

Governing Cognitive Biases

Case studies of the use of and justifications for behaviorally
informed policy tools

Abstract

In the last few decades, researchers have identified many systematic errors – cognitive biases – in the human mind. The predominant notion of human beings as rational have gradually been flawed, in one context after another. In the last few years scholars have approached the question of what policy implications should be derived from these findings and what solutions there might be. Three main philosophic lines have been argued for: soft paternalism, coercive (hard) paternalism and skeptical libertarians. The theoretical framework of this paper is composed of case relevant cognitive biases together with suggested debiasing strategies. The main aim of this study is to bridge the gap between theory and practice. By analyzing mainly official documents in three cases – the UK, the US and the EU – of in-practice use of, both a case specific and an aggregated picture are presented. Questions of *why* and *how* government administrations have chosen to work with behavioral insights are in focus. The results show that a *soft paternalist approach is the preferred strategy*, but more coercive elements are also used. The identified ways to tackle cognitive biases in practice can be concluded in the use of *default options*, *simplification* and *smart disclosure*.

Key words: cognitive biases, behavior, policy, paternalism, nudge, practice.

Words: 19872

Table of contents

1	Introduction	1
1.1	Introduction	1
1.2	Purpose	3
1.3	Research question	3
1.4	Disposition/outline of thesis	4
2	The bumps in rationality.....	6
2.1	The rational agent	6
2.2	Cognitive biases.....	7
2.2.1	Dual-process theory – System 1 and System 2	8
2.2.2	Thumb rule heuristics	9
2.2.3	Framing and loss aversion	10
2.2.4	The status quo bias	11
2.2.5	Social conformity biases.....	11
2.2.6	The scarcity-mindset an tunneling.....	12
2.3	Finding policy-relevant bumps.....	14
2.4	Relevance for governments	15
3	Debiasing philosophies	17
3.1	Libertarianism.....	17
3.2	Soft Paternalism.....	19
3.3	Hard Paternalism	21
3.4	Different types of choices.....	23
3.5	In-betweeners	24
3.6	The role of behavioralists in policymaking.....	25
4	Methodological considerations	27
4.1	A Qualitative Case-study.....	27
4.2	Case selection	27
5	The cases.....	30
5.1	The adoption into policymaking.....	30
5.2	Structure of the analysis	30
5.3	The UK– The Nudge Unit	31
5.3.1	Birth, role and justification.....	31
5.3.2	Targeted biases	32
5.3.3	Debiasing approach	36
5.3.4	Concluding the UK analysis.....	37
5.4	The US– The Bureaucrat Gatekeeper.....	37

5.4.1	Birth, role and justification.....	37
5.4.2	Targeted biases	38
5.4.3	Debiasing approach	41
5.4.4	Concluding the US analysis.....	42
5.5	The EU – The Protective Behaviorist.....	43
5.5.1	Birth, role and justification.....	43
5.5.2	Targeted biases	44
5.5.3	Debiasing approach	45
5.5.4	Concluding the EU analysis	45
6	Conclusions	47
7	Further discussion	49
7.1.1	Democratic aspects.....	49
8	Executive Summary.....	50
8.1	Subject and outline	50
8.2	Biases.....	51
8.3	Results	52
8.3.1	The how-question	52
8.3.2	The why-question	53
9	References	55
10	Appendix 1	60

1 Introduction

1.1 Introduction

In many parts of the world governments are struggling to make ends meet in their national budgets. This is no less the case since the economical crisis hit the world around 2008. At the same time populations grow and more funding is required for areas such as healthcare and education to maintain the same level of health and cultivation in the welfare states. The populations of the western world are also getting older while the healthy life-years¹ are not increasing in the same rate. Chronic diseases like diabetes is another major challenge. All this puts strong pressure on the welfare systems and national budgets around the western world. These times of demographic challenges and austerity, or even retrenchment, calls for considerations regarding new or complementary approaches to be used in policymaking. The main issue when searching for such approaches is that they will, preferably, have to be cheap. National budgets are tightened and finding political agreements on extensive reallocations of economical resources are often limited. So how may governments improve policy and welfare programs without having to increase budget items?

After what have been described as the worst economical crisis since the late twenties, prevailing economic theory have been questioned. In the light of societal challenges policymakers and public administrations are searching for viable and efficient policy options. It comes across as no coincidence that the interest in applying behavioral insights to policymaking seem to have started for real during the crisis years (around 2008). Even though findings, about the irrationality in civic decision-making, have been around for a while, the best-selling book *Nudge – Improving Decisions About Health, Wealth and Happiness* by behavioral economist Richard Thaler and law scholar Cass Sunstein, first released in 2008², was regarded by many as the first basic manual on how to apply behavioral insights³ to policy (e.g. Kahneman 2011: 372). The book seems to have been published at a very fortunate time.

The question and promise the advocators of these behavioral approaches put forward is: How can what we today know, about human behavior, help us streamline our current and coming policies by dodging counterproductive

¹ The number of years up to where people consider themselves being mainly unhealthy.

² This paper will here on refer to the second (2009) edition of this work.

³ As explained further in 2.2 behavioral insights draw on cognitive biases in human decision-making.

elements and make policy programs and regulations more efficient and possible increase citizens wellbeing?

In the last few decades behavioral research have found many – sometimes crucial – flaws in the way individuals make decisions and choices. The identification of such biases took off in the 1970s with some of the most cited articles in social science even today, written by Daniel Kahneman and Amos Tversky (e.g. 1974, 1981, 1984). While the work on cognitive biases have been going on for some time, the ideas on how to ‘debias’ these biases, and also how to gain political understanding and influence for the debiasing methods, have lingered a bit more.

It is highly important that the different aspects of cognitive biases and their implications are extensively examined. The inherent potential to influence civic behavior is strong and therefore it deserves to be scrutinized *if, where and how* it can be a valid and efficient foundation for policymaking. If proven to be a shady silver-bullet, like some critics would argue, the knowledge still has value and can help policymakers turn to other policy tools. Collecting knowledge like this requires multidisciplinary research on the subject, all the way from natural sciences to behavioral, social, political science and philosophy; from basic knowledge about human behavior to knowledge about how such behavior may affect and be adjusted into political and administrative systems.

One potential solution to these biases, suggested by Thaler & Sunstein (2009), is the use of ‘nudges’. A nudge is a *soft paternalistic* policy-intervention where freedom of choice is maintained and economic incentives are not significantly changed. By using small means and creating small changes in the social environment, policymakers may be able to adapt to the cognitive flaws and change civic behavior by inducing citizens to make better choices for themselves, choices that – at the same time – have the potential to lower the expenses of the national budget. Nudges can be used both on smaller and more comprehensive policy issues. A small nudge could be to put vegetables at eye level in the school canteen while a more comprehensive nudge could be to switch the default setting to “opt-in” in a national organ donation program.

There are however those who insist that the behavioral findings on cognitive biases require even more drastic measures. Sarah Conly (2013) argues that while nudging can help many to make better choices, many other individuals will still be making the same inferior choices and suffer from the outcomes. Therefore Conly demands a higher degree of coercive paternalism to conquer the biases.

Whichever solution gets picked, the main idea is that policymakers should base their policymaking, not on the predominant rational assumptions derived from neoclassic economics, but on the behavioral insights derived from *how people actually make decisions*.

In this paper we will look into how different administrations – the UK, the US and the EU – have worked with theories and strategies on cognitive biases, by applying behavioral insights to their policymaking. Behavioral findings are presented alongside philosophical approaches on paternalism and together this forms the theory of this paper. Finally the cases are being analyzed in the light of this theory. The main angles, from which the cases will be dealt with are: i) *the*

why-question; the reasons, justifications and legitimization for using these approaches, and ii) *the how-question*; how have the cases been working with these approaches, what strategies have been used and which biases have been specifically targeted.

1.2 Purpose

Much has been written about cognitive biases, some have been written about the use of these biases in policymaking regarding civic decisions.⁴ However, none so far seem to have put this in a wider perspective where different theories meet in-practice use in multiple countries. Therefore, the purpose of this paper have been to connect, relate and analyze:

- 1) *Cognitive biases*
- 2) *Theories on paternalism*, and
- 3) *In-practice work* by the most influential cases.

To do this it is necessary to:

- 1) Map the cases where behavioral theories on cognitive biases have been applied as regulatory tools in policymaking.
- 2) Find relevant theoretical explanations – to the use and justifications of these behavioral approaches – both by presenting the scientifically based theory and different philosophical approaches to policy interventions.

This paper is not an impact assessment of the conducted policies and regulations in the included cases; rather the focus is on the ideas and strategies used in the respective case and how these relate to the wider theoretical behavioral approaches. It has not been the intention to establish any causal explanations in this paper, however, the targeted biases and ways to intervene them *does* tell us something about what strategies tend to be used more than others or are easier to gain understanding for among policymakers.

1.3 Research question

Main research question:

⁴ Much of the previous research on these areas will be presented in the theory chapters (2 and 3).

What ideas and strategies, based on the knowledge on cognitive biases, have been applied, in practice, by government administrations in their policymaking?

The main question is answered by breaking it down into two parts:

- 1) *Why* do administrations choose to apply behavioral approaches into policymaking?
- 2) *How* have the administrations been working with behavioral theory as a policy tool?

Question 1, *the why-question* is about justification. Why have administrations chosen to work with these approaches and at this time.

Question 2, *the how-question*, is dealing with in what ways the different cases have chosen to work with behavioral theories on cognitive biases. Which biases have been in focus and what approaches to libertarianism/paternalism have been considered and used? Have the cases applied mainly theoretical or empirical approaches? The question also connects to legitimacy issues, especially when dealing with more or less paternalistic interventions.

1.4 Disposition/outline of thesis

In *chapter 2* the **cognitive biases**⁵ are presented. Initially the rationality assumptions will be discussed before moving over to the specific biases, or the ‘bumps in rationality’, that will act as the analytical seeds for this paper. Chapter 2 also deals with some theoretical implications or consequences these biases may have to citizens and policymakers.

In *chapter 3* the possible philosophical – or even ideological – **ways to react** to the biases are presented by discussing three overall approaches: libertarianism, soft paternalism and hard paternalism.

Chapter 4 will tell readers about the methodological considerations that have been used when shaping this paper. The **method** used is shortly described as qualitative case studies based on analysis of documents.

In *chapter 5* we will look at **the cases**. Focus will be on *why* the administrations have chosen to use behavioral theory in policymaking and *how* they have been conducting the work, mainly focusing on the biases and the kind of interventions used to correct biases (debias).

While most of the analysis is presented in each case, *chapter 6* is a **concluding analysis** where some comparisons between cases and overall tendencies are made.

⁵ Sometimes referred to as just “biases”.

Finally *chapter 7* presents **further discussion** points of aspects that are noteworthy but does not fully fall under the outline of this paper.

2 The bumps in rationality

It's hard to give up the idea of ourselves as completely rational. We feel as if we lose some dignity. But that's the way it is, and there's no dignity in clinging to an illusion.

Sara Conly 2013b

2.1 The rational agent

The old commonsensical understanding in economics and other sciences that human beings make rational decisions has become more and more nuanced in the last few decades. Criticizing over-beliefs in the assumptions within rational choice theory sometimes even feels like bashing-in open doors. Advocators turn the blind eye while the opponents keep on preaching to the choir. However, while the assumption of unbounded rationality *can* be feasible when making approximations of reality in well-confined cases (e.g. Ostrom 2007: 31), it might suffer from a lack of validity when dealing with more complex individual decisions and fails to explain why individuals tend to, sometimes, make inferior decisions about their own lives. The last statement is especially true if those inferior choices are not random, or due to lack of information, but due to systematic biases within the human behavior.

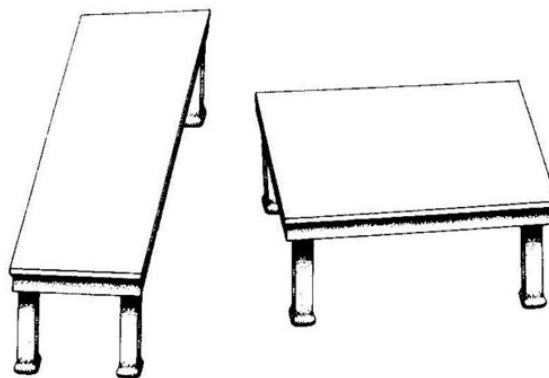
Very simplified, the assumptions used in rational choice or rational agent theory are that: *Individuals act as if they weigh costs versus benefits in order to maximize their utility within a certain delimited context.* It is also assumed that individuals have a clear order of priority for given alternatives and that they have a well defined goal. Yet another highly important assumption of rationality is that *individuals actually behave in ways consistent with their preferences.* This is the most relevant part of the rationality assumption in the light of this paper. The biases presented in 2.2 show us that preferences and behavior are not necessarily consistent.

The obvious advantage of the assumptions in rational agent theories is that researchers can explain actions made by actors (citizens, policymakers, consumers etc.) with only very little information at hand. Given a certain turnout of an action, rational choice theory can help back track the turn of events and describe the reasoning of the involved actors, using the logic of the assumptions (e.g. Allison & Zelikow 1999: ch. 1). An objection to this is that such an instrumental view of people's behavior might be too commonsensical and misses the crucial point on how people really behave. In modern decision theory models are still often

constructed with the rational actor as the ideal type (Tversky & Kahneman 1974:1130).

Thaler & Sunstein (2009: ch. 1) differs between the *rational* individual, *Econ* and the *not so rational* individual, *Human*. Both Econs and Humans can make decisions that lead to inferior results, but if Econs does so it is because they were lacking capacity or information to make a better decision at that given time. When Humans, on the other hand, make inferior decisions (which they do all the time) it is because the human mind and the human senses are prone to make systematic errors – biases – when dealing with certain issues. An Econ could never be systematically wrong. Humans are however teachable and fairly good at making decisions where the feedback is imminent and where they are allowed to re-decide in a near future. At the same time, Humans are not very good at making complex and rare decisions, such as picking an education or a retirement plan.

In this paper some of the Human biases will be presented and implications to policymaking and potential solutions will be reflected upon. Before that, let us see if you, the reader, are an Econ or a Human by looking at the so-called Shepard's table. Can you decide the proportions of these tables and tell which one is the longer?



Now of course these two tables have the exact same proportions.⁶ The way our minds fool us when trying to make sense of the size and shape of these two tables is very similar to the way the cognitive biases affects Humans regarding some decisions. An Econ would not even understand the question since it sees the tables exactly a like (Thaler & Sunstein 2009: 17f.).

2.2 Cognitive biases

The broad notion ‘behavioral theory’ is commonly used in policy-related contexts when referring to something a bit narrower. ‘Behavioral theory’ does mean, both

⁶ If you do not feel proven wrong, there is another visualization of the tables in the appendix 1

in these policy contexts and in this paper, *theory that is based on cognitive biases in human decision-making*. This part of the paper will deal with the case relevant cognitive biases and potential implications to policymaking.

