

“his most fundamental moral commitment is to the ‘free development of individuality’ as ‘one of the leading essentials of well-being’. For Mill, each human being must achieve well-being in his own way, acting on his own judgement”<sup>283</sup>

In accepting emerging empirical evidence, libertarian paternalists are dismissing revealed preferences as a model to know people’s desires, claiming that individuals are ruled by cognitive biases and that their preferences are context-dependent. In doing so, they have challenged the “Epistemic Argument of the harm principle”, which asserts that: “because individuals know their tastes and situations better than officials do, they are in the best position to identify their own ends and the best means of obtaining them”<sup>284</sup>.

Furthermore, by including the promotion of subjective preferences as a guide to paternalism and by believing that those preferences go hand in hand with an objective measure of welfare, it can be argued that libertarian paternalists are still acting on grounds of what Mill calls “general presumptions”<sup>285</sup>.

Even accepting the idea of rational latent preferences, policymakers should never assume to know what those are, nor infer that because people are acting (obviously) against<sup>286</sup> their own well-being they are acting erroneously. Especially considering that some individuals choose rationally what others see as “irrational”<sup>287</sup>.

---

<sup>283</sup> Sugden (2006)

<sup>284</sup> Sunstein (2014a) p. 7

<sup>285</sup> Mill (1863) p. 147 “(...) He is the person most interested in his own well-being (...) while, with respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by anyone else. The interference of society to overrule his judgment and purposes in what only regards himself, must be grounded on general presumptions; which may be altogether wrong, and even if right, are as likely as not to be misapplied to individual cases, by persons no better acquainted with the circumstances of such cases than those are who look at them merely from without (...)”

<sup>286</sup> In Sunstein and Thaler (2008) they present the case of obesity in America as an example of how “true” preferences seem to be obstructed “(...) We do not claim that everyone who is overweight is necessarily failing to act rationally, but we do reject the claim that all or almost all Americans are choosing their diet optimally (...) With respect to diet, smoking, and drinking, people’s current choices cannot reasonably be claimed to be the best means of promoting their well-being. Indeed, many smokers, drinkers and overeaters are willing to pay third parties to help them make better

“The inevitability of policy errors derives from the insurmountable theoretical and empirical obstacles to identifying any one person’s, let alone the distribution of all persons’, “true preferences.” One type of policy error will occur when a behavioral intervention is aimed at seemingly irrational behavior that is in fact rational for the decision maker in question. In other words, the social costs of this type of policy error flow from encouraging behavior the paternalist inaccurately believes will make individuals better off and concomitantly discouraging acts that satisfy their actual preferences.”<sup>288</sup>

#### ***4.1.5 The Libertarian Paternalists’ Two-Selves Picture***

Regarding the ideas of rationality and preferences we find, within the libertarian paternalist literature, a constant tension between the two conflicting selves of Dual Process Theory. Going back to chapter 2 we remember that this theory presupposes the existence of two systems of thought that influence behaviour. On one side System 1 is automatic and impulsive, responsible for many of our behavioural shortcomings. On the other hand, System 2 is deliberate and reflective, the one which allows us to reason. Whereas in some authors these two systems are seen as parts of one whole interactive unit, the libertarian paternalists seem to take a literal interpretation to this dichotomy of thought processing in order to justify their right to intervene.

“They take this idea very seriously and literally, and use the ‘two-selves’ language as far more than a suggestive metaphor. Indeed so ‘real’ are the two selves supposed to be, that libertarian paternalists even adopt and modify in their writings the language of externalities (...) and speak of internalities”<sup>289</sup>

---

decisions” p. 8. By focusing on the outcome of the choice, against well-being, they assume it to be a mistake, even if the choice is done rationally.

<sup>287</sup> Sugden (2017) see also Mitchell (2005); Ginsburg and Wright (2012)

<sup>288</sup> Ginsburg and Wright (2012) p. 1052

<sup>289</sup> Rebonato (2012) p. 34

Under this light, through a small step in reasoning, they can justify paternalistic interventions in accordance to Mill's harm principle<sup>290</sup>. By introducing the concept of internalities, the paternalists are opening space to protect the rational System 2 from the hedonistic System 1. Here the problem stands that libertarian paternalists seem to clearly prioritize System 2's rationality without a justification beyond the fact that it's the rational-self.

“[T]he circularity of the argument is not immediately apparent because we tend to think that listening to reason is ‘obviously’ good. However, from the logical point of view, one could after all, just as validly have said: we should listen to an individual's first, unreflective reactions. (...) Logically nothing distinguishes the two criteria to give preference to one ‘System’ over the other”<sup>291</sup>

Thus, it becomes easy to see that even though libertarian paternalists are trying to objectively define people's true preferences, they themselves are falling prey to a subjective view of “whose” preference to choose. As they believe that System 2's choices are more competent to achieve what they deem as most important – welfare – they are deliberately picking the preferences they think are best, in favour of the System 2 self and to the detriment of the System 1 self<sup>292</sup>.

It is because libertarian paternalists see the System 1 self as lacking, that they reject the idea that our emotional and intuitive side can too, make instrumentally good decisions. By infantilizing one part of the individual, who is unable to choose “true” preferences because of this child-like involuntary System 1, their position takes a traditional paternalistic stance<sup>293</sup>.

---

<sup>290</sup> *Ibid.* also see Sunstein (2014a)

<sup>291</sup> Rebonato (2012) p. 166

<sup>292</sup> *Ibid.*

<sup>293</sup> *Ibid.* see also Mitchell (2004); also, Ginsburg and Wright (2012)

But the new-paternalists go further in their assessment of this dichotomy of the self. For them our cognitive biases and inadequacies of System 1 are so deeply rooted in our neurology that to the libertarian paternalists, they are taken as an unchangeable biological given. Apparently, we are biologically determined as imperfect thinkers<sup>294</sup>. Our cognitive biases and rationality are part of our human nature and cannot be overcome. Sunstein goes further in stating that this expands beyond the abstract:

“identifiable regions of the brain are active in different task and hence it may well be right to suggest that the idea of systems has physical referent (...) ‘automatic and controlled processes can be roughly distinguished by where they occur in the brain’”<sup>295</sup>

The evidence that Sunstein is referring to has to be looked at very carefully. Despite some evidence showing that different regions of the brain are in fact related to automatic processes whereas others are more active with reflective processes<sup>296</sup>, this does not mean that those regions and functions are necessarily stagnant. Concepts of adaptability and plasticity must also be considered. Moreover, there are no studies, to this point, that show with certainty that Stanovich and West’s two systems of thought are entirely separated or independent from one another<sup>297</sup>.

---

<sup>294</sup> Rebonato (2012)

<sup>295</sup> Sunstein (2014a) p.27

<sup>296</sup> *Ibid.*

<sup>297</sup> Garvey (2016); see also Kahneman (2012); Oliver (2017); Rebonato (2012)

## 4.2 Problems Arising with Libertarian Paternalism in relation to political liberty and decisional autonomy

### 4.2.1 Nominal Freedom or Real Freedom

When Sunstein compares *nudges* to GPS systems<sup>298</sup> he is either inferring that we have an *a priori* alternative to being nudged or that we use a GPS system because, rationally, what we truly desire is to reach a given destination as fast and economically as possible (our welfare), without ever going astray. Against the latter let's recall the brief discussion in section 4.1.4; libertarian paternalists claim to know our latent subjective preferences by considering an objective measure of welfare. They further claim that individuals desire (truly) to be nudged in order to overcome what policymakers see as errors that hinder our good decision-making<sup>299</sup>. However, as has been argued, it is impossible for an outsider to know what people truly want, or to assume an action as irrational based on the quality of its outcome.

Considering the first inference and accepting it as one of the main characteristics of *nudges* - their easy-reversibility - we cannot compare the voluntary choice to use an external aid<sup>300</sup> (the GPS) as the same as accepting an already implemented intervention without our stipulated authorization.

This means that there is liberty to use a GPS and the capacity to voluntarily employ its use, but by comparison there's no choice to being "nudged", especially in the case of default rules, there is rather the possibility to actively choose *to not* be 'nudged'. The difference might seem small, but it is highly pertinent. The fact that libertarian

---

<sup>298</sup> Sunstein (2014a) p. 133

<sup>299</sup> Sugden (2017)

<sup>300</sup> Sugden, R. (2018). Better off, "as judged by themselves": A Reply to Cass Sunstein. *International Review of Economics*, Vol. 65: 9-13; against the use of the iconic GPS Sugden claims that if we consider all voluntarily hired expertise, such as a GPS, as paternalistic, paternalism ceases to be objectionable and loses most of its meaning.

paternalists assign no particular weight to liberty beyond the narrow focus of choice preservation<sup>301</sup> is illustrative.

As Sunstein (2014a) counter argues some of Mill's ideas trying to justify his position, it becomes important to understand that for Mill freedom is not only associated with alternative of choice but also with dignity, autonomy, individuality and self-development<sup>302</sup>.

“[T]o conform to custom, merely as custom, does not educate or develop in him any of the qualities which are the distinctive endowment of a human being. The human faculties of perception, judgement, discriminative feeling, mental activity, and even moral preference are exercised only in making a choice. He who does anything because it is the custom, makes no choice. (...) The mental and moral, like the muscular powers, are improved only by being used. (...). He who lets the world, or his own portion of it, choose his plan of life for him, has no need of any other faculty than the ape-like one of imitation. He who chooses his plan for himself, employs all his faculties. (...) It is possible that he might be guided in some good path, and kept out of harm's way, without any of these things. But what will be his comparative worth as a human being?”<sup>303</sup>

If we consider that not actively joining a plan is a non-voluntary action and that all actions should be presumed voluntary, except in the face of coercion, then it stands to reason that Libertarian Paternalism is traditionally paternalistic. And, even though their proposals tend to fall on soft forms of paternalism, i.e. non-forbidding and freedom preserving, it is possible to analyse how some of the proposed measures pose real threats to liberty. Perhaps as much as harder, more restrictive, policies entail.

The case of organ donation in Europe is a case in point. There is an incredible gap between countries that have a policy of presumed consent – citizens are automatic

---

<sup>301</sup> Ginsburg and Wright (2012)

<sup>302</sup> *Ibid.*; see also Donner (2010); Skorupski (1998); Sugden (2006)

<sup>303</sup> Mill (1863) pp. 112-113

donors – and those countries who have a policy of active choosing – people need to voluntarily register as a donor. Countries such as Austria, Sweden, France, Portugal, Poland and Hungary have a consent rate of donation from a minimum of 85.9% (Sweden) to a maximum of 99.98% (Austria). Other nations such as Germany, Denmark, The Netherlands and The United Kingdom display organ donation rates ranging from 4.25% to 27.5%<sup>304</sup>.

Trying to understand this numerical difference by contrasting cultural values and societal norms can only be counterproductive. Realizing that it is probably as easy to sign up for organ donation in Germany as it is to *opt out* in Austria one must ask “what is the value of the reversibility of choice set up by libertarian paternalists if freedom of choice is not exercised”<sup>305</sup>, that is, if the nudges of the libertarian paternalists are so effective. A citizen in Austria, while nominally free to choose to reverse the default, does not do so.

Libertarian paternalists face a conundrum; the more effective their policies are, the more “real” freedom of choice seems to become non-existent. This means that when the effectiveness of the nudge reaches almost 100% (as in the case of organ donation in Austria, and the SMT scheme with automatic enrolment), the easy-reversibility characteristic becomes merely abstract. Thus, “real” liberty ceases to be preserved. A default that is highly effective does not let us know how many Austrians do actually want to be organ donors and how many would actively choose it. Going back to Berlin we can remember that:

“[‘Negative liberty’] might, *prima facie*, seem to depend simply on the power to choose between at any rate two alternatives. Nevertheless, not all choices are equally free, or free at all. The mere existence of alternatives is not, therefore,

---

<sup>304</sup> Rebonato (2012)

<sup>305</sup> Rebonato (2012) p. 133

enough to make my action free (although it may be voluntary) in the normal sense of the word.”<sup>306</sup>

When faced with the effectiveness of some of the libertarian paternalistic policies, and if many other policies proposed are as effective as they claim and want them to be, then this soft paternalism is, in some cases, almost as restrictive as traditional harder forms of paternalism, if not, equally restrictive. It is here that libertarian paternalism becomes blatantly a contradiction on itself.

#### ***4.2.2 The Importance of Experience and Learning***

By proposing measures that limit and constrict decisions, libertarian paternalists dismiss the importance of experience in the process of reaching the full potential of the individual. From a liberal point of view, unencumbered actions are the best source for self-development. Libertarian Paternalism here ignores the value of the freedom to err<sup>307</sup>. People learn as much from mistakes as they do from successes and consequently grow as an individual from either. Government interference automatically limits and restricts this process of self-development, not only, because it assumes to know better, but also because it obstructs our decisional autonomy.

“In many cases, though individuals may not do the particular thing so well, on the average, as the officers of government, it is nevertheless desirable that it should be done by them, rather than by the government, as a means to their own mental education—a mode of strengthening their active faculties, exercising their judgement, and giving them a familiar knowledge of the subjects with which they are thus left to deal.”<sup>308</sup>

---

<sup>306</sup> Berlin (2002b) p. 177

<sup>307</sup> Ginsburg and Wright (2012)

<sup>308</sup> Mill (1863) pp. 211-212

Decisional autonomy exists when we are conscious of our own decisions, and when we are capable of acting on our own, by our own reflection. When a government sets out rules under principles of behavioural sciences, that more often than not we are unaware of, it becomes impossible to claim those decisions as our own, or to revel in the fruits of the experience of choosing and the responsibility that comes with the choice.

John Stuart Mill admits that even if, as outsiders, we were confident that we knew which way of life would be best for the choosers, their well-being would still be better promoted through themselves experiencing mistakes and learning from them.<sup>309</sup> He is not claiming that we can't make judgements on others choices and well-being, but rather that we shouldn't allow for those judgements to justify restricting freedom of choice<sup>310</sup>. Mill feels that in doing so would restrain their ability to learn from the experiments that lead to the fulfilment of the individual's self-realisation.<sup>311</sup>

In the case of Libertarian Paternalism this interference becomes even more problematic because it specifically focusses on targeting our automatic System 1, designing policies that circumvent our reflective selves. Thus, further restricting our capacity to critically analyse our decisions and consequently improve our “mental education”.

Furthermore, let us analyse the libertarian paternalists' upheld view of our biological determinism, as it becomes relevant to the discussion. In his book *Why Nudge*, Sunstein presents some evidence to show how the brain seems to be functionally divided even though this evidence is not as straightforward and simple as he implies. In Rebonato we are introduced to concepts of plasticity, where the brain recovers lost

---

<sup>309</sup> Sugden (2006)

<sup>310</sup> *Ibid*

<sup>311</sup> Mill (1863)

functions by rewiring (creating new pathways) instead of reactivating damaged areas; and the related concept of adaptability where the mental processes can adapt to the environmental demands of a lifetime. Under these concepts the human brain cannot be considered as made of a pre-determined and stagnant binary system. Rather Mill's speculation becomes empirically much more pertinent: "the mental and moral, like the muscular powers, are improved only by being used"<sup>312</sup>

Assuming that it is possible to change the settings of our System 1 and improve one's ability to reason, it becomes incredibly relevant to devote more time to the task. In doing so, nudges would cease to be necessary, people would be able to achieve their own true preferences and most importantly individuals would receive greater respect and dignity for the integrity and responsibility of their decisional autonomy. Alas, libertarian paternalists "tell us that experience, expertise, and incentives cannot be counted on to protect decision-makers from the influence of the choice-frame"<sup>313</sup>

To Sunstein and Thaler the deterministic view of the two systems of thought actually moves them further away from the libertarian claim to their paternalism. Emerging evidence shows that it *is* possible to be less susceptible to decision manipulation through the use of *debiasing techniques*<sup>314</sup>, but the idea of overcoming cognitive bias is absent from the libertarian paternalistic thesis.