There are many more biases than the ones presented here since the focus have been put on the biases that corresponds to the in-practice use of behavioral ideas in the cases presented in chapter 5. The biases presented below are in most cases the result of decades of research – for each and every bias – that, here, have been cut down to its main essence. Obviously such a cut will not do justice to all aspects and I would highly recommend reading the actual papers and books that these conclusions are based on.

It cannot be stressed enough that the cognitive biases are not an information problem. Nor is it the result of intentional negligence by ‘some individuals’. The biases lead to inferior irrational choices that are unintentional systematic shortcomings of our minds. The theory will show how this can be considered by policymakers when designing policy, and how behaviorally informed policy-tools can work as an alternative, or complement to other policy tools, used when trying to change civic behavior.

2.2.1 Dual-process theory – System 1 and System 2

The division in *System 1* and *System 2* has many different names. Sometimes it is referred to as the automatic system and the reflective system, sometimes as fast and slow thinking (Kahneman 2011), the “doer” and the “planner” (Thaler & Sunstein 2009), or as the founders of “System 1 and 2”, Keith Stanovich and Richard West, nowadays prefer to call it: Type 1 and Type 2 processes (Stanovich & West 2000; Stanovich 2011: 19ff.). These different names on two types of cognitive ‘systems’ tell us a lot about what it is about. Research has found that humans are processing information and making choices and decisions in two different ways, one intuitive (System 1) and one reflective (System 2). The relation between System 1 and System 2 can sometimes function as a cognitive bias in itself, but also as a base for other identified biases. This division is commonly known as ‘dual-process theory’.

System 1 – the fast, intuitive doer – is in general good at what it does, dealing with familiar situations and short-term predictions. Often it is important that we can make decisions rather swift and intuitive, however this is also what makes System 1 conflicted with systematic errors. System 1 makes us answer simpler questions than the actual question and has serious problems dealing with statistical and logical problems. The really serious twist; *System 1 can never be turned off* (Kahneman 2011: 25). Therefore, the intuitive reasoning will always interfere with your reflective processes. To demonstrate this, try to answer the simple problem below. Even if you probably manage to get it right, you will hopefully feel the conflict between the intuitive answer and the reflective.

A bat and a ball cost 11 €.

The bat costs 10 € more than the ball.

*How much does the ball cost?*⁷

System 2 on the other hand – *the slow, reflective planner* – is responsible for self-control and for taming the impulses of System 1. It is in System 2 we make computations that require attention and is therefore easily disrupted (by System 1). While System 1 is always running, *System 2 is lazy*. When dealing with the theme of this paper it is of great importance that System 2 is “often associated with the subjective experience of agency, choice and concentration.” (Kahneman 2011: ch. 1).

No one is able to protect herself from System 1 and 2 biases by always keeping processes in System 2. That method would be way too slow and take up way too much mental strength. There is however the possibility that we can, to some extent, learn to recognize situations in which we are likely to make crucial systematic errors. One of the main premises presented by Kahneman (2011) is although that *we are more likely to see other individuals' errors than our own*. This premise opens up an implicit way for different kinds of paternalistic or deliberative decision-making approaches, as we will discuss further in chapter 3.

The baseball problem above also touches another related problem regarding intuitive decisions. While many possible solutions to cognitive biases may be fairly straightforward, and in line with our intuition, solutions to cognitive biases can sometimes be more sophisticated or even counter-intuitive.

2.2.2 Thumb rule heuristics

Being somewhat pioneers in the field of cognitive biases Amos Tversky and Daniel Kahneman (1974) presented some groundbreaking research. Tversky and Kahneman mapped out three heuristic principles that humans use when we create rules of thumb or judgments, on which we often base our decisions. These heuristics is in play for example when we make predictions or estimations of the probability or risk for certain events to occur. The biases that can be derived from these heuristics are very similar to when humans make estimations of size or distances (Tversky & Kahneman 1974: 1124), something similar to what we saw in the simple example with the Shepard's table in 2.1. It should be pointed out that the heuristics presented here often works when creating rules of thumb, however they also have systematic shortcomings. Recently it has been found by scholars that the biases in these heuristics are the result of the interplay between System 1 and System 2 (Thaler & Sunstein 2009: 23).

One of the most important biases when discussing thumb rule heuristics is the *availability bias*. This bias is the flaw in peoples minds that when rare events

⁷ Did you get it right? Well obviously the answer is not that the ball costs 1 € and the bat costs 10 €. However, that is often the intuitive answer. The correct answer is given through the equation $X+(X+10)=11$ where “X” is representing the price of the ball. This gives us $X=0.5$, the ball costs 0.50 € and the bat 10.50 € making the bat cost 10 € more than the ball.

occur people are overestimating the risk/chance of such an event re-occurring. Common examples are major storms or Tsunamis (Tversky & Kahneman 1974: 1127).

Anchoring and adjusting is one of the most robust findings in experimental psychology. When thinking of a number before being faced with an estimation problem, people will mind that number even if it is completely unrelated to the problem (Kahneman 2011: 119ff.; Tversky & Kahneman 1974: 1128). Depending on what anchoring point we use, we will end up with different estimates.

Other biased thumb rule heuristic:

- Misconception of representativeness
- Insensitivity to prior probability of outcomes
- Insensitivity to sample size
- Misconceptions of chance
- Insensitivity to predictability
- The illusion of validity (too small samples)
- Misconceptions of regression (regression to the mean)

(Ibid.)

Lessons learned:

- The over belief in the reoccurrence of an event poses a risk for citizens if exploited by others
- A more or less random starting point can have a major impact on what decisions we make

2.2.3 Framing and loss aversion

If you are a rational being, like the Econ, how a decision problem is framed should have no effect on your decision. Still research shows that *how* a problem is framed can have a significant effect on people's decisions. The power of *framing* can be even stronger if it is connected to other biases, which seems to be especially true if connecting to the *loss aversion bias*⁸ (Tversky & Kahneman 1991). Tversky & Kahneman (1981) showed that framing an experimental problem to respondents in different ways had different results. Even though the numbers are the same, the respondents seem to respond different depending on which reference point that was used when framing the issue. *Is the problem framed as a gain or a loss in relation to the chosen reference point?* If the relation to a reference point is expressed through "saving 200 out of 600 lives" (a gain), that alternative is more attractive than if the same option is framed as "400 out of 600 will die for sure" (Ibid.). Other studies have given similar results when framing opportunities as financial incentives to individuals as gains versus losses.

⁸ *Loss aversion*: that individuals consider a loss to reduce their utility more than a gain, of the same size, would benefit their utility.

When looking at framing and loss aversion from a policy perspective it should be seriously considered to, for example, create programs that take away parts of future gains rather than taking away current assets from individuals.

Lessons:

- People are loss averse
- Mind the reference point and frame accordingly

2.2.4 The status quo bias

Unlike what is sometimes suggested by economic theory or rational choice theorists most decisions come with a status quo alternative. This alternative is i) to do nothing or ii) to maintain a current or previous decision. In 1988 William Samuelson and Richard Zeckhauser found that the status quo alternative was disproportionately sticky when individuals even made important decisions regarding their lives such as picking health plans or retirement programs. They entitled this cognitive flaw: the *status quo bias* (1988:8).

If policymakers, marketers or other actors see potential in exploiting this bias there is a powerful way to do so by designing well thought out default options or by making a product become the default option in that specific market (Thaler & Sunstein 2009: 35). While private companies might sometimes use the default rational in a deceiving way, with regards to the cognitive biases, it should be noted that the worst defaults of that kind are most likely eliminated by the market competition itself (Sunstein 2013: 113).

To policymakers the status-quo bias could be intervened in four main ways:

- 1) To use the power of a default option or a default rule to engage citizens in a policy program.
- 2) To use coercive means to ban all alternatives but one or by making one option universal.
- 3) To ‘counter nudge’ against nudges conducted by actors on the private market who might exploit the status quo bias in citizens.
- 4) To require individuals to make an *active choice* (Ibid.: 119).

Lessons:

- A no-action (passive) alternative is never a neutral option and has potential to change behavior if designed properly.

2.2.5 Social conformity biases

It should not come as news to anyone that we, humans, are social creatures that profoundly live our lives in the light of identities, group belonging, norms and institutions. While ‘being social’ in most cases is considered a great thing – meaning that the individual has adapted and submitted herself, in the expected

way, to the social environment – there are also downsides to social behavior. Terrible examples such as the Nazi- regime or the mass suicides of sects like the People’s Temple are telling examples of the inherent potential in social conformity. We do not have to draw on extremes like that to make a case for cognitive biases of social desirability however. Thaler & Sunstein divides the social biases into two categories: social *information* biases and *peer pressure* (2009: 54).

The *information bias* is related to the effect that people tend to use the action of others to establish how they ought to think or act, even if this means doing the most horrible or irrational things. When people, in experiments, are asked to decide on their own without seeing the results of others they are rarely failing on very simple tasks. However, if individuals are first presented with the answers of others they are failing to a much higher extent (e.g. Asch 1955). In a more recent study in Japan it was found that individuals were seven times more likely to chose an socially unacceptable option if presented with the information that the majority of the preceding respondents picked that option (Kondo et al. 2010). Note that both these findings were found by eliminating the element of peer pressure completely. There is even some evidence implying that when people say that they see things in the same way as others do in conformity experiments like this, analysis of their brains suggest they actually do see it that way (Berns et al. 2005).

The *peer pressure* bias on the other hand is related to the implicit wrath of other people if you chose a path that is not endorsed by them. This is probably a more obvious way of looking at social conformity and is commonly seen in the forms of bullying, group pressure settings of different types and the way most people dress.

Lessons:

- We *inform* ourselves using the actions of others. Even if their action is sometimes really harmful
- We are biased towards acting like others due to *peer pressure*. Even if their action is sometimes really harmful

2.2.6 The scarcity-mindset an tunneling

Most people would agree that things should be kept simple. Whether we are dealing with a ticket system for public transportation in a foreign town during a summer vacation or with an application to apply for a school or financial aid, simple is in general a good thing for the individual faced with a task. While the pros of keeping things simple might be obvious, behavioral research in the last few years have shed a new light on why reducing hassle and simplifying things can be of crucial importance, especially for underprivileged individuals.

By conducting experiments and gathering experimental findings from real life situations Sendhil Mullainatan & Eldar Shafir (2013) have gouged out a theoretical approach based on cognitive biases that are in play when humans face scarcity. Scarcity in this theory mainly refers to the lack of money (or resources),

time or social connections, but also to other types of scarcity – like scarcity of calories when on a diet – are suggested to have similar effects.

After having conducted experiments around the world Mullainatan & Shafir show that individuals, whether it is sugar cane farmers in India or mall visitors in the US, are getting their mental “bandwidth” significantly “taxed” when facing scarcity. This forces them into a constant and mentally challenging economic tradeoff reasoning. The mental bandwidth are defined and operationalized by looking at two factors: *cognitive capacity* and *executive control*. Cognitive capacity is measured using standardized IQ tests before and after scarcity interventions, while executive control is measured using simple two-button tests (right and left) examining the individual’s self-control and potential to withstand immediate reactions when the reactions are wrong. Results show that, due to the scarcity effect, executive control and IQ levels are seriously reduced. The IQ is taxed by 13-14 points on the standardized IQ scale.⁹ 13 points are enough to go from the category “average” to the category “borderline deficient” (Mullainatan & Shafir 2013: ch. 2, p. 52). This mechanism is in play, for example, before and after farmers gets paid for the year’s harvest.

While scarcity also has advantages – it actually makes you better at the task connected to managing your scarcity (called the *focus dividend*) – it severely taxes your mental bandwidth. When focusing on the sole task of reducing scarcity you are also “tunneling” on that task and are leaving out other, more or less important tasks, in your life (Ibid: 29). If suffering from scarcity over a longer period of time the tunneling effect can trap individuals, facilitate inferior decisions about other parts of their life and possibly even lead to scarcity in other areas that will amplify the effects incrementally. As an example, a widow and mother of two, who is having a hard time managing to put food on the table each day, might tunnel even crucial things such as going for free vaccine for her children. The lack of vaccination is then likely to lead to other issues for the family and the mental bandwidth is taxed even more. While snowball effect stories like this one are really unfortunate there are some ideas and means on how to counteract the scarcity traps and negative effects of tunneling.

Some would argue that findings like the ones presented by Mullainatan & Shafir (2013) calls for a total restructuring of our modern societies into more equal arrangements.¹⁰ In this paper we will not go in depth on macro solutions like that, but look more into micro level solutions that have shown to have a real impact. Two main solutions will be in focus here: *incentives* and *simplification*.

Well-designed *incentives* are incentives that – although the real policy issue is outside of the tunnel for the targeted individuals – are immediate and inside of the mental tunnel. The incentive does not have to be very large. For example, when facing the policy issue of having too many children unvaccinated in rural India, researchers found that just a kilogram of lentils proved to be an efficient incentive that made parents vaccinate their children (Banerjee et al. 2010; Mullainatan &

⁹ The mean value being 100 and standard deviation 15.

Shafir 2013: 173). When carrying out aid to the third world, free is not always going to be enough. If the advantages, of for example immunization of children, lay to far ahead in time, a visit to a medical facility is likely to be tunneled out or put aside for more immediate threats. The inconsistency when making tradeoffs, between costs and benefits, that occur in different places in time is a general flaw in the human mind and not limited to underprivileged individuals in the third world (Cappelen et al. 2010: 356). The same phenomenon can be spotted in the developed world when people are not saving enough for retirement or do not look after their medical conditions properly.

The second solution, *simplification*, does, as was pointed out initially in this part, come across as very sound and immediate. But if we can be sure of why we are in fact making choice-environments simpler for individuals we may also be able to make simplified redesigns more efficient. When trying to solve policy issues that involve individuals trapped in scarcity mindsets, simplification can be especially important. Very simply put, simplification (*ceteris paribus*) is a way to keep the mental bandwidth tax to a minimum and reduces the risk of important tasks being neglected due to procrastination, myopia or tunneling. One area with large potential for simplification is application processes and the filling out of forms. A study in the US showed that simplifying the application process for financial aid, to different degrees, had a major impact on whether or not low-income high school graduates enroll in college. When receiving personal help for filling out the forms, enrollment to college increased by 29 percent (Haycock 2006). Another experimental study showed that students from families without college experience tripled their enrollment rate if receiving help in filling out the applications (Bettinger et al. 2012). While simplifications most likely affect us all positively, it seems to be especially positive for people that are worse off.

A factor to keep in mind, when dealing with scarcity mindsets, is that *the timing of a policy intervention* can have a big impact on the turnout. This is even more important if the policy action aims at teaching or informing since that requires at least some mental bandwidth to have any impact. Approaching farmers after rather than before harvest or approaching the poor after rather than before Christmas would most likely be preferable and increase the positive impact of the intervention (Mullainatan & Shafir 2013: 219).

Lessons:

- Aiming incentives towards the inside the mental tunnel
- Simplification of procedures is especially helpful to people who are worse off in society
- Timing of policy interventions should be strategically picked

2.3 Finding policy-relevant bumps

The cognitive flaws, or bumps in rationality, that have been presented above do however not tell us anything about how individuals will act regarding a specific

new policy or regulation. To reach a point where it is possible to speak about *policy informed by how people actually make decisions* it is necessary to first theorize and then analyze. Based on the research findings concerning cognitive flaws a researcher or analyst can build a hypothesis that is possible to test in a real situation. These tests are usually conducted in the form of experiments like Randomized Controlled Trials (RCTs). Using RCTs, however, require a very thought out structure of the organization that are assigned to work with behavioral insights. Therefore there will be two kinds of behaviorally based insights. The *theory based* policy and the *empirically based* policy. The theory-based version borrows scientific findings from other similar contexts and uses it to try and change behavior in closely related contexts. The empirically based version uses theory from another context to create a hypothesis and then tests this hypothesis, using RCTs, on the policy intervention in question. Then there are obviously degrees of how theoretical or how empirical a certain policy foundation is. An even more thorough empirical approach would re-test and if necessary re-shape the hypothesis.

When dealing with human beings as the object of study there are always problems involved due to *learning*. The physicist does not very often face the dilemma of an atom having learned something new since the last time, but for a social scientist it happens all the time. People learn, change their minds and their intentions might not be very constant over time. To deal with this issue in behavioral tests such as RCTs it is probably sound to treat the test results like perishables and to be adaptive in the policymaking.

2.4 Relevance for governments

Why then is it necessary for governments to take cognitive biases into consideration? The answer is two fold. *First*, if building policy solely on the rational foundation that tells us citizens are rational, the effectiveness and efficiency of such policies will not be optimized. This will lead to citizens that are worse off than they needed to be and to governments wasting efforts without getting satisfying results. *Secondly*, because it should be a basic right to be able to make decisions in a sound and enabling environment. Or put in a different way; *how would it be justified to leave citizens in a suboptimal state if the tools to increase their wellbeing were at hand and were cheap to use?*

Why then have we, humans, come this far in terms of economic development, democratization, freedom of speech, health and wellbeing, child mortality rates, increasing life spans and so on, *despite* our biased minds and our unawareness of it up until recently? As mentioned in the introduction, the ‘2008 crisis’ were the hardest blow to western economic development since the late 1930s. In this 80 years most of the western world, especially after the Second World War, were having golden ages where economic development in terms in GDP were sky rocketing. This was not really a time where western citizens and governments had to be very efficient to make ends meet. Today an economic golden age probably

seems distant to most people and to maintain our levels of wellbeing we will have to be a bit more efficient and innovative in our economic trade-offs.

A final argument is that the research on cognitive biases and related areas has shown strong evidence, in the last few decades, on how our minds really work. This knowledge have not been available before that and therefore many traditional ways of reasoning, such as the free, autonomous and rational choice and the ways governments traditionally have been governing, could be seen as lacking a bit behind the reality that modern research depicts.

3 Debiasing philosophies

The cognitive flaws in human decision-making, such as those presented in chapter 2 are widely accepted within many fields of science, especially psychology and behavioral economics. The bumps in rationality are thus not the main issue. The controversy is rather about *if* and *how* our societies should respond to those flaws. This chapter presents different philosophical approaches to the cognitive flaws; different debiasing strategies. It also discusses the theoretical relation to the biases. The main focus is on the degrees of paternalism, stretching from *libertarianism* – almost no paternalism – to *coercive paternalism*.

The following philosophical fundamentals (3.1- 3.4) are used as ideal types when discussing and analyzing different levels of regulation. It stands for itself that in all the current systems of regulation in the world (nation states being the most prominent one), there are mixes of those three fundamentals. The extent to which, one or the other, is used might however differ a lot.

3.1 Libertarianism

Libertarianism might not be a philosophical foundation to any solution to the cognitive bias problem; rather it is the status quo from which regulation is trying to move citizens. Nevertheless, it is an important starting point for discussions on different approaches to paternalism.

The basic principles of the libertarian approach are two: First, the individual is rational enough to make her own decisions; and secondly, if the individual fails in her decisions, there is no miracle machine in the government that gives the policymakers a lesser fail-rate while taking the decisions on behalf of the individual.

The base of libertarian thoughts about the state and the state's relation to citizens goes back to the Age of Enlightenment and thinkers like John Locke, and later, John Stuart Mill – perhaps the most well-known opponent against paternalism. The Enlightenment is sometimes called the Age of Reason, which indicates the importance and strong beliefs in the power of a deliberative individual mind. An important idea drawn from this time, expressed some 100 years later, by Mill in *On Liberty* (2001 [1859]), which opposes paternalistic thought, is that it can never be justified to protect an individual against herself. It can only be justifiable to intervene when an individual is bringing harm to others. Commonly known as the “harm principle”. To be fair, Mill admits that it can be justified to stop an individual from crossing a broken bridge unknowingly. This however, raises a demarcation problem of when paternalistic intervention to

protect the individual from herself is justified, and when it is not. If accepting that the person walking on the bridge is not suicidal, this is presumably an information problem. The person walking on the bridge does not see the crack that the person below does, and because of the lack of information the person on the bridge is unknowingly in danger. While there is most likely no universal solution to these kinds of information problems, there are ideas on how societies can best disperse information and knowledge. *Education* is one example. Another example on how to disperse information regarding more economical issues is, as Fredrich von Hayek (1945) suggested, the *price mechanism*. Hayek was a true believer in the potential of market competition to solve societal issues. To put his ideas in the light of cognitive biases we turn to the dispersion of knowledge and information in society. What makes the market, or more specifically the price mechanism, a better solution than a solution created by a few decision makers? Knowledge and information are widely spread among individuals of our societies. Therefore a few experts will always have less knowledge than the society as a whole. The solution to this problem, according to Hayek, is not to have politicians or government planners make decisions based on a few expert opinions, but to let the market mechanism act as a collector of societal knowledge. The quality of a product is reflected in its price in relation to other products and so on (Hayek 1945). Others, who mainly agree with Hayek, have argued that Hayek had a point but that the tables have turned since the early-mid 1900s and that the information technologies of today present us with better alternatives to solve the issue of dispersing information in society (Sunstein 2013: 80).

The biases presented in chapter 2 stands uncorrected however, and there might be a good explanation for this in the light of market mechanisms. One very strong reason that market mechanisms will not always work towards, but sometimes rather against, a solution to the bias problem has to do with market incentives, often using marketing as a main tool. It is not news to anyone who has studied psychology or related subjects that cognitive biases have been used to convince people, often successfully, for a long time in advertising. While stating that the market competition often do a lot of good, Thaler & Sunstein highlights that in some cases market actors have strong incentives to exploit the flaws of individual decision-heuristics. Insurances are pointed out as an especially treacherous area (2009: 74, 78ff.). After a flood, that rarely occurs, for example, insurance companies are likely to – in an efficient way – take advantage of the availability heuristic (2.2.2) since people will heavily overestimate the risk of a new flood. The advantage can be taken both in advertising and overpricing of insurances.

Market mechanisms also have some ways of getting rid of situations where consumers could have been exploited through their biases, like driving companies, with more extreme examples of such exploitations, out of business (Sunstein 2013: 113). At the same time, exploiting the bias to a lesser extent, as in the flood example, remains a risk for suboptimal outcomes for our societies. Or for Hayek's price mechanism which will be infected with these biases as well.

The biases presented in chapter 2, and many more like them, are in play even if individuals are informed or educated. There is of course the idea that people should be educated in how cognitive biases work so that they might be able to

recognize situations where their minds might be fooling them. This would probably be a smart thing to do, but just as we are blind for our biases, we are also blind for our blindness of the biases, especially regarding our own biases (Pronin et al. 2002; Kahneman 2011). This makes it, in some cases, harder for *you* to determine which action is the best action for *you*. The thought can obviously be hard to accept since many of us are prone to believe (blindly) in our own senses, heuristics and judgments (Pronin et al. 2002). This overestimation of our own capacities is of course nothing but the result a cognitive bias in itself.

While educating and informing are evidently important tools in our societies, for a number of reasons, they only work all the way on rational agents such as Econs. For the rest of us, Humans, education is of great importance as well, but the cognitive biases will often leave us with what Sendhil Mullainathan calls “the last mile problem” (Idea42 2012). We know what we prefer and we believe we know how to act on what we prefer, but we do not act. And sometimes we do act, but the action is not consistent with our preferences. If we do not act or act in ways that are inconsistent with our preferences, we might just need a nudge in the right direction? Different kinds of paternalistic interventions might offer such help.

3.2 Soft Paternalism

Soft paternalism is probably the version of paternalism that has gained most new ground in the last decades. Often considered a middle way, soft paternalism comes in a few different forms. In philosophical terms ‘soft paternalism’ is usually thought of in the way Joel Feinberg expressed it in his work *Harm to Self* (1986). Feinberg uses *soft paternalism* for paternalism exercised towards an individual whose choice is insufficiently voluntary to be her own. *Hard paternalism*, on the other hand, is paternalism exercised towards sufficiently voluntary choices. In less philosophical ways *soft paternalism* is usually used as a notion for paternalism that is not coercive and that does not close any paths of action for an individual. The type of soft paternalism used for discussion here will mainly be this one. In 2003 Camerer et al. used a similar definition for what they called “Asymmetric Paternalism” in their article *Regulation for conservatives: Behavioral economics and the case for ‘Asymmetric Paternalism’*. The same year Thaler & Sunstein released their first take on soft paternalism under the name: *Libertarian Paternalism* (2003). Thaler & Sunstein later elaborated their ideas on libertarian paternalism further in their bestselling book *Nudge – Improving Decisions About Health, Wealth, and Happiness* in 2008¹¹ where they also established the notion “nudge” as a verb for applying libertarian paternalism to

¹¹ First released in 2008. However the version referred to later on in this paper is copyrighted in 2009.

policymaking. More skeptic scholars, like Whitman & Rizzo (2007) use the straightforward name “New paternalism” to categorize these middle ground ideas.

Applying libertarian paternalism by the use of nudges follow two main characteristics. A *nudge* is a change in the social environment that is:

1. *Non coercive and choice preserving*. The choices are still the same and no alternatives are taken off the table
2. *No severe shift in economic incentives* for individuals or market actors

The most important difference between hard paternalism and soft paternalism is that the latter one does not refuse the individual any choices or alternatives, it is choice preserving. Possible choices to choose from are always the result of a *choice architect*. The choice architecture might be either intentional or unintentional, but *no architecture is neutral* (Thaler & Sunstein 2009: 83ff.).

There has been some debate regarding if a nudge has to be intentional or not to be considered a nudge (for an overview see Hansen & Jespersen 2013), mainly from critics of nudging. Much of the criticism rests on the argument that intentional nudges are not neutral while “accidental” nudges are. This would, one might suspect, be followed by the definitions of *intentional* and *accidental*; which is not the case. Instead, accidental and intentional choice-environments tend to be separated only by the complexness of its structures. When a structure is too complex it is categorized as “accidental” and thus neutral. A second flaw in this line of criticism is the difference between *government action and inaction*. We all agree that pushing someone off a cliff is wrong and therefore this type of action is reflected in our laws and enforced by coercive paternalist means. But what if we are just watching someone slip down from a cliff to a certain death when we could easily have saved that someone? Having insights and tools for the redress of the cognitive biases, but not using them, has to be considered a non-neutral choice by governments. At least when dealing with inferior choices that will lead to major health threats or life threatening situations for citizens. Even if we might not want to punish people who chose the path of inaction, by law, it could hardly be considered wrong to save a person from falling off a cliff.

What should be added to this is as Cass Sunstein points out; a nudge can be both a paternalistic and a non-paternalistic intervention. A nudge is paternalistic if it tries to protect choosers from themselves, but a nudge is non-paternalistic if it tries to prevent harm to others (Sunstein 2013: 192). It is very likely that also John Stuart Mill would agree on this dual role of the nudge concept.

Information and education can act as simple and powerful nudges, although it is usually more a question of how things are informed of or educated that is of interest for a soft paternalist. The strategic way in which the information is presented is usually of great importance as well. Strategic use of bias-related factors such as framing, tunneling, timing and social conformity will have an impact when designing information campaigns. This becomes evident when regarding, for example, the difference between general information about the dangers of smoking and putting (scary) visual information on cigarette packs.

Another critic against soft paternalistic approaches like nudging is that they might work a couple of times until people get the ‘hang of it’ and might change their behavior. The counter argument to this critic is that these interventions requires adaptive governments where the soft policy interventions have to be tested and retested, preferably through experiments, to maintain their reliability and efficiency.

A more sophisticated nudge could be to draw lines just before sharp turns on roads creating the illusion that the car is moving faster than it is, which makes the drivers slow down. A simpler but often powerful nudge is to make use of the status quo bias and put in serious efforts to design smart default options. To investigate the power of default options and framing Johnson & Goldstein (2003) looked at the rate in which people donate their organs in different European countries. Similar countries – like Sweden (86%) and Denmark (4%); Austria (100%) and Germany (12%) – differed severely in donation rates. All this because of how the question, when asked to donate, was framed and whether the default alternative was to donate or not donate.

Cost-benefit analyses is sometimes brought up as a nudge that can act, both as a safeguard from rash intuitive decisions, and as a spurring nudge, to make choices or engage in policy matters that at first glance may look (intuitively) unattractive (Sunstein 2013: 152f.). The idea of using cost-benefit analysis as a nudge rests on the potential it has to force individuals or institutional bodies to make decisions in System 2 rather than System 1 – to use self-control instead of emotions when making complex decisions. As we will see when turning to the cases in chapter 2, there are many more examples of countering biases in a soft paternalistic fashion.

If the American federal gun law, stated in the famous second amendment, is a safeguard against misuse and oppression after having concentrated and delegated power, then libertarian paternalism is presented as a way to change civic behavior while maintaining a safeguard against meddling governments. Meddlesome policymakers might, in the eyes of libertarians, start to ban certain actions.

3.3 Hard Paternalism

Banning cigarettes, by law, constitutes a textbook example of hard or *coercive paternalism*, where the choices or actions are simply taken off the table and violations are being enforced, outmost, by threat of violence.