Different tests have shown successful outcomes for overcoming cognitive biases and for expanding the reasoning of System 2. By dismissing the role of experience and learning, libertarian paternalists are further hindering our liberties and, therefore, our potential for self-development. An extended quote by Rebonato is appropriate to

---

<sup>312</sup> *Ibid.* p.113

<sup>313</sup> Mitchell (2004) pp. 8-9

<sup>314</sup> *Ibid.*

conclude how the dismissal of experience and of learning by oneself has serious implications for liberty:

“It is therefore reasonable to conclude that individuals can learn to become System-II thinkers, or at least to apply System-II thinking to more areas of their decision-making than libertarian paternalists seem to assume. (...) studies also suggest that *not* encouraging individual to make conscious, deliberate, well-reasoned choices can reinforce the System I (...) [T]he more decisions are reached by, say, exploiting a lack of understanding of the effect of inflation (...), or by relying on the intent acceptance of a default option, the less the neural pathways associated with our reflective mode of thinking are activated – and the more our brains will *become* ill-equipped to reason rationally and critically about the next choice. (...) It is in this sense that the cognitive deficiencies of individuals can truly become, by their systematic exploitation, a self-fulfilling prophecy”<sup>315</sup>

The fact is that Libertarian Paternalists seem to dismiss completely an educational side to their libertarian view, to the extent that there is no interest in helping people overcome problems in reasoning in order to become more capable and free to choose in the future, “(...) to create conditions in which those who lack them will be provided with opportunities to exercise those rights (freedom to choose) which they legally possess, but cannot, without such opportunities, put to use (...)”<sup>316</sup>. As such it seems obvious that the preference Sunstein and Thaler show in using our unreflective selves to make their policies more effective, makes their position less liberal than they claim.

#### ***4.2.3 Transparency and Salience***

In order to employ democratic safeguards over abusive and restrictive governments, policies should be transparent and easily monitored by citizens. Hard

---

<sup>315</sup> Rebonato (2012) p. 219

<sup>316</sup> Berlin (2002a) p. 46

paternalistic measures tend to be self-evident, as they take the form of taxes, bans, penalties, etc. As such, these types of policies are constantly subjected to the scrutiny of individuals. By being less transparent citizens might find more difficult to monitor such policies, which in turn can create two problems. First, less accountability from government by hiding behind psychological mechanisms and consequently less control over government by civil society. Secondly, people might succumb to nudges more easily if they are unaware of them.

Nudge interventions, in the form of soft paternalism, are employed in ways that exploit our impulsive and automatic thought-processes. These measures are not often advertised by governments before or after implementation, and when they are, the framing of the disclosure is left to the planner's endeavours. As such, these interventions tend to fly under the radar altogether<sup>317</sup>.

Sunstein for his part states that the problem with soft paternalism is not the lack of transparency but rather of salience. When applying soft measures such as nudges, they are visible, public and observable, the difference lies in that they do not attract attention and so do not suffer public discussion.<sup>318</sup> However, he admits that for the most part nudges don't promote scrutiny, they are not done in secret, but operate by influencing or appealing to System 1, and he invokes justifiability as a defence to the accusation of lack of transparency.

The libertarian paternalist argument is weak, at best. They allude, to the publicity principle that "bans governments from selecting a policy that it would not be able or willing to defend publicly to its own citizens"<sup>319</sup> Nonetheless, it stands to reason, that the prospect of a future need of justification for when a policy comes to light, i.e.

---

<sup>317</sup> Glaeser (2006)

<sup>318</sup> Sunstein (2014a)

<sup>319</sup> Sunstein and Thaler (2009) p. 244

becomes transparent, hardly accounts for the policy's actual transparency<sup>320</sup>. This principle, as presented in Sunstein and Thaler, does not imply that a given policy is visible, but rather that it can become visible. They consider the matter of future transparency as a defence and use the image of a benevolent ruler as its justification.

Nudge interventions exploit cognitive blunders that people are unaware in order to guide towards more sensible decision-making. Thus, it may happen that a nudge will only work with a certain degree of covertness, appealing to people's unconscious and possibly becoming unavoidable, restricting the freedom to ignore it.<sup>321</sup> The lack of its transparency now becomes more challenging to the libertarian position to the extent that, not only, is the public scrutiny of government hindered but also, because of its covert characteristics individuals are more susceptible to the effectiveness of the nudge.<sup>322</sup>

As civil society's capabilities of recognising and discerning such policies becomes somewhat hindered by its lack of transparency and salience, the risk of the increasing state control and manipulation becomes a conceivable reality as limits imposed on government by civil society decrease.

#### ***4.2.4 The Slippery Slope of Libertarian Paternalism***

So far, we have seen how behavioural sciences have identified different cognitive biases. Libertarian Paternalism designs policies that targets these biases in order to make them more effective. As such, it's only to be expected that they should also use their knowledge of behavioural insights to promote their thesis<sup>323</sup>. The concept

---

<sup>320</sup> Rebonato (2012)

<sup>321</sup> Oliver (2013a), "if people are meant to face the changes in the choice architecture *unconsciously* – e.g., unconsciously face fruit rather than cheesecake at eye level as they are about to pay for their lunch in their local canteen – how can they make the conscious decision of non-participation?" p. 3

<sup>322</sup> *Ibid.*

<sup>323</sup> Whitman (2010)

of extreme-aversion claims that people tend to stay clear of the extreme choices – when framed between opposites people tend to choose the middle ground. Libertarian Paternalists have presented their movement as “the Third-way”, a moderate middle ground between libertarianism and hard paternalism. Nevertheless, as Whitman points out, the middle ground position tends to shift directions over time. Once a regulation is in place it can become the *status quo*, the accepted new norm, and as this middle ground changes it can demand other forms of interventions, perhaps more restrictive in nature.

Tobacco regulation is one of the most illustrative cases<sup>324</sup>. Banning smoking on airplanes, on train compartments and in school classes was once the middle ground position, nowadays some cities are considering banning the act in some outdoor areas<sup>325</sup>. All these changes in regulations tend to come about small but increasingly paternalistic policies that over time shift, an initial compromise becomes the accepted norm.

Phenomenon, such as framing or the small-change tolerance - “a willingness to tolerate changes perceived as relatively small movements from the *status quo*”<sup>326</sup> - can be used, to one’s own advantage, to gradually expand from soft paternalistic measures to hard ones. Of course, recalling the discussion in the previous section, we can understand that from a democratic standpoint harder forms of paternalism can be more attractive to individuals as they are easier to track, more visible to scrutinize and consequently more attractive to a liberal position as it can lead to a more effective monitorization of government.

By comparison, from a public official’s standpoint we can understand an attraction to soft paternalism as hard paternalism demands scrutiny and deliberation,

---

<sup>324</sup> For a very brief history of the different tobacco regulations see Ginsburg and Wright (2012) and Glaeser (2006)

<sup>325</sup> Ginsburg and Wright (2012) alluding to New York City Health Commissioner’s proposal to ban smoking at city parks and beaches.

<sup>326</sup> Voloch (2003) *The Mechanism of the Slippery Slope*. quoted in Rebonato (2012) p. 120

public officials when devising a paternalistic policy of this kind risk public controversy and alienating the electorate<sup>327</sup>. Thus, soft measures that are not as salient, and therefore, don't risk losing voters support might appeal to the policymaker. The more effective the measure turns out to be (as in the cases of defaults) the more likely it will be for the public official to try and expand the scope of soft paternalism.

With that expansion lies a much more worrying slippery slope critique: not the possibility of shifting from a soft-to-hard paternalism, but rather the possibility of becoming, not only, less overt, but more manipulative and also gradually more restrictive. When considering the use of more intrusive subliminal interventions, Sunstein and Thaler's position is very tentative, they do not say they agree, but rather that they "tend to think" that it is objectionable due to lack of monitorability<sup>328</sup>. As libertarian paternalists do not present any coherent acceptability line to their programme, the slippery slope is open to less visible and more intrusive interventions.

"The less overt these modes of influencing are, the less an individual is in a position of reversing the 'default option'; the greater the risks in wrongly guessing what the individual really wants (...); the smaller, in general, the degree of monitorability of the whole process. In short, the slippery slope I am worried about does not lead from the reversal of default options to stiff fine and prison sentences, but to subliminal advertising."<sup>329</sup>

As the constant risk of being on a slippery slope advances, the more real Mill and Berlin's fear becomes: the gradually increasing limitless state. Furthermore, this fear is deepened due to the behavioural aspect of these new-paternalisms, as it invites governments to focus on our cognitive mistakes and our subconscious cues for their own advantage. As policies of this kind risk becoming less and less visible to public

---

<sup>327</sup> Glaeser (2006)

<sup>328</sup> Sunstein and Thaler (2009) p. 246

<sup>329</sup> Rebonato (2012) p. 131

scrutiny and more effective in attaining governments' objectives; the more likely it becomes that the "nominal freedom" of these policies ceases to be relevant as most humans will not act in favour of it, which in turn might end up atrophying<sup>330</sup> our critical thinking.