There are two obvious reasons for using coercive or law-enforced paternalism. The first one is the main rational behind the cigarettes example, to protect citizens from themselves and their inferior choices. The other rational that may justify the use of coercive paternalism is the one used when regulating prostitution in many countries. Due to certain structural imbalances of power between certain social groups of individuals the government has decided to protect the exposed individuals both from themselves and from such power structures. Another example of the structure-protection rational is the issue of second hand smoke. If a

child is brought up by two heavy smoking parents it will be exposed to serious health issues without ever having a choice.

There might also be a third alternative used when justifying coercive paternalism. Returning to the cigarette example, it might be possible to justify a ban on cigarettes due to increasing health costs that could jeopardize the popular support and economic base for national welfare programs.¹² Let us call this third version *paternalistic welfare-protection*. In this paper paternalistic interventions will, however, mostly relate to the poor-choice protection.

In her work *Against Autonomy – Justifying Coercive Paternalism* (2013) Sarah Conly takes a stand for coercive paternalism drawing on behavioral research on cognitive biases. Conly states that anti paternalists like Mill never were aware of such biases and therefore Mill's take on the 'harm principle' is outdated and it is likely that even Mill would revise his view of the justifications for paternalistic interventions. Conly argues that these scientific findings not only justify us to question the free and rational choice, but to act on them in a coercive way. The main objection from Conly against a softer paternalism such as Thaler & Sunstein's libertarian paternalism (2009) is that, while maintaining every option for citizens and making citizens – in general – better off, it will still leave a bunch of individuals – who have made poor decisions – in a bad situation. For example, even if nudging people towards healthier food, there will still be some that will stick with the unhealthy food if such food is still an option. The coercive paternalist approach Conly represents *does not want to sacrifice the well being of a few to maintain the liberty of the many* (Conly 2013: 31).

The obvious comment on this criticism by Conly is that coercive paternalism cannot really take away any options at all, that would require total supervision and safeguarding (resting on the threat of violence) of each and every citizen around-the-clock. Since this is not even remotely realistic, or what Conly argues for, options to citizens will remain in a coercive paternalist approach, but they will remain with higher or different costs. Punishments and the likely development of black-market alternatives will be a major issue to deal with for the coercive paternalist. If not dealt with efficiently enough, it will undermine the whole argument behind the justification of coercive paternalism over soft paternalism (Conly 2013: ch. 5).

In most western societies today there are some deeply rooted applications of coercive paternalism. Two examples being prescription medicine and parents right to decide for their children (or for the state to decide over both parents and children if parents fail in their care). This brings up the question on whether certain areas are deemed more or less suitable or justifiable for coercive paternalist interventions. While this is indeed a normative question the examples make it clear that coercive interventions are widely accepted in some areas by most western populations. Part 3.4 deals a bit more with paternalism in relation to different types of choices.

¹² Note that the cigarette example here is a theoretical example; in reality the costs/benefits to society regarding smoking are disputed.

Another point that Conly has a hard time handling, and that Thaler & Sunstein (unrelated to Conly) comment on, is that hard regulation does not always change the intended behavior. Nudges may sometimes have a more efficient impact on civic behavior than coercive laws (Thaler & Sunstein 2009: 101). In Sweden for example, we have a law on wearing seat belts on both public and commercial buses when travelling on the freeway. In Conly's coercive paternalist words, this means that there is no choice for the bus passengers; to wear seat belts have been decided for them. However, recent investigations show that as many as 7 out of 10 passengers do not use seat belts on public buses, while travelling on the highway (SVT 2014). This indicates at least some form of choice. There might be a number of reasons for this, but the interesting thing for this paper is that this specific paternalistic regulation – whether or not it is morally right or wrong – seems to be terribly ineffective sometimes. Even though this is a fairly new law, in a case like this it is very likely that a soft paternalistic approach would be more efficient. For example, pictures of bus crashes on billboards with number of deaths and number of seat belt wearers along the freeway would probably have a bigger impact on passenger behavior than the law.

Another alternative is to combine a coercive paternalist law with the soft paternalist billboards. It is of course possible to apply more efficient nudges for the seat belt scenario, but having a flashing light with a yelling sound on each seat felt to excessive even as an example. But then again, the extent of an intervention is always in relation to the risk of a certain behavior in a certain context. If someone failed to lower the bar on a roller coaster ride, the light-and-sound nudge would probably not be excessive.

Who then should make the decisions for us? And what makes these individuals better suited for making those decisions than the individual herself? A common understanding on paternalistic interventions is that some kind of elite, experts or democratically elected politicians are going to tell us what to do and what not to do to a greater extent. While this is of course true to some extent – because those are the types of actors we generally turn to when dealing with potential collective concerns or decisions – it is not the important part of taking away certain decisions from individuals. The important part is that these decisions are decided by individuals who act in an environment that is – as far as possible – safeguarded for the cognitive biases. Conly argues that what we need when creating paternalistic constraints, is for people who are not currently in the automatic decision system [1] to create those constraints; in other words, we would prefer if the planner rather than the doer created the paternalistic constraints (Conly 2013: 38).

3.4 Different types of choices

Drawn from the biases, and necessary to take into consideration before looking at practical solutions, we can see that there are basically two types of choices citizens or policymakers can face while trying to govern civic behavior. The first

one is the *aggregated choice*, which means that the impact of an action – for example drinking poison – will have a negative impact on all individuals who chose that action. The second type of choice is the *personalized choice*. When facing the choice of designing your retirement plan, it is very likely that (even if we assume you make a rationally optimal choice) other individuals have other optimal choices due to diversity of situation, age and so on. The subjective opinion on what type of choice is an aggregated choice and what type is a personalized choice can however make this classification a bit blurry.

The different types of choices are most likely to be of importance when looking at policy alternatives such as libertarian paternalism vs. coercive paternalism. Banning an aggregated choice, like drinking poison, is less likely to have as many adverse effects and opponents, as a ban on a personalized choice like carrying a firearm.

If faced with an aggregated choice there is, at least, one more theoretical approach when dealing with the cognitive biases. Through deliberative means and decision-making by civic interest group, there is the possibility that crucial decisions, that are otherwise likely to be made in System 1, on an individual level, can be debiased by interpersonal reflection and slower decisions. This *deliberative approach* has been presented as the “Think Think”-approach and can sometimes be seen as an alternative to nudging or coercive paternalism (John et al. 2009).

The firearm example touches on another problematic dimension of regulating or liberating citizen’s choices; the issue of *negative externalities* that comes with some individual choices. Since most of our current laws and justice systems are designed to enforce one individual’s choice from having negative spillover-effects on other individuals, this issue is to a wide extent already dealt with. And since even John Stuart Mill agreed, in the harm principle, that protecting others from this kind of adverse effects where fair game for paternalistic interventions, non of the three main approaches (3.1- 3.3) really oppose such cohesive actions.

3.5 In-betweeners

Looking into the conceptual world of different behavioral debiasing approaches it becomes clear that not all interventions can be defined as either soft paternalism or hard paternalism. There are in-betweeners.

One common example of an *in-between*er is regulating the sizes of cigarette packs or (as in New York¹³) soda cans, in an attempt to change civic behavior by drawing on the human bias that makes us eat and drink more if containers are larger (see e.g. Sunstein 2013: 40). While this can certainly have an effect on civic behavior it could also change economic incentives. Even if people can still buy their desired amount of the product, there is a possibility that regulations like this

¹³ The regulation was recently overruled, but the theoretical example holds (NBC 2014).

will alter the economic conditions for producers. Consequently, if this intervention does in fact change the economic incentives significantly, it disqualifies itself as a nudge, even though freedom of choice is maintained. An interesting point here is however, where to draw the line for changing economic incentives? If forcing market actors to redesign consumer contracts to be less confusing and prohibitive (to have people understand what they really buy), this would most likely inflict some damage on some market actors. But at the same time, most people would agree that state regulations are a profound necessity for any well functioning market. By establishing and enforcing market basics, such as contract code and property rights, it allows market actors to compete on similar terms. Laws on contract codes and property rights are obviously the result of profound coercive actions. It is in the zone of these question marks that the most interesting in-betweeners seem to be found.

3.6 The role of behavioralists in policymaking

To be able to create, at least some, economic slack in the stressed out budgets it is likely that policymakers will have to look for solutions outside of the box to apply behavioral insights to their policies. Bavel et al. (2013) suggests four “stages” to apply behavioral insights to policymaking:

- 1) When designing policy drafts
- 2) When implementing policy or making cost-benefit analysis or impact assessments
- 3) To make small pilot studies that may lead the way, *after* having decided upon an option
- 4) When critically evaluating existing policies

While these possible stages for action should be considered, it might also be relevant to approach the phenomenon in a broader way. New ideas does not always have to be radical or very innovative in themselves to be good, sometimes approaching issues from a different angle might do the trick, or from all angles at the same time. During the 2006 Finnish presidency of the European Union, the Health in All Policies (HiAP) was introduced to the EU policymaking (Health Inequalities Portal). The idea is simple and takes a ‘wide stance’ on health. Healthcare is a tough challenge and a huge cost for western societies. Since so many behaviors in a society affect the health of its citizens we should try to reduce negative health impacts everywhere possible. Consequently areas such as consumer behavior, environment and climate, agriculture, public safety, housing and education can be included when trying to tackle the healthcare challenge (EC web HiAP).

If policymakers have i) accepted the existence of the cognitive biases in the decision-making of citizens and ii) chosen a principle to use in terms of more or less libertarian or coercive paternalism, how and where should they apply these

decisions to work with these tasks? One suggestion drawn from the HiAP example above is to use a wider stance to reach maximal impact of the behavioral approach. Even though a behavioral approach would be instrumental rather than thematic (like the HiAP). There are a few ways to make sure that behavioral insights are widely considered when it is potentially feasible. The first way suggests that the legislative or executive body has to take an active stand on the potential for behavioral adjustments in almost every single policy. One downside to this way is that it risks to be neglected if falling into the trap of inertia. A less cynical but more practical problem is that policies and regulations that have already been decided will not be affected by the potential efficiency improvements. Another potential role for behavioralists in policymaking is to create a unit that acts as gatekeeper in areas with high potential for behavioral approaches. This unit could also act as a cold-case group and investigate the potential in old policies and regulations. The way in which any of these options could be implemented, however, has a lot to do with the characteristics of the political and administrative system in which it acts.

4 Methodological considerations

4.1 A Qualitative Case-study

As Yin (2003) suggests, case studies are often the preferred method when approaching *why* and *how* questions, when researchers has little control over events and when analyzing “a contemporary phenomenon within some real-life context”. Since all these three suggestions fit strikingly well with the purpose and subject of this paper, the decision to conduct a *case study* was fairly straightforward.

Data-wise this paper is an *analysis of documents*. When *collecting data* for this paper a lot of time was spent on understanding the population of potential cases and to collect all possible information regarding these cases. To gain understanding everything from newspaper articles to, policy documents, research papers and internal evaluations were processed. After having delimited and picked the cases for the study the collected data was narrowed down to a handful documents for every case. The main sources, from which the cases are analyzed, are official documents from the governments and administrations. These are, to a smaller extent, complemented with statements and texts from the officials that are not formally official. The unofficial documents are mainly used to gain an understanding on underlying explanations, but are *not* used to establish occurrences of actual events. A non-quantifiable processing of the material makes this paper a qualitative case study.

4.2 Case selection

The most crucial methodological consideration of this paper has been the picking of cases for studying the use of behaviorally informed policy-tools as a phenomenon. When scooping for possible cases it was soon obvious that potential cases were rather few and thereby selecting by, for example, random sampling would not be possible. The attention, in the search for valid cases, was then turned to a more purposive mode of sampling that borrows inspiration from the work of Jason Seawright & John Gerring (2008). In Seawright & Gerring’s paper “Case Selection Techniques in Case Study Research – A Menu of Qualitative and Quantitative options”, they explain how random picking (and pragmatic selection) many times is not possible to apply. At the same time as randomization has its insufficiencies, so do purposive case selection. Picking cases in a purposive manner will bring some intrinsic problems – often in play when working with

small-*N* samples – regarding the generalization of the cases and analytical results. Even so, purposive methods can, by helping researchers chose the most suitable cases, still make important contributions (Seawright & Gerring 2008: 295). The picking of cases in case-study research has the same two main objectives as random sampling due to Seawright & Gerring. To find:

- 1) A representative sample, and
- 2) Useful variation (in the light of the applied theory)

(2008: 296)

To be able to study the phenomenon of behaviorally informed policy-tools it is necessary to delimit the study to a few study objects. In this paper the choice was made to analyze objects that had not only been working with a single policy, but had a more concentrated use of behaviorally informed policy-tools. The picking of cases was conducted after the following characteristics/delimitations:

- 1) A use of behavioral approaches in the public sphere
- 2) Cognitive biases as a base for certain policy interventions
- 3) An explicit statement of the use of behaviorally informed policy-tools
- 4) A concentrated use of behaviorally informed policy tools
- 5) A time span, starting with the economic crises in 2007-2008 up until today (May 2014)
- 6) The behavioralist applications should have, or be directly tied to, legislative and executive powers in their respective administration

After this delimitation the three most *influential cases* were chosen from the rather small population. The picking influential cases is most commonly used when analyzing small or moderate sized samples and when the researcher has reasons to believe that results are being driven by a few cases (Seawright & Gerring 2008: 303f.). The reason for choosing influential cases is that the studied phenomenon here is fairly new and the population is small. There are reasons to believe that these influential cases are driving the development of the use of behavioral policy tools, today as well as some time ahead. The aim of selecting influential cases is “to explore cases that may be influential vis-à-vis some larger cross-case theory” (Ibid.). In plain English this means that if smaller and less influential cases were added in the investigated scope it would not change the results in any major way.

The reason for choosing *three cases* is a matter of balance between the limitations of words and the ability to analyze and present each case in a representative way. Together these delimitations *will* include many of the most influential applications of behaviorally informed policy and have as good representativeness as it gets, given the small population. An (unintended) effect of the case selection is that all the cases are western welfare states.

If having sufficiently met the requirement of a “representative sample”, as stated by Seawright & Gerring (2008: 296), the next step is to consider if there is any “useful variation” among the cases. Useful variation is an often unrecognized

requirement in case study research (Ibid: 294). The need for useful variation is naturally crucial if case-study researchers are trying to establish causal mechanisms. While this paper somewhat isolates factors – the biases derived from the use of behaviorally informed interventions – there has not been any intention to establish any causal explanations. This paper focuses on theories, ideas and strategies that have been boiled down to an essence by using the cases as a filter. The most interesting variation between the cases in this paper is the different roles the case units have in their respective administration, and what policy areas their competence includes. This will be discussed further in the introduction to each case.

The generalizability of the cross comparisons of the cases in this paper is most likely limited. It should also be emphasized that it is the phenomenon – behaviorally informed policy – that is in focus and not the cases themselves, making it more of a theory consuming study rather than a theory testing.