As a side note, we should admire libertarian paternalists response to the slippery slope critique. Their stance is to dismiss the abstract discussion<sup>331</sup> on grounds of their policies' results, and because they implicitly trust a benevolent government that is limited by a publicity principle that only intends to design wise, "worthwhile" policies.

"If our policies are unwise, then it would be constructive to criticize them directly rather than to rely on the fear of a hypothetical slippery slope. And if our proposals are worthwhile, then let's make progress on those, and do whatever it takes to pour sand on the slope (assuming that we really are worried about how slippery it is)"<sup>332</sup>

#### **4.3 Libertarian Paternalism. Is it libertarian? Is it Paternalistic? Or is it, simply put, Welfarist?**

We have seen how diverse libertarian paternalistic interventions can be, their policies range from tackling savings, food-habits to organ donation; they shift from individuals health to climate change and employ different means to achieve their policies, some, admittedly, more intrusive to people's decisional autonomy than others.

In some ways they are paternalistic because some of their measures clearly tackle their concept of *internalities*, understood as harms to oneself. However, this can be more ambiguous than at first sight, as Libertarian Paternalism focusses immensely on their intended outcomes, it seems to shift from a means to ends type of paternalism

---

<sup>330</sup> *Ibid.*; see also Ginsburg and Wright (2012); Poole (2014)

<sup>331</sup> Sunstein and Thaler (2009)

<sup>332</sup> *Ibid.* p. 236

more often than not. As paternalism is defined by a matter of intention, where there is a purpose to guide people towards decisions that increase their well-being by ameliorating *internalities*. On the other hand, one can interpret these paternalistic policies as deriving from a social welfare perspective, not centred on the individual and to the benefit of government.

Tackling smoke addiction or unhealthy food habits benefits the government in reducing investments on obesity related diseases or lung cancer treatments<sup>333</sup>. In turn this can allow for the redistribution of the National Health Service (NHS) means to either focus on other NHS treatments, on other government departments, or even towards tax reductions<sup>334</sup>. Going back to chapter 3, this was evident to the authors of the Mindspace Report who implied, more than once, that the proposed measures were beneficial, not from a paternalistic position, but in order to improve welfare on the whole, in this case justifying the intervention in order to increase public health and save public resources by reducing the burden on the NHS<sup>335</sup>.

In the end these lines become blurry as, from an outside view, no citizen can claim to know the real intentions of public officials. The libertarian paternalists assumption of a state governed by benevolent planners becomes dangerous, as the intentions are never clear and as political leaders have a number of goals that go beyond that of improving individuals' well-being<sup>336</sup>. When faced with these policies the intention can become ambiguous. Where one might see a paternalistic measure that benefits an individual's well-being another might see a measure that serves governments' self-interest, hidden behind justifications of soft paternalism, which is less salient and transparent to civil society.

---

<sup>333</sup> Oliver (2013a, 2017)

<sup>334</sup> Rebonato (2012)

<sup>335</sup> Dolan et al. (2010); also in Oliver (2017)

<sup>336</sup> Glaeser (2006)

On another, more straightforward hand, some of these policy proposals clearly focus outside the scope of negative *externalities*, but rather on a more generalized idea of society's welfare. The most evident case being that of organ donation, as the donor, being declared dead, will never benefit from the increased availability of organs that libertarian paternalists are trying to reach. Even, the caricature nudge of the Schipol housefly mentioned in chapter 3 and often quoted by Sunstein and Thaler, is welfarist in kind, as it does not benefit the individual directly, but rather other users - the facility's management, the cleaning department, etc. But many other of the proposed initiatives that went unmentioned focus on tackling climate change. Given it is mostly future citizens and societies that will benefit from these "green" interventions which impose costs on present consumption these measures cannot be considered paternalistic<sup>337</sup>.

These specific cases are clear-cut examples of a non-paternalistic position as they don't try to target self-behaviour, but rather the best outcome for society, even at the expense of some, preferably few, individual's preferences. If these policies' aim at benefiting society at large, the libertarian paternalists must have a pre-existing idea of what is good for society and how to aggregate the welfare of all citizens. Ultimately, in assuming to objectively know individuals' real preferences, in emphasizing the benevolence of the government rulers, in prioritizing welfare above any other value, it becomes plausible to assume Libertarian Paternalism as somewhat welfarist in nature, especially when we realize that Libertarian Paternalism defines welfare as its objective norm and proposes to influence and guide individuals' behaviour in regards to that precise norm<sup>338</sup>, not only to improve individual's well-being, as they see it, but also to increase society's welfare.

---

<sup>337</sup> *Ibid.*

<sup>338</sup> Mitchell (2005); see also Sugden (2008; 2017)

## *CONCLUSIONS*

For the past five decades, developments in the different behavioural fields have challenged established models of rational theory. By demonstrating how people decide in certain situations, these sciences have contributed to the belief that people, often, make systematic mistakes, and as such, don't always act of their own self-interest.

Theories of bounded rationality, that claim our rational actions to be limited by either inner constraints or external constraints, were put forth; the role of cognitive biases and heuristics in our behaviour has been thoroughly studied by psychology, culminating in extensive literature that detail hundreds of these automatic responses which, allegedly, tend to hinder our rational actions; Dual Process Theory that postulates an automatic and impulsive system of thought and another of reflective and deliberate action has been hypothesized, scientifically explaining the long held belief that we have at least two “personalities” that directly influence our choices.

All these insights and many other inputs from different fields of psychology have opened a greater space for paternalism, appealing to fields of law, economics and policymaking<sup>339</sup>. Insofar as paternalism means for citizens to not act fully voluntarily, these cognitive biases and limits to our rationality justify novel forms of paternalism as they consider our rational actions to be hindered by our infantilized impulsive selves. The libertarian paternalistic intention is evident; “bad choices” (as judged by public officials) demand external interference so people can be guided towards what they themselves consider “right choices” (from the external objective point of view of the public official).

---

<sup>339</sup> Glaeser (2006); also, recalling Rachlinski's quote in see also Mitchell (2004), p. 4: “Virtually every scholar who has written on the application of psychological research on judgment and choice to law has concluded that cognitive psychology supports institutional constraint on individual choice.”

Because of its wide audience, due in part to the success of Sunstein and Thaler's book of 2008 - *Nudge*, Libertarian Paternalism has become the most well-known thesis of these emerging "new" forms of paternalism.<sup>340</sup> They have presented themselves as the real Third Way, the middle ground between the anti-paternalist and the paternalist<sup>341</sup>.

We have seen that libertarian paternalists claim their position to be libertarian because, to their eyes, their measures do not encroach on people's liberties. At the same time, they describe themselves as paternalistic because their main intention is to guide people towards their own well-being<sup>342</sup>. In chapter 2, I have exposed how libertarian paternalists justify themselves under the belief that anti-paternalists share one false assumption and two misconceptions<sup>343</sup>. They highlight the anti-paternalists' assumption, that individuals are the best arbiters of their own lives and, as such, always act according to their own self-interest and claim it to be proven false by all the developments in the different behavioural fields. However, here Sunstein and Thaler are refusing to accept that liberal theorists don't consider people to be infallible, for liberals like Mill and Berlin to make mistakes is, on itself, beneficial.

By assuming as justification for their paternalistic stance that individuals obviously don't act in their best interest and that an external expert can be better equipped to choose for them, they go against the anti-paternalistic objections of Mill and Berlin. However, the fact that men do not always act without harming themselves, do not always choose the most preferable choices and can be prone to making mistakes

---

<sup>340</sup> Asymmetric Paternalism proposed by Camerer *et al.* (2003), briefly discussed in chapter 2, has also influence the present debate however this proposal and others are not as widely known outside the realm of academia

<sup>341</sup> Sunstein and Thaler (2009)

<sup>342</sup> Sunstein and Thaler (2003a)

<sup>343</sup> *Ibid.*

is acknowledged by both these authors. Our impulsive selves are intrinsic to human nature, not separated from our rational-selves.

“It is not because men's desires are strong that they act ill; it is because their consciences are weak. There is no natural connection between strong impulses and a weak conscience. The natural connection is the other way. (...) Strong impulses are but another name for energy. Energy may be turned to bad uses; but more good may always be made of an energetic nature, than of an indolent and impassive one.”<sup>344</sup>

Yet, despite this human nature, Berlin and especially Mill believed that even though it is possible to make external judgements over others' decisions, there lies no necessity to interfere with individuals' decisions in accordance to those judgements, people have the right to err.