The three cases might not be very generalizable one by one, but together they should be attributed a fair amount of representativeness regarding how governments in the western welfare states (can) use behavioral policy-tools strategically. *When they use it*. As this could be considered a quite new approach by governments, it is however likely that new ways of applying behaviorally theories will emerge. This would of course temper with the generalizability of this paper and complementary studies would be needed to maintain representativeness. At the same time, looking at the most influential cases – that are likely to influence future work on the behavioral area – might give the study life for yet a while. The main contribution of this paper lies not in cross comparisons between the cases, but in the collected examples of in-practice use all together.

A note on ‘selection’ should also be made regarding *the cases within the cases* in this paper. Since the actors in each case have conducted many different behavioral interventions, how have the presented examples been chosen? The answer is that the examples are not directly analyzed themselves. Approaches and strategies are analyzed and the cases within the cases are mainly used as pedagogical means when communicating the analysis. Indirectly the case examples are analyzed as a part of the overall approaches and strategies.

5 The cases

5.1 The adoption into policymaking

So, how have the behavioral theories regarding the cognitive biases and the suggested solutions been received in government administrations around the world? Since many of the ideas are still fairly new it is likely that the ideas will, if not have a greater deal of influence, at least be ventilated more in political deliberations in the future. As an example, the European Commission has shown an increased interest in applying behavioral sciences to the EU policymaking, both by arranging recurrent international conferences¹⁴ on the theme and investigating the potential in in-house reports (Bavel et al. 2013).

As we have also seen in chapter 2 the status quo and the default heuristics make a possible answer to why business is conducted as usual also in policymaking. However, some administrations seem to have been able to challenge the status quo somewhat. In the US Barack Obama recruited Nudge-author Cass Sunstein to the White house under some controversy in 2009. The other Nudge-author Richard Thaler has been acting as policy advisor to Denmark, France and the UK. In the UK, Prime Minister David Cameron even established a *Behavioral insights team* in 2010. The UK should be considered the pioneer and the most extensive and influential user of behavioral approaches in policymaking to this date, by far. This makes the UK an obvious case and a natural starting point when looking into cases where behavioral approaches have been applied. The US is mainly considered an influential case because of Nudge advocate, Cass Sunstein, who has had the role, both as a theorist and a practitioner within the US context. Finally the smaller EU case, is picked because of its outspoken willingness to apply behavioral ideas and because their influential and reoccurring high level conferences on the subject.

5.2 Structure of the analysis

To connect the analysis of the empirical cases to the theoretical chapters (2 and 3), and later be able to compare the relevant parts of the cases, a well-structured case

¹⁴ The 3rd behavioral economics conference organized by the European Commission was held at October 30th 2013 (Mimica 2013).

presentation is crucial. The empirical cases chosen for this paper have been analyzed using 3 main questions:

- 1) How is the work with behavioral insight *justified* and what role does the behavioral applications have in the respective case?
- 2) On *which cognitive biases* are the behaviorally informed work based and how?
- 3) *Which debiasing approach* are used?

Question 1) is meant to give a short background to work on behavioral insights in each case, but also to connect – if possible – to the introduction chapter (1) regarding the economical crisis and alternative low-cost policy. The first question also addresses the purpose of the behaviorally informed work as well as the role of the behavioral applications in the different administrations. *Question 2)* is answered by analyzing the empirical material using the knowledge about cognitive biases presented in chapter 2. The biases are presented in a theoretical way but are drawn from practical work. Examples of such work are also presented briefly. *Question 3)* corresponds with the debiasing approaches that were presented in chapter 3 and the reasoning on soft versus hard paternalism as solutions.

5.3 The UK– The Nudge Unit

5.3.1 Birth, role and justification

In the Pensions Act 2008 the UK switched the default regarding occupational pension plans to a system where workers were automatically subscribed to a pension plan instead of having to sign up for a plan (opt-in). Workers can then choose to opt-out if they prefer. This action by the UK parliament, which came into force in 2012, could be seen as one of the first major steps taken on the grounds of behavioral theory and default options as a nudge in the UK. After having seen very bleak results when using information and advice, politicians turned to this approach as a solution to the ‘last mile problem’ (Dolan et al. 2010: 46).

The Behavioral Insights Team (BIT) – or the ‘Nudge unit’ – in the UK was set up in July 2010. Before David Cameron became Prime Minister it is said that he was a believer in the behavioral ideas put forward by Thaler and Sunstein in their 2008 bestseller ‘Nudge’ (The Economist 2014-02-07). After taking office, as Prime Minister, in May 2010 it only took the cabinet some two months to set up the BIT.

In 2009 the Cabinet Secretary and Head of the Home Civil Service¹⁵ ordered a review of the implications of behavioral theory for policymaking. The task was assigned to the Institute for Government who investigated how behavioral theory was currently being used and what challenges were related to it (Dolan et al. 2010: 6). The MINDSPACE framework, written by researchers in related subjects continued to be the overall guide to BIT in the initial years, and to a lesser extent it is still being used in 2014 (BIT 2014: 3).

The BIT works mainly together with other parts of the UK administration to investigate, test and change different interventions on different behaviors. From this aspect the BIT could be considered an inter-institutional task force or consultant. In February 2014 it was declared that the BIT would become a partnership outside of the UK government. The state will maintain the new company's biggest customer but it is now able to sell its services to other governments and administrations as well (Nesta web 2014, BIT 2014).

Why then did the UK government turn to the findings of behavioral science and why at this very time? In the MINDSPACE report (Dolan et al. 2010) we are given some answers to this question:¹⁶

- A time of fiscal constraints calls for new approaches in policymaking.
- New relevant research with strong potential
- New innovative ways to change behavior in areas that have previously proved to be hard to change (p. 13)¹⁷
- A cheap but potentially powerful way to change civic behavior to the better
- Soft paternalistic ways such as nudging is a “low pain way” to change civic behavior (p. 7)

5.3.2 Targeted biases

In the mnemonic “M-I-N-D-S-P-A-C-E” the most prominent biases, targeted by the UK, are concluded as¹⁸:

- ❖ **Messenger** – we are heavily influenced 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

¹⁵ Sir Gus O’Donnell

¹⁶ These bullet points (•) are a conclusion of the justifications and not a quotation from an actual bullet point list in the analyzed documents.

¹⁷ E.g. exercising and antisocial behavior.

¹⁸ These bullet points (❖) are present in the analyzed documents, but have been narrowed down to its essence.

- ❖ **Saliency** – our attention is drawn to what is novel and seems relevant to us
- ❖ **Priming** – our acts are often influenced by sub-conscious cues
- ❖ **Affect** – our emotional associations can powerfully shape our actions
- ❖ **Commitments** – we seek to be consistent with our public promises, and reciprocate acts
- ❖ **Ego** – we act in ways that make us feel better about ourselves

(Dolan et al. 2010: 8)

The BIT, in their work, initially used these nine elements then gradually discovered that it was a bit too complex (“to policymakers”) and that it, together with much of the academic literature, lacked some relevance when working with these theories in practice. The main effect that was lacking seems to have been the effect of simplification. Reducing friction was resulting in greater impact and the mnemonic was therefore reduced to four elements in the EAST framework (BIT 2014:3). The new mnemonic and its main elements want adopters to make interventions:

- ❖ **Easy** – use the power of defaults; reduce the ‘hassle’ to an action; simplify messages
- ❖ **Attractive** – attract attention by using images, color and personalization; well designed rewards and sanctions
- ❖ **Social** – draw on social conformity; use networks; encourage people to make a commitment to others
- ❖ **Timely** – prompt when people are receptive; consider the immediate costs and benefits; Help people plan their response to events

(BIT 2014: 3-6)

The analysis of the use of biases in the UK case will focus mainly on the updated and reduced EAST-framework and the work related to it. The framework can be explained by the cognitive biases theory presented in chapter 2 in the following way:

Easy – The BIT states that while the power of *default options* is widely exploited in the private sector – e.g. by the use of default settings on social networks sites or having unreasonably high ‘default tariffs’ on energy or telephone services – it is not taken into consideration by policymakers enough, yet. Automatic opt-in approaches in organ donation and pension plans have proven to have a substantial potential of increasing the impact. But, as the BIT points out, it can be highly debatable who should set such defaults and how. Because of this, vigorous discussions about the relation between the use of behavioral insights and democratic aspects is a welcomed occurrence (2014: 10f.).

Another important part of the ‘easy’ element is *simplification*. As we saw when discussing the scarcity mindset and tunneling in chapter 2, simplification tends to have positive effects on anyone but especially on people who are worse

off and might be stuck in scarcity traps. As we also saw, in the example of simplified forms for applications to financial aid for college and what is sometimes believed to be an information problem (which it often *also* is): simplification reduces the taxation of the mental bandwidth and increases the potential for individuals to make *more* balanced choices and decisions. The BIT works with this by reducing “hassle” and “friction cost” for users in relation to a certain action. Certain potential is seen in enabling consumers to more easily be able to switch suppliers (BIT 2014:12, 15). In the EAST framework it is, however, also pointed out that there are situations where simplification can have adverse effects. Such effects can sometimes also be used as a nudge. When regulation required paracetamol pills to be sold in blister packs the extra effort it took to get each pill out appears to have reduced the cases of paracetamol poisoning by 43 percent and deaths by 765 over a 11 year period (Hawton et al. 2013; BIT 2014: 12). A main conclusion by the BIT, drawn from their work with simplification, is that it is often fruitful to adopt changes in stages rather than trying to change an entire behavior at once from scratch (BIT 2014: 18).

Attractive – The attractiveness of a policy intervention mainly *appeals to* the intuitive and emotional parts of *System 1*. Successful commercials rarely aim at engaging the viewers reflective system. Also here the BIT states that the private sector is a much more frequent and extensive user of applying the biases connected to attraction. By personalizing messages, reframing the way in which incentives and choices are presented, the potential in designing attractive choice environments can be substantial also in the public sector. The framing of an option can potentially have a large impact on how people make their choices as discussed in 2.2.3.

In an effort to increase the number of tax-registered vehicles in the UK the BIT worked together with the Driver & Vehicle Licensing Agency on redesigns of messages sent out to targeted citizens. The agencies created two new letters and tested the efficiency of those compared to the old letter. The first new letter was simplified and more threatening while the second new letter also had a picture of the person’s car on it. While the simplified version had fairly little impact the attached picture increased tax registration from the previous 40 percent to 47 percent (BIT 2014: 20). The picture made the possibility of losing the car more salient than just threatening (with words) to take the car away. By combining the power of *framing* with the *emotional* reactions of *System 1* and *personalization*, the agencies were able to raise tax obedience with very small means.

Another way to make an action more attractive is to connect *rewards* or incentives to them (BIT 2014: 18). Such rewards does not have to be a direct result of the benefits of a program but can be completely unrelated to the program and only work as a motivation tool to get citizens to engage. Rewards like this can make individuals engage where they otherwise will procrastinate or tunnel out the engagement in such a program. As we saw in chapter 2 unrelated rewards such as a kilogram of lentils could have a major impact on the larger policy problem; to vaccinate children in development countries (Banerjee et al. 2010).

Social – Due to the BIT social influences often go unnoticed. However, if policymakers can get insights on how these social influences operate it would most likely lead to more efficiently designed policies.

Drawing on the *social information bias* (2.2.5) the BIT suggests working out nudges that appeal to the fact that ‘most other people’ behave in a certain way. In cooperation with HM Revenue & Customs the BIT have conducted a series of trials on this theme. By including information, on dispatches regarding tax payment, that most people had already paid their taxes, the tax payment rates increased significant (Hallsworth et al. 2014).

Making promises to friends or social connections can also be an efficient way to have people act in a certain way or to follow a program until its end (BIT 2014: 34). This mechanism can be considered in the light of the *peer-pressure bias* presented in 2.2.5. It has also been shown that individuals have a hard time meeting deadlines that are too far off in the future. If individuals set out part time deadlines and create material or social losses, if they fail to meet the deadlines, they are more likely to submit (Mullainatan & Shafir 2013: 22f.). This should be especially true if the loss would mean losing something the individual is already in possession of, as opposed to missing out on future gains (see 2.2.3 regarding *loss aversion*). Creating nudges in the form of ‘social deadlines’ should therefore be a potential way to induce behavior. Together with a job center in the UK the BIT introduced an intervention where job seekers were committing through written statements to their job seeking advisors. Interestingly enough, the BIT has also introduced a similar commitment device for the members of their own team (BIT 2014: 35).

Timely – By bringing in the aspect of timing, the BIT wants to put more emphasis on the way in which policies are implemented and not regard this as a matter of smaller details. The BIT is even contemplating the concept of testing different points of interventions experimentally to see when impact is best (BIT 2014: 37f.). Together with HM Courts Service the BIT conducted a trial where the effects of well-timed text messages were used as a way to increase fine payment rates. A moment was identified where there was still a chance to pay the fines and avoid a bailiffs procedure. The results showed that simply sending a well-timed text message doubled the payment rate and that if the messages were personalized the payment rate increased as much as three times compared to no text message at all (Heynes et al. 2013). Yet again we see that the use of a combination of biases seem to have a powerful potential when designing policy-interventions/nudges.

As we saw when discussing the scarcity mindset in chapter 2, the point in time of an intervention can be crucial to the potential impact of the intervention. An informative intervention, with the intention of educating citizens regarding a certain task, requires more mental bandwidth which makes the importance of the ‘when-factor’ even more tangible. The BIT points out that when identifying moments for interventions, periods of change in citizens’ lives should be considered scrutinously. When people are having a child, moving to another town or go to university they tend to be more receptive to interventions that can change their behavior. Since transition periods like these often also involve some contacts

with public administrations, the chances are often good to reach individuals at such times (BIT 2014: 39).

5.3.3 Debiasing approach

If the MINDSPACE and EAST frameworks have set out the theoretical approach to the UK's and BIT's work with behavioral approaches, there is another document that stated the methodological approach. The work of BIT is to a vast degree based on an *empirical approach* where behavioral theory and Randomized Controlled Trials¹⁹ (RCTs) are used to test behavior in a specific context. The BIT's own main strategy paper: *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*, very distinctively points out the use of RCTs as the main tool for identifying and strategically counter cognitive biases, mainly by the use of nudges (BIT 2012: ch. 1).

When analyzing the cognitive biases targeted in the BITs work, focus was mainly put at the EAST framework. However, when looking into the overall approach and strategy, most of the ideas have not changed since the MINDSPACE framework. In the final chapter of the MINDSPACE report it is explained how policymakers can put the presented behavioral strategies into practice and target the cognitive biases. Already in the early steps of the work with applying behavioral insights in a more specific way, the idea of *soft paternalistic methods* were suggested as the main approach in the UK (2010: 7).

Important to note is that there is an *emphasis on the complementary role of behavioral tools* rather than replacing such tools (Ibid: 10). While this complementary tool might be merely soft paternalistic it could i.e. be combined with coercive tools. Such coercive tools (legislation and regulation) however, goes outside of what the BIT is officially working with.