As we have seen in chapter 1, so long as people don't act outside of what agreed laws dictate, infringing on other people's rights, there is no need for outside intervention in a person's individual life. This will only hinder individuals' freedom to act, to learn, to develop and to take away the responsibility for their choices.

Libertarian Paternalism has also created two alleged misconceptions for the anti-paternalists: that there exists a viable way to paternalism and that paternalism always involves coercion. Regarding the former, Sunstein and Thaler argue that there is always some form of external influence on behaviour, as behaviour tends to be context-dependent.

For the new paternalists behavioural insights have shown that most of our preferences are context-dependent and as such they argue that preferences do “not

---

<sup>344</sup> Mill (1863) p. 115

formally exist<sup>345</sup>. They claim that if every detail matters, and because we don't live without context, a form of influence is always present and as such it ought to be benevolent, it ought to promote individual welfare. Furthermore, in the literature libertarian paternalists do not provide a concrete and coherent reason to why a less benevolent alternative (such as prioritizing profit) is less viable when having the power of choice-architecture (the design of contexts)<sup>346</sup>. But it is because of this benevolent streak that welfare - "whatever choosers would think make their lives go well"<sup>347</sup> - becomes the normative value to their policymaking and the only viable option.

From here the authors conclude that a proper alternative to some kind of paternalism - "at least in the form of an intervention that affects what people choose"<sup>348</sup> - does not exist. However, we have seen that this argument is far-fetched as paternalism cannot be defined as any intervention, but rather a purposeful one with the *intention* to change behaviour in a supposedly better direction.

With this thought in mind, libertarian paternalists accept, on one side, preferences to be subjective - "as judged by themselves" - and on the other side, dismiss that our automatic revealed preferences can actually be our "true" preferences. Ultimately, they put themselves in the difficult position of objectively stipulating what people "truly" desire,<sup>349</sup> dismissing the subjectivity and diversity of individuals' preferences, even our rational ones, and consequently dismissing Mill's ardent defence for experiencing and individuality.

---

<sup>345</sup> Sunstein and Thaler (2003a)

<sup>346</sup> Sugden (2008; 2017)

<sup>347</sup> Sunstein (2014a)

<sup>348</sup> *Ibid.* p.5, emphasis added

<sup>349</sup> Sugden (2017; 2018)

“If a person possesses any tolerable amount of common-sense and experience, his own mode of laying out his existence is the best, not because it is the best in itself, but because it is his own mode.”<sup>350</sup>

Regarding the second misconception – that paternalism always involves coercion - it can be concluded that indeed soft forms of paternalism might not involve extensive coercive measures. However, in this case, it seems more pertinent to ask whether if by not being coercive, paternalism can still have no implications for liberty and thus be called truly libertarian?

It has not been my intention with this to continue the never-ending debate between the liberal and the paternalistic positions. I haven't felt obliged to discuss the importance of welfare in society or to better define Libertarian Paternalism (as it can be argued on some grounds that it is neither libertarian nor paternalistic). My purpose was rather, to challenge the claims that Libertarian Paternalism has (1) no implications for liberty and decisional autonomy because their policies preserve freedom of choice; (2) that anti-paternalistic objections are unwarranted because people have now been proven to not always act in their best interest and because paternalism doesn't need to be coercive; and (3) because choices are context-dependent, soft paternalism is the only viable option available<sup>351</sup>.

Despite these paternalistic assumptions not being entirely new, they use evidence of relatively new fields of research to support their thesis. Because of this, the idea to “take advantage” of behavioural empirical evidence in order to attain government's' goals, (which by libertarian paternalists' standards should be welfare), deserves the same scrutiny and reservations as seen in Mill and Berlin.

---

<sup>350</sup> Mill (1863) p. 130

<sup>351</sup> Sunstein and Thaler (2003a; 2003b; 2009) Sunstein (2014a): the anti-paternalists' false assumption and two misconceptions.

Libertarian paternalists are blatantly inviting governments to use behavioural insights to their advantage in order to design more effective policies, thus giving more power of manipulation to the state<sup>352</sup>. As we have seen, these types of policies tend to lack transparency, try to use cognitive biases and predictable behaviour to their own gains, and mostly target our automatic cues. Thus, the tension that should exist between governments and their civil societies becomes more relaxed.

Ultimately, it seems that Libertarian Paternalism is calling for greater state influence over people's individual lives without setting proper limits in place against the abuse of government practices. As measures of this type are less transparent, the risk of an all-encompassing state that disrespects individualism and that dismisses the benefits of personal autonomous experiences<sup>353</sup>, becomes more ever-present.

Libertarian paternalists fail here to understand the benefits of decisional autonomy for the good of society, and because they use a very narrow concept of negative liberty – the mere preservation of choice – they seem unable to grasp that the more effective their policies become, the less “real” freedom to voluntarily choose what people truly prefer will exist<sup>354</sup>. But more importantly, libertarian paternalists are unaware that governments might use these types of policies in order to gain more power by exploiting the lack of transparency and by developing more manipulative and intrusive policies<sup>355</sup>. As they, themselves, only reluctantly refuse the idea of using subliminal means to achieve their goals because those types of measures are more invisible and impossible to monitor, it is not inconceivable to consider that a government might feel the attractiveness of such policies.

---

<sup>352</sup> Rebonato (2012)

<sup>353</sup> Rebonato (2012)

<sup>354</sup> *Ibid.*

<sup>355</sup> *Ibid.*; also Glaeser (2006)

Despite all its risks and possible implications, it must be concluded that so far “nudging” is hardly a dystopian tool of state control, it is ruled on grounds of preserving choice of action, it tends to value freedom of choice and many of the initiatives proposed are relatively less intrusive than harder measures. It is also easy to understand why their proposed measures are attractive, as it becomes difficult to oppose initiatives that intend to raise organ donations or to tackle obesity and climate change.

We can accept that libertarian paternalists are indeed liberal in kind, or at the very least more liberal than hard-core paternalists. They value freedom of choice and their measures tend to be guided by soft forms of paternalism, but this does not mean on itself that their proposal and underlying assumptions have no implications for political and individual liberty. We have seen in chapter 3 how even some of the less intrusive measures, such as merely providing information for students to enrol at university, diminishes the importance of decisional autonomy and gives power to the state to use psychological mechanisms to comply with certain government objectives. Thus, when considering government goals and party intentions it must be understood that Libertarian Paternalism in theory does not necessarily account for certain variables in its practical application.

But the fact that they argue against the risk of a slippery slope on grounds of benevolent rulers and because of the assumption of justifiability as their limit to government action – where government is bounded only according to their ability to justify their practices to citizens – makes one believe that libertarian paternalists have a deep-rooted trust in government officials to do only that which is right towards their citizens. At the same time, libertarian paternalists are those presenting an interpretation of empirical evidence that allows governments to look at citizens as not fully competent thinkers. If this is true, libertarian paternalists become those enabling governments to,

justifiably, disrespect citizens reflective autonomy, contributing to potential further abuses of trust and the dismissal of public discussion.<sup>356</sup>

Even if a group of persons supports a given goal for themselves, even if the government is able to justify their measures to citizens, people might still not support covert initiatives intended to guide them to that given goal. From a liberal point of view, the end in itself is important but *how* it is reached is too.<sup>357</sup> As Sunstein noted, impersonal default rules are only acceptable by the chooser when there is trust in the planner<sup>358</sup>. Nudge policies are also better supported when targeting more technical decisions, where people lack the expertise to make a properly informed choice (as in the case of a savings rate for retirement). This support seems to go in the opposite direction when nudges are targeting different types of decisions (such as organ donation)<sup>359</sup> or individual choices (such as the choice to eat fattening food)<sup>360</sup>.

From a policymaking point of view, it would be wrong to dismiss all the developments in the behavioural fields because of the fear of how this knowledge can be used to manipulate and encroach on people's freedoms. It rather seems essential to understand the evidence that has come to light in order to design effective and beneficial policies. However, it is my understanding that this should not be done by ways of using these findings for governments' own advantage as libertarian paternalists propose, by exploiting people's automatic cues without their knowledge and by interfering in areas that perhaps should not fall under a governmental scope.

It is with this in mind, that in closing, I wonder if liberty is as important to Sunstein and Thaler as they say it is. From a different perspective it is easy to see

---

<sup>356</sup> Poole (2014)

<sup>357</sup> Oliver (2013a; 2017)

<sup>358</sup> Sunstein (2013)

<sup>359</sup> *Ibid.*

<sup>360</sup> Sugden (2018)

measures that also use behavioural insights to design policies that protect individuals from certain behaviours, and that can help them overcome some of their internalities and most importantly that don't seem to carry such strong implications for political liberty and decisional autonomy.

For example, when faced with all the behavioural findings, paternalists that value freedom could design an interim educational proposal. This would be intended to improve citizen's behavioural literacy in order to overcome, or at the very least, become aware of their cognitive biases/heuristics and how these allow for external manipulation by central planners<sup>361</sup>. In making their position temporary, they would still be descriptively paternalistic as their purpose would still be to enhance quality of choice towards welfare, yet they would not impose policies on people but rather leave them free to act as they desired, whether in favour of System 1 or of System 2, once their educational period would end.

In using *debiasing* techniques<sup>362</sup>, libertarian paternalists could empower people to learn and expand their critical abilities, overcome decisional blunders, become more autonomous by their own standards, as our System 2 selves would expand<sup>363</sup>, and therefore make preferable choices. The idea of educating citizens to expand their rationality, to overcome their cognitive biases, to call attention to possible manipulation from outside parties, instead of using unconscious and subconscious cues to guide people towards their future well-being, can go much more hand in hand with a libertarian position. Even Mill, who looked at centralized education as a threat to individualism, acknowledged that education should be used to cultivate individuals' mental capabilities and virtues, at least during citizens' formative years.

---

<sup>361</sup> Mitchell (2004)

<sup>362</sup> *Ibid.* see also Rebonato (2012); Ginsburg and Wright (2012)

<sup>363</sup> Oliver (2013a; 2015); see also Poole (2014); Rebonato (2012)

Libertarian paternalists propose instead to use known behavioural fallibilities to comply with their objective idea of what individuals “think is good for themselves”. Having this in mind it bears quoting Mill’s last thoughts in *On Liberty*:

“The worth of a State, in the long run, is the worth of the individuals composing it; and a State which postpones the interests of their mental expansion and elevation, to a little more of administrative skill, or of that semblance of it which practice gives, in the details of business; a State which dwarfs its men, in order that they may be more docile instruments in its hands even for beneficial purposes, will find that with small men no great thing can really be accomplished.”<sup>364</sup>

But, considering behavioural insights, there might lie another possible way to become legitimately paternalistic, one that does not preclude the use of citizen’s biases for governments own gain. Instead paternalists could try to protect people by focusing on those who might, and often do, exploit people’s errors in order to promote their own self-interest. Different industries are more than aware of the empirical evidence shown in chapter 2 and the different tendencies of human behaviour. Many of these industries have chosen to employ different means of behavioural manipulation in the direction of their own interests to garner larger profits, and less in the directions of individuals’ own good. As these industries are often not dictated by moral limits or benevolent justifications, this exploitation of our cognitive biases/heuristics can itself harm individuals and can also be considered an abuse of private power over consumers.

Thus, it seems plausible to propose that governments are justified to regulate the abuse and excesses of this type of private manipulation. As even Isaiah Berlin claimed that “freedom for the wolves has often meant death to the sheep” and that a position of absolute non-interference in markets “armed the strong, the brutal and the unscrupulous

---

<sup>364</sup> Mill (1863) p. 223

against the humane and the weak” failing to “provide the minimum condition in which any degree of significant ‘negative’ liberty can be exercised by individuals (...)”<sup>365</sup>.

In this regard, in 2011, for example, France made an addendum to their bioethics code stating that: “brain imaging methods can be used only for medical or scientific research purposes or in the context of court expertise”<sup>366</sup>. Also relating to this matter, Oliver (2013a; 2015; 2017) proposes, in contrast to nudges, the concept of “*budge policy*”: a behavioural economic-informed regulation designed to budge the private sector away from socially harmful acts.

“An education in behavioural economic concepts can sometimes give a legitimate justification – a theoretical foundation – for public officials to regain the confidence to lead, by offering potential insights into where and how their citizens’ cognitive limitations are being exploited excessively, and by also imbibing politicians with a sense of awareness of the bounds on their own rationality. Behavioural economic- informed regulation – or budge policy – calls for stronger and clearer input from behavioural economics (...) and recognises that policy makers should often aspire to budge profit-oriented industry (...) rather than nudge citizens if they are to serve more effectively the social good.”<sup>367</sup>

In contrast to these proposals, Libertarian Paternalism is inviting governments to exploit our automatic biases in order to design more effective policies. These paternalists present an idea of citizens as incompetent thinkers, with compromised rationality and as a result, governments might, under this assumption, start to treat citizens differently to one that respects people’s reflective autonomy<sup>368</sup>, succumbing to the slippery slope of paternalism.

---

<sup>365</sup> Berlin (2002a) p. 38

<sup>366</sup> Garvey (2016)

<sup>367</sup> Oliver (2013a) p. 698

<sup>368</sup> Poole (2014); see also Rebonato.

By proposing measures that are not very transparent; by dismissing the benefits of learning through experience; by applying policies that makes choices for citizens; and, most importantly, by implicitly trusting government officials to act on behalf of citizens' own well-being, Libertarian Paternalism's implications for liberty become unsettling for liberal ideals in democratic societies.

## **BIBLIOGRAPHY**

- Akerloff, G. A.; Yellen, J. L. (1987). Rational Models of Irrational Behavior. *The American Economic Review*, Vol. 77, No. 2: 137-142.
- Ariely, D. (2009). *Predictably Irrational: The Hidden Forces that Shape our Decisions* (revised and expanded ed.). New York: HarperCollins.
- Baddeley, M. (2017). *Behavioural Economics: A very short introduction*. Oxford: Oxford University Press.
- Bai, M. (17-7-2005). "The Framing Wars" *The New York Times Magazine* retrieved from <<https://www.nytimes.com/2005/07/17/magazine/the-framing-wars.html>> (last accessed April 2018).
- Barros, G. (2010). Herbert A. Simon and the concept of rationality: Boundaries and procedures. *Brazilian Journal of Political Economy*, Vol. 30, No. 3: 455-472.
- Behavioural Insights Team, The. (2013). *Applying Behavioural Insights to Organ Donation: preliminary results from a randomised controlled trial*. London: Cabinet Office.
- Behavioural Insights Team, The. (2010). *Applying Behavioural Insights to Health*. London: Cabinet Office.
- Behavioural Insights Team, The. (2012). *Annual Update 2011-2012*. London: Cabinet Office.
- Behavioural Insights Team, The. (2015). *Update Report 2013-2015*. retrieved from <https://www.behaviouralinsights.co.uk/publications/the-behavioural-insights-team-update-report-2013-2015/> (last accessed March 2018).
- Behavioural Insights Team, The. (2017). *Update Report 2016-2017*. retrieved from <<https://www.behaviouralinsights.co.uk/publications/the-behavioural-insights-team-update-report-2016-17/>> (last accessed May 2018).
- Benartzi, S. Thaler, R. H. (2004). Save More Tomorrow: Using Behavioural Economics to Increase Employee Saving. *Journal of Political Economy*, Vol. 112, No. 1:164-187.
- Benedictus, L. (2-5-2013). "The nudge unit – has it worked so far?" *The Guardian*: <<https://www.theguardian.com/politics/2013/may/02/nudge-unit-has-it-worked>> (last accessed March 2018).
- Berlin, I. (2002a). Introduction in H. Hardy (Ed.) *Liberty: Incorporating Four Essays on Liberty* (pp. 4-54). Oxford: Oxford University Press.
- Berlin, I. (2002b). Two Concepts of Liberty in H. Hardy (Ed.) *Liberty: Incorporating Four Essays on Liberty* (pp. 166-217). Oxford: Oxford University Press.