In the MINDSPACE report it is also reflected upon the potential legitimacy problem when trying to shift civic behavior. This connects directly to the theories regarding paternalism that were presented in chapter 3 and what kind of harm that calls for paternalistic interventions. The most prominent point made is that legitimacy is higher when trying to prevent harm to others (e.g. safer communities) than when trying to promote benefits to self (e.g. healthier lifestyle) (2010: 65). However, the obvious problem of being able to always separate an action that causes harm to others from an action that causes harm to self is not reflected upon. Leading an unhealthy lifestyle in a universal welfare state could, in the longer run, very well be seen as harm to others by undermining the premises of the very foundation of the welfare state. Reflecting further on self-harm in the report, the authors state that before conducting intervening nudges, that affect individuals, it should be proven that the targeted behavior is actually reducing the individual's well being. The final line, regarding the harm issue, is that *soft paternalism requires less justification than coercive paternalism* (Ibid).

¹⁹ Experimental studies.

Because of this, soft paternalistic tools can be seen as quicker and easier to use, something that can be important when policymakers look down in the policy toolbox.

So far trials and nudges have only been investigating civic behavior as a single group without any segmentation. In the future the BIT aims at creating trials that enables them to see how different social groups may react differently to certain kinds of interventions. The strategy initially also was to work with fairly simple behaviors because they were easier to measure. But moving forward with works on job centers etc. the unit aims at target more complex behaviors (BIT 2014: 50).

5.3.4 Concluding the UK analysis

- Justification: see bullet points in 5.2.1
- Mainly an *empirically based approach* to behavioral theory. Behavioral theory is used to test specific cases and then apply changes rather than applying wide theories directly into policymaking
- The BIT targets a number of biases: the status quo bias, social conformity biases, scarcity and tunneling biases, framing bias and emotional biases in System 1.
- *Soft paternalistic approach* or, more specifically, libertarian paternalism and nudges. Open for complementary soft/hard paternalistic strategies
- The BIT has a corporative role as an internal task force (up until now)

5.4 The US– The Bureaucrat Gatekeeper

5.4.1 Birth, role and justification

Much like David Cameron in the UK, it is said that Barak Obama has been a long time fan of nudging and behavioral insights to policymaking; even that President Obama is a “nudge politician” or “behavioralist” (e.g. NYRB 2008).

In 2008 nudge theorist and law scholar Cass Sunstein was elected to the post as administrator for the Office of Information and Regulatory Affairs (OIRA) by president Obama. The process of getting Sunstein into office was however everything but quick and easy since it is widely understood that OIRA is a little office with a big impact. While the libertarian-paternalism concept aims at reaching out to a bipartisan audience, many republicans seemed to have a hard time accepting Sunstein entering the White house on such a crucial position. This resulted in Sunstein being blocked by several republican senators in a row before finally being allowed to enter the White house as the OIRA Administrator in 2009 (e.g. New York Times 2009; Sunstein 2013).

OIRA is a federal office established by Congress in the 1980 Paperwork Reduction Act (PRA). It is part of the Office of Management and Budget, an agency within the executive office of the President. There are both political

appointees and civil servants working in OIRA. The head of the office is the OIRA Administrator. The formal duties of OIRA are set out, mainly, in the PRA and in executive order 12866 (Clinton 1993), which were reemphasized by Obama in executive order 13563 in 2011. Given in the PRA, OIRA has the task to review all collections of information carried out by the federal government. No federal authority has the right to issue any new paper work on the American people if OIRA has not approved of it. Under executive order 12866 the unit has the task to oversee draft regulation, which means in practice that OIRA has the power to delay or halt regulations. Finally, a less formal part of OIRAs work is its power to help shaping the president’s agenda (Exec. Order 12866; Exec. Order 13563; Sunstein 2013: 1-5).

To put the role of OIRA in the context of the legislative and executive powers, it is best seen as a hub between the legislative power (Congress) and the executive power, when dealing with regulations. When Congress has decided upon regulation there is sometimes a need to clarify certain aspects regarding the implementation before handing it over to the executive agencies. The relevant agencies and other stakeholders are also supposed to have a chance to give their opinion on a draft regulation. This is where OIRA comes in (Sunstein 2013: 30). In policy analysis the pros and cons of writing detailed policies compared to more ‘open’ policies are commonly discussed. Here units like OIRA acts as gatekeepers who (hopefully) reduce the hassle for the agencies that are assigned the implementation of the regulation.

Apart from the leading OIRA documents this analysis will also look into memorandums^{20 21 22} from OIRA and Sunstein’s own descriptions (2013) of the time as administrator from 2009 to 2012.

A relevant question is, of course, *why* the US government, through OIRA, chose to direct efforts towards a behavioral perspective on regulation. *If* they actually did so and *how* are dealt with in the sections below, but the main reasons for this direction seem to be:

- Often more efficient than direct regulation
- *Low cost* way to change behavior (Exec. Order 13563)
- *Low pain* way to change behavior, by maintaining freedom of choice (Ibid)

5.4.2 Targeted biases

In the case of OIRA and the US government, the work on behavioral theory is not stipulated directly in the main documents (see 5.4.1) that state OIRA’s duties. However, the orders leave room for the potential use of such ideas. There are parts

²⁰ Memorandum on ”Disclosure and Simplification as Regulatory Tools” (OIRA 2010/6).

²¹ Memorandum on “Executive Order 13563, ’Improving Regulation and Regulatory Review’” (OIRA 2011/2).

²² Memorandum on “Informing Consumers through Summary Disclosure” (OIRA 2011/9).

of executive order 12866 that are more or less implying the use of behavioral theory:

Each agency shall identify and assess available alternatives to direct regulation, including providing economic incentives to encourage the desired behavior, such as user fees or marketable permits, or providing information upon which choices can be made by the public.

(Section 1, principle 3)

In executive order 13563 President Obama enhances the wish for nudge-tools to be used, when possible, by stating: “[E]ach agency shall identify and consider regulatory approaches that reduce burdens and maintain flexibility and freedom of choice for the public”. Examples are also presented in the order pointing out *default rules* and *disclosure requirements*. To reduce burdens and maintain freedom of choice is almost the definition of a nudge as explained in 3.2. The formulations correspond very well with libertarian paternalism and the use of nudges as a less direct and less coercive form of policy. The expressions in executive order 12866 and 13563 are translated into more behavioral terms in OIRAs memorandum 2010/6. There are mainly two ways that OIRA advocates to adapt to cognitive biases: *simplification* and *disclosure*.

Simplification could, in some sense, be seen as an OIRA duty even before an apparent behavioral perspective was added through a memorandum in June 2010 (OIRA 2010/6). These elements are explicitly drawn from the statues of executive order 12866, but are also forged together with a behavioral theory derived from cognitive biases. Even though “simplification” in itself might not be a very specific nudge, it can include other more specific nudges. One powerful nudge that OIRA have put under the simplification banner is the use of “default rules, such as automatic enrollment, to simplify people’s decisions” (2010/6: 9). Agencies should consider, what would be the outcome if citizens make no choice at all in a specific situation. The argued use of default rules to simplify is justified by drawing on human flaws such as procrastination and inertia. The default rule also has potential to be regarded as the endorsed option. The bottom line of the recommendation from OIRA is to use default rules or options – as the desired alternative – together with an opt-out possibility (Ibid.). Procrastination is closely linked to the tunneling issue presented in 2.2.7 and the power of inertia in relation to endorsed default options is explained in 2.2.5.

Another side of the simplification element is that simplification can have crucial impacts on whether or not people are engaging in certain programs. If application procedures are to daunting or complex many citizens are likely to not participate or enroll. This is another reason why OIRA are focusing on simplification (OIRA 2010/6: 10). As discussed both in the theory section (2) and in the UK case, simplification of forms and applications to programs are likely to be helpful to anyone, but especially helpful to people who are worse off, have low income or no experience of tertiary education in the family.

Disclosure (as a regulatory tool) should not be seen as having any real intrinsic nudge-value regardless of what kind of disclosure we are talking about.

Just pouring out information in society is not a meaningful nudge. If designed in a proper way, by using behavioral insights, it could however have good nudge potential:

People have limited time, attention, and resources for seeking out new information, and it is important to ensure that relevant information is salient and easy to find and to understand. There is a difference between making a merely technical disclosure — that is, making information available somewhere and in some form, regardless of its usefulness — and actually informing choices. Well-designed disclosure policies are preceded by a careful analysis of their likely effects.

(OIRA 2010/6: 3)

The memorandum continues by explaining two main types of disclosure: *summary disclosure* and *full disclosure*. A summary disclosure is often related to the purchase of a product, such as nutrition facts on packages, while full disclosure is what agencies often use in an act of transparency (OIRA 2010/6: 3). In the memorandum, signed by Cass Sunstein, OIRA wants agencies to contemplate over their use of disclosure and, by doing so, increase the efficiency when facilitating new or changed regulations.

If simplification and disclosure are the two main type of nudges used by OIRA and agencies connected to them, *how are these nudges explained by the theories on cognitive biases, presented in chapter 2?* Starting off, simplification and disclosure are interconnected. A user-friendly disclosure of information is most likely a rather simple one, at least when viewing it from a citizen's perspective. As mentioned above **simplification** is helpful to any one, but often especially to the people who are worse off in society. Simplified procedures can engage more people from groups that are considered "worse off". This draws on the *scarcity mindset* and the concept of *tunneling* presented in 2.2.7. Another part of the simplification principles deals explicitly with *default option* nudges drawn from the status quo bias presented in 2.2.5. While a default option would not have any advantages over other alternatives in the eyes of a rational agent (Econ), it does have significant advantages when an actual person is making a choice – or abstains from making any choice.

Disclosure on the other hand can be used as a simpler form of nudge. As Sunstein writes: "Summary disclosure should be designed for System 1, not System 2" (2013: 93, 79). In the US case disclosure as a nudge draws on mainly a few cognitive biases. First there is the rather obvious advantage that disclosure might have on individual decision-making (if the information is adopted): choices are more likely to be made in System 2 (reflective) than in System 1 (intuitive). A well-designed disclosure can then, potentially, induce individuals to make an informed and deliberated choice and bypass the biases of System 1.

Ironically though, certain types of disclosure can target the shortcomings of System 1 and instead induce decisions on a more emotional basis (much like some commercials does). An example of this, as suggested by OIRA, is the use of vivid descriptions and persuasive images when disclosing information (2010/6: 4). In

practice this strategy has been used in many countries when putting on horrific images on cigarette packs.

Yet another cognitive bias, targeted by OIRA, is the power of framing disclosed information in a smart way. This means to, when opportunity is given, use the fact that Humans are *loss averse*. By framing potential outcomes as a loss rather than a gain it will attract more attention and citizens are more likely to engage (OIRA 2010/6: 4).

Much like in the BIT, OIRA sees timing of interventions, like disclosure, as a crucial element to facilitate regulation or behavioral change directly. This can mean to consider how to get the fuel-economy information out in the right moment just before a purchase of a car (OIRA 2010/6: principle a4).

5.4.3 Debiasing approach

When entering the position of OIRA administrator in 2009 it was no secret from Sunstein's side (or to any one else) that the work of OIRA would be directed a bit towards nudging and libertarian paternalism, when opportunities were given (Sunstein 2013: 2). This can also be established by looking at the more formal work of OIRA where there are obvious pushes for the use of *soft rather than coercive paternalistic* interventions when trying to reach policy goals.

As cited above, from executive order 13563, “[E]ach agency shall identify and consider regulatory approaches that reduce burdens and maintain flexibility and freedom of choice for the public”. This is pretty much as close as it gets without putting the actual labels ‘nudge’ or ‘libertarian paternalism’ on the approaches. OIRA on their hand expresses the use of specific nudges like default rules.

To promote regulatory goals, agencies should consider whether it is appropriate to use default rules (such as automatic enrollment) as a substitute for, or as a supplement to, mandates or bans.

(OIRA 2010/6: 10)

The use of softer means – as compared to bans – are justified by their ability to *maintain freedom of choice* and to *allow a diversity* where individual circumstances makes a ban unlikely to fit everyone. It is also considered a *low cost* intervention compared to mandates. Default options are especially fruitful when “agencies have reason to be confident about the appropriate default rule, and when preferences and situations are not relevantly diverse” or when the technicalities and complexity implies that citizens will lack experience and expertise to make an appropriate choice (Ibid: 10f.).

At the same time, the *complementary role of nudges* is pointed out. Soft paternalistic interventions – like nudges – are not a complete solution that is always optimal, but one that should seriously be considered in many cases (Ibid.). Even if coercive means – which require citizens to engage by law – are used, this is no insurance against people still not obeying that law for different reasons. This

is where coercive and libertarian paternalism can have an important complementary potential.

As an alternative to default rules, with an automatic opt-in, OIRA argues for the use of *active choosing* as a nudge. This approach is recommended when there is a lack of knowledge about what is a preferred choice for a certain population²³ (OIRA 2010/6: 11; Sunstein 2013: 119ff.). An active choosing procedure requires, or even coerces, individuals to make an active choice. Since there is no way to avoid making a choice, there is no default option. How should people then be forced to make active choices? Suggestions claim that it could be solved by connecting these kinds of choices to a final step before getting certain licenses, benefits or before being allowed to enter a new job. Before getting your driver's license you will have to make an active choice on organ donation etcetera (Sunstein 2013: 119). But then again, outcomes will depend heavily on how the questions are framed and which default is used when framing the question (see 2.2.3; 2.2.4; Kahneman 2011: 373).

In section 3.4 there was a presentation of different types of choices. Mainly two choices were highlighted: the aggregated choice and the personalized choice. What is suggested in the US case is that *default options should be used when deciding on aggregated choices and active choosing should be applied when dealing with more diverse personalized choices*. Obviously there are no hard lines made and this should be regarded as more of a thumb rule. There are also obvious limitations to the use of active choosing because of its mandatory nature.

One important difference between the BIT and OIRA is that the former works mainly with nudges and behavioral approaches while the OIRA have wide possibilities to use these approaches but they are not delimited to only such tools. On the other hand, OIRA do not have the capacity to conduct their own behavioral experiments like the BIT does. OIRA is limited to suggest to other parts of the public administration to, where it is possible, conduct experiments to guide their work (Exec. Order 12866: 5; OIRA 2010/06). This makes the US a case of *theory-based approach* rather than an empirical approach such as in the UK.

5.4.4 Concluding the US analysis

- Soft paternalistic interventions are *low cost, low pain* ways that may be *more effective than coercive regulation*
- *Theory based* approach; no direct capacity of conducting trials
- Targeted biases by OIRA is mainly: the status quo bias, loss aversion and framing bias, tunneling bias and System 1 and System 2 biases (through disclosure)
- *Simplification, defaults and disclosure* in focus
- The role of a *gatekeeper* and *gentle bureaucrat*

²³ Population as a group of citizens, not the population of a whole country.

5.5 The EU – The Protective Behaviorist

5.5.1 Birth, role and justification

The EU was in fact, in a minor way, one of the first adopters of behavioral approaches when the European Commission (EC), in October 2008, challenged the default-option rational by suggesting a ban on the use of pre-checked boxes in consumer contracts. Extra charges when purchasing online products were i.e. swapped from sometimes being pre-checked (default) to always being unchecked, requiring an active consumer choice (EC MEMO/11/675; EC 2008/0196).²⁴

The EU does not have a specific application working with behavioral insights. The EU does, on the other hand, have a number of initiatives that derives explicitly from behavioral findings and theories. Since it is the Commission that has the power to both initiate and execute regulation it is here we find the work informed by behavioral theory in the EU. Most of the work has been conducted by the Directorate General for Health and Consumers (DG SANCO). Since both health policy and consumer protection policy are areas with large theoretical potential for the use of behaviorally informed regulation, this was no accidental starting point. The approaches have since spread to other parts of the Commission. Consumer behavior remains the most prominent area for the use of these approaches in the EU, whether it is DG SANCO, DG Environment or DG Research and Innovation that use it for their respective perspectives (Ciriolo 2011: 2). Both studies and behaviorally informed policy initiatives are expected in the near future due to the Commission's in-house science service, the Joint Research Centre (JRC) (Bavel et al. 2013).

The role of the behaviorist applications in the Commission has so far been very thematic. As Ciriolo points out, *consumer behavior* is the lowest common denominator when it comes to behaviorally informed work in the EU (2011: 2). This likely derives from the potential to use behavioral theory in this specific policy area. Consumer policy, however, has the potential to affect many other, for the EU, relevant policy areas such as health and environmental policy.

Presented by the European Commissioner for consumer policy, Neven Mimica, there are a number of challenges that behavioral insights can meet that justifies its use that:

- An increased diversity of products and services make the number of choices each individual has to make overwhelming

²⁴ The directive was formally adopted by the Parliament and member states in October 2011 (EC MEMO/11/675; EC 2011: art. 22).

- The time in which a purchase can be made is today down to seconds
- The importance of enabling consumers to feel confident, empowered and to have a fair chance within the competitive environment in which they make their choices.

(Mimica 2013)

5.5.2 Targeted biases

The EUs focus on consumer behavior and protection is easily recognized in their in-practice use of behavioral theory. These are their main interventions draw from behavioral theory on cognitive biases.

As mentioned initially in this section, EC drew on behavioral theory when **banning pre-checked boxes** for additional costs in consumer contracts. It was in the Consumer Rights Directive (EC 2011/83: article 22) that the Commission introduced this regulation, meaning no additional costs – as a result of pre-checked boxes – are a legally binding part of the contract and should therefore be reimbursed (Ibid). The action taken by the EU targeted (among many other things) the exploitation, by traders, of the *status-quo bias*. By forcing traders to have all ‘additional-cost boxes’ unchecked the behavioral theory tells us that this will likely have a major impact on the number of additional costs paid.

The next behaviorally informed intervention has to do with something as specific as **web browsers** for PCs. For a long time software manufacturer Microsoft delivered their operating system Windows with their own web browser Internet Explorer preinstalled [as the default option]. The Commission was – in line with their antitrust efforts – uneasy with the potential consequences of what Microsoft’s market dominance in the PC market would lead to. They argued that such dominance deprives consumers of choice. In 2009 Microsoft offered to remove this barrier for competition and in 2010 users were supposed to face the new *active choice design* and face a new “choice screen”²⁵ when launching their Windows Internet application. In 2012, the Commission expressed criticism on Microsoft’s non-compliance on the legally binding agreement from 2009 and in 2013 the Commission fined the company €561 million for the non-compliance (EC web 2013; EC 2013/6; EC 2012/10). Instead of a pre-checked box, this approach requires consumers to make an active choice. Evidently this nudge intervention draws on the fact that the *status quo bias* will cause many consumers to stay with the default browser; not because it is their preferred choice but because it is the endorsed default that requires no effort to use. An active choice approach, to counter the status quo bias, is likely to be beneficial if the choice is not too complex and in situations where users can be forced to make an active choice (cf. 2.2.5 and 3.4).

²⁵ The choice-screen makes it possible for users to choose one or more out of 12 possible web browsers. The browsers will be presented with information from their producers. The choice screen shall be available for 5 years (EC web 2013).

In the line of harmonization initiatives from EC, the Health and Nutritional Claims Regulation (No. 1924/2006) was presented in 2007 and aimed at regulate the reference points [or anchoring points] for certain **nutritional claims** on products and in marketing campaigns. Among other streamlining features for the inner market, this regulation targeted the *framing bias* where products with 20% fat previously could have been presented as 80% fat free (Ciriolo 2011: 2; EC 1924/2006). In this case EC did not, itself, use the fact that humans are susceptible to framing. Instead EC countered commercial exploitation of the framing bias. Since most of the EUs work with behaviorally informed policies have been to stop this kind of consumer exploitation, drawn on default and framing biases, ECs in-practice strategy can be concluded as *mainly protective*.

Another side of the Health and Nutritional Claims Regulation is the obvious disclosing purpose. By regulating the use of *summary disclosures* allowed on products the consumers will have an easier task when making informed choices. *Simplifying* and presenting comparable information on foodstuff makes it possible to make reflective *System 2* choices. If there is no comparable information consumers can either put a lot of effort into making informed choices, or just make intuitive choices in *System 1*.

5.5.3 Debiasing approach

The European Union has through the Commission started to work with behaviorally informed policy. While EC have put out calls on behavioral experiments both in their framework programs (e.g. Consent 2013) and internally, they have not yet used experimental methods such as pilot studies to guide their work on behavioral policy-tools. The EUs take to the use of this kind of tools should thus, so far, be described as a *theoretical approach*.

The theoretical approach used is mainly the soft paternalistic but there are also hard paternalistic elements, such as in the web browser case. While this behaviorally informed action, by the Commission, is justified through the importance of fair competition on the European market, it is a fairly *coercive intervention*. It does increase the number of choices for consumers but it also, for sure, changes the economic incentives for the targeted trader (Microsoft). Therefore it cannot be considered a nudge, but should be regarded more as a *hard paternalistic regulation*. As discussed in section 3.5 not all cases are obvious examples of a soft respectively hard paternalistic approach. The ban on pre-checked boxes in 5.5.2 could possibly be considered an ‘in-betweenner’, to some traders this could force them to completely rearrange their marketing strategies. Since policy actions like this are carried out primarily for the sake of harmonization of the European inner market, the aim is to level the playing field for all traders to avoid asymmetrical competition between EU countries.

5.5.4 Concluding the EU analysis

- *Protecting and empowering consumers* in their choice environments
- Thematic role with *consumer behavior* in focus
- EU challenges the *framing bias*, the *status-quo bias* and the *exploitation of System 1* decisions
- *Mainly protective* use of nudges or 'counter nudges'. However, some interventions are coercive by law for traders
- *Theoretical approach* when applying behavioral insights

6 Conclusions

The roles of the behavioral applications in the different cases differ quite a bit. What was not as expected was that also the justifications – *or the answer to the why-question* – differed quite a lot. The behavioralist applications in the cases of this paper have chosen to work mainly with soft paternalistic nudges and the most profound justification for that is to *maintain citizens freedom of choice*. Another important justification is that it is an often inexpensive way with large potential to change civic behavior. This connects to the reasons for applying behavioral insights to policymaking at this very time, 2007-2014. Just as a non-legislative policy is likely easier to find democratic support for, a low cost policy tool is more likely to be voted through, and thus easier to use. Since the economic crises of 2008 an inexpensive policy tool could be found attractive, both out of economic and democratic bargaining reasons. The most evident examples of hard paternalistic interventions are found in the EU case where economic incitements for traders are somewhat sacrificed for the good of market competition and empowerment of consumers.

Moving over to the *how-question*; the most prominent interventions, drawn from behavioral insights on cognitive biases, can be concluded as: *default settings*, *simplification* and *smart disclosure*. The use of default options or default rules, as a way to counter the status quo bias, have in some cases been put aside in favor of an active-choice procedure that acts as a nudge on the same bias. Simplification and smart disclosure have a more complex connection to the biases. *Simplification* is a way to reduce unnecessary taxation of citizens' mental bandwidth, which is especially important when trying to engage underprivileged groups or the uneducated in policy programs. *Smart disclosure*, a behaviorally informed use of summary disclosure, should be designed for the intuitive System 1, be easy to understand and compare to other products, services or alternatives. A smart disclosure also includes appealing to the framing bias and sometimes also to social conformity biases.

As shown in the cases, behavioral insights on cognitive biases can be used as a way for governments or public administrations to:

- Use *engaging nudges* to gain attention and engage citizens in programs – as in the case with OIRA and their advocating of the use of framing and loss aversion when formulating messages to citizens. Or as in the case with default options, used both in the US and in the UK.
- Regulate commercial exploitation of the biases through *protective interventions* – as in the EU case when regulating the use of framing in marketing concepts such as “fat-free” or when safeguarding the market competition and consumer choice in antitrust cases like the web browser case.

Behaviorally informed policy tools, or at least framed in this way, is a fairly new concept and many of the impacts from its interventions cannot be properly evaluated yet since the impact requires yet some time to set in the systems. However, we can already see some lines of strategy taking shape where behavioral insights have been applied to policymaking.

7 Further discussion

7.1.1 Democratic aspects

While not being a main concern for the topic of this paper, there are interesting reflections to be made in the outskirts of the topic. There is an obvious distinction in how for example OIRA and BIT are working that could be analyzed in the light of the respective democratic system. While the UK government has a large degree of power of action in their own hands, the US government is to a much larger extent depending on bipartisan agreements. This has likely caused a sensitiveness regarding the work of OIRA, especially in its role to influence the precedent's agenda. This of course makes the analysis of OIRA's role as presidential adviser very hard and treacherous.

A more overall point regarding the democratic aspect has to do with politics and accountability of politicians. Using behavioral theory by empirically testing the policy outcomes before applying them on the whole population has one major upside and one major downside to politicians (agents). The upside is that since experimental testing like RCTs are widely considered to be the ultimate evaluation method when assessing impact and causal effects, politicians can just raise their shoulders and say "we tried our best", if a policy intervention goes south. On the downside, and maybe a bit more cynically – politicians may sometimes have a need to create policy that they know are not very good policy – because it might be a great political or rhetorical move. If an RCT is necessary to conduct before making such a move it could act as an unwanted gatekeeper from the perspective of the politician. More research on these and other democratic and policy process aspects should be fruitful in the future when the behaviorally informed interventions have had some time to set in the administrations.

8 Executive Summary

8.1 Subject and outline

During the last few decades, research on human behavior – or more specifically the mapping of cognitive biases – has made some groundbreaking discoveries. The start is often attributed to the research of Amos Tversky & Daniel Kahneman in their 1974 work *Judgment under Uncertainty: Heuristics and Biases*. These biases are systematic errors of our minds and do not only affect some unfortunate individuals with bad character. They affect all of us; usually without us knowing.

The implications of cognitive biases can be dealt with both on an individual and a collective level. It is the collective way to handle these biases, through policy and regulation, that is focused in this paper.

The year 2008 is often considered the start of the economic crisis that still today haunts citizens and governments around the world. The same year scholars Richard Thaler & Cass Sunstein presented a basic manual on how to apply behavioral insights to policymaking, in their book *Nudge – Improving Decisions About Health, Wealth and Happiness*. The ideas from this book are likely, by far, the most influential theory on cognitive biases that have been adopted and applied by governments. The approach argued for by Thaler & Sunstein is a soft paternalistic version entitled *Libertarian Paternalism* (2003, 2008). The libertarian paternalistic solution to the cognitive biases is to use nudges as smaller changes in citizens' choice environments. If strategically picked, such small changes could potentially have major impacts on civic behavior. To be defined as a nudge, an intervention has to be *choice preserving* and *non-coercive*. Other scholars, like Sarah Conly (2013) have argued that the knowledge on cognitive biases calls for more drastic policy-implications and government actions. While soft paternalistic actions might be able to save many citizens from making crucial biased decisions regarding their lives, it will still be possible for citizens to make such inferior decisions. Thus, a number of citizens *will* still make inferior decisions and suffer the consequences. Conly calls for a coercive response to this problem and argues that we cannot sacrifice these failing citizens for the greater good of maintaining the illusion of autonomous individual free choice.

Regardless if we prefer a more paternalistic or a mainly libertarian approach, the inherent potential of cognitive biases and its policy-implications calls for multidisciplinary research on the subject. Social and political scientists, natural scientists as well as philosophers need to converge and contribute with their angles to the bigger picture. We might not have stumbled upon a magic cure to social problems but if, how and where these behavioral insights can be used in a valid way, deserves to be examined thoroughly.

By conducting case studies of the main adopters of behavioral theory in the UK, the US and the EU there is a hope that this paper can somehow contribute to the examination of the potential of behaviorally informed policy tools. The main contribution consists of scrutinizing the discrepancy between theory and practical use of the cognitive biases.

This paper is focused on the policy implications of cognitive biases and how these can be, and have been, ‘debiased’ in policymaking. How such work *can be* conducted is answered by presenting a theoretical framework that forges together scientific knowledge, on the cognitive biases (ch. 2), with philosophical approaches on paternalism (ch. 3). The question of how such work *has been* conducted is answered by looking at in-practice use, of this kind of behaviorally informed theory, in governmental institutions that have chosen to apply it to parts of their policymaking.

The research question for this paper is:

What ideas and strategies, based on the knowledge on cognitive biases, have been applied, in practice, by government administrations in their policymaking?

This question is answered by breaking it down into two parts:

- 1) *Why* do administrations choose to apply behavioral approaches into policymaking?
- 2) *How* have government administrations been working with behavioral theory as a policy tool?

The questions are answered by analyzing public documents together with some complementary documents from previous research and narratives from the work of the studied units.

8.2 Biases

It is hard to face the fact that we, humans, are not completely rational. This questions the notion about human behavior that have dominated and rooted itself in our culture for centuries. The good news however, is that we are sometimes rational and make pretty good judgments to guide our decisions. So where do we go wrong?

One crucial feature of a rational agent is that it has preferences that are consistent with its behavior. If the rational agent diverges from its preferences it is because it lacks necessary information or the practical capacity to reach its preferred goals. A more humanlike creature, on the contrary, often makes the same mistakes as the rational agent, but also sometimes fails to align its behavior to its preferences.