- Berlin, I. (2002c). John Stuart Mill and the Ends of Life in H. Hardy (Ed.) *Liberty: Incorporating Four Essays on Liberty* (pp. 218-251). Oxford: Oxford University Press.
- Boroditsky, L.; Thibodeau, P. H.; (2011). Metaphors We Think With: The Role of Metaphor in Reasoning. *PLOS ONE*; Vol.6. No. 2: e16782.
- Boseley, S. (12-12-2017). “Jeremy Hunt launches opt-out organ donation plans in England” *The Guardian*: <<https://www.theguardian.com/society/2017/dec/12/jeremy-hunt-launches-opt-out-organ-donation-plans-in-england-and-wales>> (last accessed September 2018).
- Bowles, S. (2016). *The Moral Economy: why good incentives are no substitute for good citizens*. New Haven and London: Yale University Press.
- Brown, P. (2012). A Nudge in the Right Direction? Towards a Sociological Engagement with Libertarian Paternalism. *Social Policy and Society*, Vol. 11, No. 3: 305-317.
- Cain, D. M.; Lowenstein, G.; Sah, S. (2013). Confessing one’s sins but still committing them: transparency and the failure of disclosure in A. Oliver (Ed.) *Behavioural Public Policy* (pp. 148-159). Cambridge, UK; New York: Cambridge University Press.
- Camerer, C.; Issacharoff, S.; Lowenstein, G.; O’Donoghue, T.; Rabin, M. (2003). Regulation for Conservatives: Behavioural Economics and the case for Asymmetric Paternalism. *University of Pennsylvania Law Review*; Vol. 151, No. 3: 1211-1254.
- Cameron, D. (Feb. 2010). “The Next age of Government”. *TED2010: What we Need Now...* available online at <[https://www.ted.com/talks/david\\_cameron](https://www.ted.com/talks/david_cameron)> (last accessed May 2018)
- Castello Branco, J. T. (2004). Isaiah Berlin: Da liberdade negativa à sociedade decente in J. C. Espada e R. Rosas (Eds.) *Pensamento Político Contemporâneo* (pp. 60-80). Chiado, Lisboa: Bertrand Editora.
- Cialdini, R. B. (2006). *Influence: The Psychology of Persuasion*. New York: HarperCollins
- Chandle, R.; Sanders, M.; Shelley, E. (2017). *Encouraging People into University: Research Report*. London: Department for Education.
- Conly, S. (2013). *Against Autonomy: justifying coercive paternalism*. Cambridge: Cambridge University Press.
- Constant, B. (2001). *A Liberdade dos Antigos Comparada à dos Modernos* (António Araújo, Trad.). Coimbra: Edições Tenacitas.
- Cudd, A. E. (2002). Preference, Rationality and Democratic Theory in R. L. Simon (Ed.) *The Blackwell Guide to Social and Political Philosophy* (pp. 106-127) Malden, Mass.; Oxford: Blackwell.

- Department of Health and Social Care. (2018). *New Approach to organ and tissue donation in England: government response to public consultation*. retrieved from <<https://www.gov.uk/government/consultations/introducing-opt-out-consent-for-organ-and-tissue-donation-in-england>> (last accessed September 2018).
- Dolan, P.; Hallsworth, M.; Halpern, D.; King, D.; Vlaev, I. (2010). *MINDSPACE: Influencing Behaviour Through Public Policy*. London: Cabinet Office and Institute for Government.
- Dolan, P. (2013). Influencing the Financial Behaviour of Individuals: the mindspace way in A. Oliver (Ed.) *Behavioural Public Policy* (pp.191-208). Cambridge, UK; New York: Cambridge University Press.
- Donner, W. (2010). Autonomy, Tradition and the Enforcement of Morality in Ten C. L. (Ed.) *Mill's On Liberty: A Critical Guide*. Cambridge: Cambridge University Press.
- The Economist*. (24-3-2012) Nudge Nudge, Think Think. available online at: <<https://www.economist.com/finance-and-economics/2012/03/24/nudge-nudge-think-think>> (last accessed December 2018).
- Espada, J. C. (2008). *A Tradição Anglo-Americana da Liberdade: Um Olhar Europeu*. Cascais: Principia.
- Frederick, S. (12-4-2010). "Of Frogs and Man" (a response to Whitman). *Cato Unbound*, April Issue, 2010. available online at: <<https://www.cato-unbound.org/issues/april-2010/slippery-slopes-new-paternalism>> (last accessed December 2018).
- Frey, B.; Eichenberger, R. (1994). Economic Incentives Transform Psychological Anomalies. *Journal of Economic Behaviour and Organization*, Vol. 23, No. 2: 215-234.
- Frey, B.; Oberholzer-Gee, F. (1997). The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out. *American Economic Review*, Vol. 87, No. 4: 746-755.
- Frey, B. (2013). How should people be rewarded for their work in A. Oliver (Ed.) *Behavioural Public Policy* (pp. 165-183). Cambridge, UK; New York: Cambridge University Press.
- Garvey, J. (2016). *The Persuaders: The Hidden Forces that are Trying to Change your Mind*. London: Icon Books.
- Gigerenzer, G. (2015). On the Supposed Evidence for Libertarian Paternalism. *Review of Philosophy and Psychology*; Vol. 6, No. 3:361-383.
- Gigerenzer, G. (2016). Taking Heuristics Seriously in Samson, A (Ed.) *Behavioral Economics Guide 2016*. retrieved from: <<http://www.behavioraleconomics.com/>> (last accessed July 2017).

- Gigerenzer, G.; Hertwig, R. (1999). The Conjunction Fallacy Revisited: How Intelligent Inferences look like Reasoning Errors. *Journal of Behavioural Decision Making*; Vol. 12, No. 4: 275-305.
- Ginsburg, D. H.; Wright, J. D. (2012). Behavioural Law and Economics: Its Origins, Fatal Flaws, and Implications for Liberty. *Northwestern University Law Review*; Vol. 106, No. 3: 1033-1090.
- Glaeser, E. L. (2006). Paternalism and Psychology. *University of Chicago Law Review*; Vol. 73, No.1: 133-156.
- Goldcare, B.; Haynes, L.; Service, O.; Torgerson, D. (2013). *Test, Learn, Adapt: Developing Public Policy with Randomised Control Trials*. London: Cabinet Office.
- Goldstein, D.; Johnson, E. J. (2003). Do Defaults Save Lives? *Science*; Vol. 302, No. 5649: 1338-1339.
- Henderson, D. R. (2014). Libertarian Paternalism: Leviathan in Sheep's Clothing?. *Society*; Vol. 51, No. 3:268-273.
- House of Lords. (2011). *Behaviour Change*. London: The Stationary Office.
- Jolls, C.; Sunstein, C. R.; Thaler, R. (1998). A Behavioural Approach to Law and Economics. *Stanford Law Review*, Vol. 50, No. 5: 1471-1550.
- Kahneman, D. (2012). *Thinking, Fast and Slow*. London: Penguin Books.
- Kahneman, D.; Tversky, A. (1974). Judgements Under Uncertainty: Heuristics and Biases. *Science*; Vol. 185, No. 4157 in D. Kahneman (2012). *Thinking Fast and Slow* (Appendix A, pp. 419-432).
- Kahneman, D.; Tversky, A. (1979). Prospect Theory: An analysis of Decision under Risk. *Econometrica*; Vol. 47, No. 2: 263-91.
- Kahneman, D.; Tversky, A. (1984). Choices Values and Frames. *American Psychologist*; Vol. 34 in D. Kahneman (2012) *Thinking fast and slow*. (appendix B, pp. 433-448) London: Penguin Books.
- Kahneman, D.; Tversky, A. (1986). Rational Choice and the Framing of Decisions. *The Journal of Business*; Vol. 59, No. 4, Part 2: The Behavioral Foundations of Economic Theory: 251-278.
- Klick, J.; Mitchell, G. (2006). Government Regulation of Irrationality: Moral and Cognitive Hazards. *Minnesota Law Review*, Vol. 90, No. 6: 1620-1663.
- Klick, J. (9-4-2010). "The Dangers of Letting Someone Else Decide" (a response to Whitman). *Cato Unbound*, April Issue, 2010. available online at: <<https://www.cato->

- unbound.org/issues/april-2010/slippery-slopes-new-paternalism> (last accessed December 2018).
- Korobkin, R. (2009). Libertarian Welfarism. *California Law Review*; Vol. 97, No. 6:1651-1685.
- Lakoff, G.; Johnson, M. (1980). *Metaphors we Live By*. Chicago: University of Chicago Press.
- Laupies, F. (2005). *A Liberdade*. Mem-Martins, Sintra: Publicações Europa-América.
- Le Grand, J. (1997). Knights, Knaves or Pawns? Human Behaviour and Social Policy, *Journal of Social Policy*; Vol. 26, No. 2: 149-169.
- Linden, S. Van. (2013). A response to Dolan in A. Oliver (Ed.) *Behavioural Public Policy*. (pp. 209-215). Cambridge UK, New York: Cambridge University Press.
- Lourenço, J. S.; Ciriolo, E.; Almeida, S. R.; Troussard, X. (2016). *Behavioural Insights Applied to Policy: Report 2016*. Brussels: European Union.
- Lowenstein, G.; Brennan, T.; Volpp, K. G. (2007). Asymmetric Paternalism to Improve Health Behaviors. *The Journal of the American Medical Association*; Vol. 298, No. 20: 2415-2417.
- Malik, S.(30-4-2013). “Jobseekers made to carry out bogus psychometric test” *The Guardian*: <<https://www.theguardian.com/society/2013/apr/30/jobseekers-bogus-psychometric-tests-unemployed>> (last accessed September 2018).
- Malik, S. (6-5-2013). “Jobseekers Psychometric test is a Failure”. *The Guardian*: <<https://www.theguardian.com/society/2013/may/06/jobseekers-psychometric-test-failure>> (last accessed October 2018).
- Marlin, R. (2013). *Propaganda and the Ethics of Persuasion* (2<sup>nd</sup> Ed.). Peterborough: Broadview Press.
- Mill, J. S. (1863). *On Liberty*. Boston: Tickner and Fields.
- Mitchell, G. (2004). *Libertarian Paternalism is an Oxymoron*. FSU College of Law, Public Law Research Paper No. 136; FSU College of Law, Law and Economics Paper No. 05-02.
- Nelson, W. (2002). Liberal Theories and Their Critics in R. L. Simon (Ed.) *The Blackwell Guide to Social and Political Philosophy* (pp. 197-217). Malden, Mass.; Oxford: Blackwell.
- Oliver, A. (2013a). From Nudging to Budging: Using Behavioural Economics to Inform Public Sector Policy. *Journal of Social Policy*; Vol. 42, No. 4: 685-700.
- Oliver, A. (Ed.) (2013b) *Behavioural Public Policy*. Cambridge, UK; New York: Cambridge University Press.

- Oliver, A. (2015). Nudging, Shoving and Budging: Behavioural Economic-informed Policy. *Public Administration*; Vol. 93, No. 3:700-714.
- Oliver, A. (2017). *The Origins of Behavioural Public Policy*. New York: Cambridge University Press.
- Orlowski, A. (6-2-2014). "Nudge Unit flies into Nesta's arms: Is the Hype Justified?" *The Register*: <[https://www.theregister.co.uk/2014/02/06/the\\_nudge\\_myth/](https://www.theregister.co.uk/2014/02/06/the_nudge_myth/)> (last accessed October 2018).
- Poole, S. (22-9-2014). "Not So Foolish". *Aeon*. available online <<https://aeon.co/essays/we-are-more-rational-than-those-who-nudge-us>> (last accessed May 2018).
- Rabin, M. (1998). Psychology and Economics. *Journal of Economic Literature*, Vol. 36, No. 1: 11-46.
- Rebonato, R. (2012). *Taking Liberties: A Critical Examination of Libertarian Paternalism*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
- Ryan, A. (1998). Mill in a Liberal Landscape in J. Skorupski (Ed.) *The Cambridge Companion to Mill*. Cambridge: Cambridge University Press.
- Salvat, C. (2008) *Is Libertarian Paternalism an Oxymoron?: A Comment on Sunstein and Thaler (2003)*. Working Paper. retrieved from: <<https://hal.archives-ouvertes.fr/hal-00336528>> (last accessed August 2018).
- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*; Vol. 69, No. 1: 99-118.
- Simon, H. A. (2009). *An Empirically Based Microeconomics*. New York: Cambridge University Press.
- Skorupski, J. (Ed.) (1998). *The Cambridge Companion to Mill*. Cambridge: Cambridge University Press.
- Sugden, R. (2006). What we desire, what we have reason to desire, whatever we might desire: Mill and Sen on the value of Opportunity. *Utilitas*; Vol. 18, No. 1:33-51.
- Sugden, R. (2008). Why Incoherent Preferences Do Not Justify Paternalism. *Constitutional Political Economy*; Vol. 19, No. 3: 226-248.
- Sugden, R. (2017). Do People Really Want to Be Nudged Towards Healthy Lifestyles? *International Review of Economics*; Vol. 64, No.2:113-123.
- Sugden, R. (2018). Better Off "as judged by themselves": A Reply to Cass Sunstein. *International Review Economics*; Vol. 65, No.1: 9-13.
- Sunstein, C. R. (1997). *A Behavioural Analysis of Law*. University of Chicago Law Review; Coase-Sandor Institute for Law & Economics, Working Paper No. 46.

- Sunstein, C. R. (2012). *Impersonal Default Rules vs. Active Choices vs. Personalized Default Rules: A Triptych*. (unpublished manuscript) retrieved from:  
<<https://dash.harvard.edu/handle/1/9876090>> (last accessed November 2018).
- Sunstein, C. R. (2013). The Storrs Lectures: Behavioural Economics and Paternalism. *Yale Law Journal*; Vol. 122, No. 7: 1829-1899.
- Sunstein, C. R. (2014a). *Why Nudge: The Politics of Libertarian Paternalism*. New Haven and London: Yale University Press.
- Sunstein, C. R. (2014b). Choosing Not to Choose. *Duke Law Journal*; Vol. 64, No. 1.
- Sunstein, C. R. (2016). *The World According to Star Wars* (E-Pub Ed.). HarperCollins.
- Sunstein, C. R.; Thaler, R. H. (2003a). Libertarian Paternalism is not an Oxymoron. *The University of Chicago Law Review*; Vol. 70, No. 4: 1159-1202.
- Sunstein, C. R.; Thaler, R. H. (2003b). Libertarian Paternalism. *American Economic Review*; Vol. 93, No. 2: 175-179.
- Sunstein, C. R.; Thaler, R. H. (2009). *Nudge: Improving Decisions about Health, Wealth and Happiness*. New York: Penguin Books.
- Taylor, C. (1985). What's wrong with Negative Liberty in *Philosophy and the Human Sciences: Philosophical Papers Vol. II* (pp. 211-229). Cambridge: Cambridge University Press.
- Thaler, R. H. (7-4-2010). "Fear of Falling" (a response to Whitman). *Cato Unbound*. April Issue, 2010 retrieved from <<https://www.cato-unbound.org/issues/april-2010/slippy-slopes-new-paternalism>> (last accessed December 2018).
- Veetil, V. P. (2011). Libertarian Paternalism is an Oxymoron: an essay in defence of Liberty. *European Journal of Law and Economics*; Vol. 31, No.3: 321-334.
- Wertheimer, A. (2002). Liberty, Coercion and the Limits of the State in R. L. Simon (ed.) *The Blackwell Guide to Social and Political Philosophy* (pp. 38-59). Malden, Mass.; Oxford: Blackwell.
- Whitman, D. G.; Rizzo, M. J. (2007). *Paternalist Slopes*. New York University Law and Economics Working Papers. Working Paper No. 87.
- Whitman, D. G. (2010). "The Rise of New Paternalism" *Cato Unbound*, April Issue, 2010. available online <<https://www.cato-unbound.org/issues/april-2010/slippy-slopes-new-paternalism>> (last accessed December 2018).
- World Development Report. (2015). *Mind, Society and Behaviour*. Washington, DC: The World Bank.

## Selected Websites

Aeon. <<https://aeon.co/>>

The Atlantic. <<https://www.theatlantic.com/>>

BBC. <<https://www.bbc.com/>>

The Behavioural Insights Team. <<https://www.behaviouralinsights.co.uk/>>

Cato Unbound - A Journal of Debate. <<https://www.cato-unbound.org/>>

The Economist. <<https://www.economist.com/>>

The Guardian. <<https://www.theguardian.com/>>

New York Times. <<https://www.nytimes.com/>>

The Stanford Encyclopedia of Philosophy. <<https://stanford.library.sydney.edu.au/>>

TED. <<https://www.ted.com/>>

UK Government. <<https://www.gov.uk/>>

The World Bank. <<https://www.worldbank.org/>>