The now established *dual-process theory* regards the human thinking in two different systems; the intuitive *System 1* and the reflective *System 2*. System 1 is conducting most of our decisions and can never be turned off. System 2 sets in when we face more complex tasks or computations and is often bothered by the intervening emotions and intuitive arguments of System 1. It is in System 1 and in the fight between the systems that we find many of the cognitive biases. Researchers studying cognitive biases have found hundreds of different biases that make humans chose inferior actions. Since this paper is focused on the practical use of the bias-related insights, the number of biases included has been narrowed down. The relevant presented biases needed to explain the work and justifications conducted in the three cases are:

- *The status quo bias* – people are systematically overvaluing passive alternatives, which are disproportionately common compared to alternatives that require action.
- *Scarcity and tunneling* – when facing scarcity of money, time or social life, people are getting increasingly effective on managing that scarcity. However, at the same time other parts of life are being tunneled out.
- *Framing and loss aversion* – The same alternative pitched in different ways gets chosen disproportionality, especially if one alternative is pitched as a loss.
- *Social conformity* – we use fellow citizens to inform ourselves, we are also sensitive to peer pressure. This applies even if it leads us to unreasonable actions.

8.3 Results

8.3.1 The how-question

The dominant approach to adapt to cognitive biases is, by far, the use of soft paternalistic nudges. Especially in the EU case there are however interventions with elements of a more coercive nature. The coercive elements have not been used on civic behavior, but to regulate private market actors. There is also a general believe among the cases that combinations of soft and coercive paternalist interventions might often be a smart strategy. Altogether the following findings answer the **how-question** of this paper:

- The *status quo bias* is widely appealed to; mainly by the use of delicately designed default options, but also by the use of active-choice procedures. *Default options* (or rules) tend to be used for complex cases where individuals might have diverse preferences and would have needed expertise knowledge to make an optimal choice. *Active choice* procedures tend to be used for simpler choices where individuals easily can gain understanding of the overall picture. Active choice however, requires

choice architects to have a snag on the targeted citizen since there can be no passive or default choice.

- Strategic *simplification* is another common strategy, especially in the UK and US cases. Citizens are facing thousands of minor decisions each day, all of which are taxing their mental capacity to some extent. Simplification can help individuals in general to make better choices, but seems to be especially effective when targeting underprivileged citizens. If trapped in *scarcity traps* their mental capacity can get heavily taxed and risks falling into a slippery slope where one biased decision makes yet another inferior decision more likely.
- A third group of bias-related interventions used in practice is *smart disclosure*. Smart disclosure is designed by: appealing to *System 1* with summary disclosures; regarding the *framing and loss aversion biases* and in some cases the *social conformity biases*. Another part of smart disclosure can be to mind the *tunneling bias* and carefully select the timing of a planned intervention.

It is also noteworthy that governments tend to use insights drawn from cognitive biases in two main ways:

- 1) To *engage citizens* in policy programs and gain attention for certain issues.
- 2) To use *protective interventions* that prevent other actors, often traders, from exploiting biased decisions by citizens.

8.3.2 The why-question

The justifications for the use of behaviorally informed policy tools were investigated to answer the **why-question** of this paper. The behavioral applications studied in the three cases differ a bit in their roles, work and justifications.

In the UK case the *Behavioral Insights Team* have acted as an internal taskforce with its base in the Cabinet Office. Cooperating with other authorities the team has used Randomized Controlled Trials on everything from paracetamol poisoning to tax obedience. The justifications for the use of soft paternalistic interventions are:

- A time of fiscal constraints calls for new approaches in policymaking
- New relevant research with strong potential
- New innovative ways to change behavior in areas that have previously proved to be hard to change
- A cheap but potentially powerful way to change civic behavior to the better

- Soft paternalistic ways such as nudging is a “low pain way” to change civic behavior
- Soft paternalistic interventions require less justification than coercive interventions

In the US case the federal *Office of Information and Regulatory Affairs* (OIRA) has acted as a bureaucrat gatekeeper with distinct elements of behaviorally informed strategies. OIRA’s work drawn from cognitive biases is mainly theoretical and the unit does not have capacity to conduct trials to lead their policy work like in the UK case. OIRA does however have the potential to induce agencies and authorities to act in certain ways and follow certain procedures. Through its position in the executive office of the President, OIRA also has the potential to influence the Presidents agenda. The justifications for the use of behaviorally informed nudges are:

- Often *more efficient* than direct (coercive) regulation
- *Low cost* way to change behavior
- *Low pain* way to change behavior, by maintaining freedom of choice

In the EU case the *European Commission* (EC) have worked with behaviorally informed tools mainly in the area of consumer behavior and consumer protection. EC are also arranging reoccurring conferences on the subject and there are explicit ambitions to extend the use of behaviorally informed policy to other policy areas. Because of the thematic use of behavioral insights so far, the justifications rest on grounds that differ a bit from the other cases. What justifies the use is that:

- An increased diversity of products and services make the number of choices each individual has to make overwhelming
- The time in which a purchase can be made is today down to seconds
- The importance of enabling consumers/citizens to feel confident, empowered and to have a fair chance within the competitive environment in which they make their choices.

The behavioral insights, drawn from cognitive biases, and its in-practice adoption into policymaking needs to be examined further. Most of the impacts from behaviorally informed interventions cannot yet be evaluated due to its recent implementation. The results of such evaluations will, together with continuing discussions on the philosophical aspects of paternalism, constitute a vindication for future use of behaviorally informed policy-tools.

9 References

- Allison, Graham & Philip Zelikow, 1999 [1971]. *Essence of Decision: Explaining the Cuban Missile Crisis*. 2nd ed. New York: Longman.
- Asch, Solomon E., 1955. "Opinions and Social Pressure". *Scientific American*, Vol. 193, No.5, pp. 31-35.
- Banerjee, Abhijit Vinayak, Esther Duflo, Rachel Glennerster & Dhruva Kothari, 2010. "Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives". *British Medical Journal*, Vol. 340, No. 7759, p. 1291.
- Bavel, René van, Benedikt Herrmann, Gabriele Esposito & Antonios Proestakis, 2013. "Applying Behavioural Sciences to EU Policy-making". *Joint Research Center (JRC) Scientific and Policy Reports*. Brussels: European Commission.
- Berns, Gregory S., Jonathan Chappelow, Caroline F. Zink, Giuseppe Pagnoni, Megan E. Martin-Skurski & Jim Richards, 2005. "Neurobiological Correlates of Social Conformity and Independence During Mental Rotation". *Biol Psychiatry*, No. 58, pp. 245-253.
- Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos & Lisa Sanbonmatsu, 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment". *The Quarterly Journal of Economics*, pp. 1205-1242.
- BIT 2014= UK Cabinet's Office's Behavioral Insights Team, 2014. *EAST – Four simple ways to apply behavioural insights*. [Electronic] Available: <http://www.behaviouralinsights.co.uk/publications/east-four-simple-ways-apply-behavioural-insights>. Downloaded: 2014-08-11.
- Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue & Matthew Rabin, 2003. "Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism". *University of Pennsylvania Law Review*, Vol. 151 No. 3, pp. 1211-1254.
- Cappelen, Alexander, Ottar Mæstad & Bertil Tungodden, 2010. "Demand for Childhood Vaccination – Insights from Behavioral Economics". *Forum for Development Studies*, Vol. 37, No. 3, pp. 349-364.
- Ciriolo, Emanuele, 2011. "Behavioural economics in the European Commission: past, present and future". *Agenda*. Brussels: Oxera.
- Conly, Sarah, 2013. *Against Autonomy – Justifying Coercive Paternalism*. Cambridge: University Printing House.
- Conly, Sarah 2013b= Sara Conly, 2013-03-24. "Three Cheers for the Nanny State". *The New York Times*. Opinion Article. [Electronic] Available: <http://www.nytimes.com/2013/03/25/opinion/three-cheers-for-the-nanny-state.html>. Downloaded: 2014-08-11.

- Consent 2013= Consent project web page, 2013. Research project under the European Union's 7th Framework Program (FP7). [Electronic] Available: <http://www.consent.law.muni.cz/>. Downloaded: 2014-07-24.
- EC 2011/83 (COD)= European Commission, 2011. "Directive 2011/83/EU of the European Parliament and of the Council on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council". Brussels: European Commission.
- Dolan, Paul, Michael Hallsworth, David Halpern, Dominic King, Ivo Vlaev, 2010. *MINDSPACE – Influencing behaviour through public policy*. United Kingdom: Cabinet Office through the Institute for Government. [Electronic] Available: <http://www.behaviouralinsights.co.uk/publications/mindspace>. Downloaded: 2014-08-11.
- EC 1924/2006 (COD)= European Commission, 2006. "EU regulation 1924/2006 on Nutrition and Health Claims made on Food". Brussels: European Commission.
- EC 2008/0196= European Commission, 2008. "Proposal for a Directive of the European Parliament and the Council on Consumer Rights". COM(2008) 614 final. Brussels: European Commission.
- EC 2012/10= European Commission, 2012-10-24. "Antitrust: Commission sends Statement of Objections to Microsoft on non-compliance with browser choice commitments." Press release. Brussels: European Commission.
- EC 2013/6= European Commission. 2013-03-06. "Antitrust: Commission fines Microsoft for non-compliance with browser choice commitments". Press release. Brussels: European Commission.
- The Economist* 2014-02-07. "Nudge unit leaves kludge unit". Newspaper Article. [Electronic] Available: <http://www.economist.com/blogs/freexchange/2014/02/market-paternalism>. Downloaded: 2014-08-10.
- EC MEMO/11/675= European Commission, 2011-10-10. *New EU rules on consumer rights to enter into force*. Brussels: European Commission.
- EC web HiAP= European Commission web page. "Health in All Policies". [Electronic] Available: http://ec.europa.eu/health/health_policies/policy/. Downloaded: 2014-07-24.
- EC web 2013= European Commission web page, 2013. "Web browser choice for European consumers" [Electronic] Available: http://ec.europa.eu/competition/consumers/web_browsers_choice_en.html. Downloaded: 2014-07-24.
- Exec. Order No. 12866, 58 Fed. Reg. 51735 (1993-10-30).
- Exec. Order No. 13563, 76 Fed. Reg. 3821 (2011-01-18).
- Feinberg, Joel 1986. *Harm to Self: The Moral Limits of the Criminal Law*. Oxford University Press.
- Guldborg Hansen, Pelle & Andreas Maaløe Jespersen, 2013. "Nudge and the Manipulation of Choice – A Framework for the Responsible Use of the Nudge

- Approach to Behaviour Change in Public Policy”. *European Journal of Risk Regulation*, No.1, pp. 3-28.
- Hallsworth, Michael, John A. List, Robert D. Metcalfe & Ivo Vlaev, 2013. ”The Behavioralist As Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance”. *NBER Working Paper Series - National Bureau Of Economic Research*, No. 20007.
- Hawton, Keith, Helen Bergen, Sue Simkin, Sue Dodd, Phil Pocock, William Bernal, David Gunnell & Navneet Kapur, 2013. “Long term effect of reduced pack sizes of paracetamol on poisoning deaths and liver transplant activity in England and Wales: interrupted time series analyses”. *British Medical Journal*, Vol. 346, no. 2, p.15.
- Haycock, Kati, 2006. ”Promise Abandoned: How Policy Choices and Institutional Practices Restrict College Opportunities.” *Report by The Education Trust*, Washington, USA.
- Hayek, Fredrich A., 1945. “The Use of Knowledge in Society”. *The American Economic Review*, Vol. 35, No. 4, pp. 519-530.
- Haynes, Laura C., Donald P. Green, Rory Gallagher, Peter John & David J. Torgerson, 2013. “Collection of Delinquent Fines: An Adaptive Randomized Trial to Assess the Effectiveness of Alternative Text Messages”. *Journal Of Policy Analysis And Management*, Vol. 32, no. 4 pp. 718 – 730.
- Haynes, Laura C., Owain Service, Ben Goldacre & David Torgerson, 2012. *Test, Learn, Adapt – Developing Public Policy with Randomised Controlled Trials*. London: UK Cabinet Office’s Behavioural Insights Team. [Electronic] Available: www.gov.uk/government/organisations/behavioural-insights-team. Downloaded: 2014-02-11.
- Health Inequalities Portal= European Portal for Action on Health Inequalities, Web page. [Electronic] Available: http://www.health-inequalities.eu/HEALTHY/EN/policies/health_in_all_policies/. Downloaded: 2014-08-07.
- Idea42 2012. “Solving Social Problems with a Nudge”, video speech by Sendhil Mullainathan. [Electronic] Available: <http://www.ideas42.org/the-last-mile-problem-sendhil-mullainathans-ted-talk/>. Downloaded: 2014-07-23.
- John, Peter, Graham Smith & Gerry Stoker, 2009. “Nudge Nudge, Think Think: Two Strategies for Changing Civic Behaviour”. *The Political Quarterly*, Vol. 80, No. 3, pp. 361-370.
- Johnson Eric J. & Daniel Goldstein, 2003. ”Do Defaults Save Lives?”. *Science*, Vol. 302, No. 5649, pp. 1338-1339.
- Kahneman, Daniel, 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kondo, Chisuzu, Chiaki Saito, Ayaka Deguchi, Miki Hirayama & Adam Acar, 2010. “Social Conformity and Response Bias Revisited: The Influence of ‘Others’ on Japanese Respondents. *Human Affairs*”, No. 20, pp. 356-363.
- Mill, John Stuart, 2001 [1859]. *On Liberty*. Ontario: Batoche Books Limited.
- Mimica, Neven, 2013. *Applying Behavioural Insights to Policy-Making: Results, Promises and Limitations*. SPEECH/13/761 (2013-09-30) by the Commissioner for Consumer Policy. Brussels: European Commission.

- Mullainathan, Sendhil & Eldar Shafir, 2013. *Scarcity – Why Having Too Little Means So Much*. New York: Time Books.
- NBC 2014-06-26 “Top New York Court Wont Reinstate City’s Big Sugary Drink Ban”. Newspaper Article. [Electronic] Available: <http://www.nbcnewyork.com/news/local/>. Downloaded: 2014-08-07.
- Nesta web 2014= Nesta webpage, 2014-02-05. “Behavioural Insights Team: our new partnership”. [Electronic] Available: <http://www.nesta.org.uk/blog/behavioural-insights-team-our-new-partnership>. Downloaded: 2014-08-11.
- New York Times* 2009-09-10. “Senate Confirms Regulatory Chief”. [Electronic] Available: <http://thecaucus.blogs.nytimes.com/2009/09/10/senate-confirms-regulatory-chief/>. Downloaded: 2014-08-11.
- NYRB 2008= *New York Review of Books* 2008-06-12. ”Economics: Which Way for Obama?”. Newspaper Article. [Electronic] Available: <http://www.nybooks.com/articles/archives/2008/jun/12/economics-which-way-for-obama/>. Downloaded: 2014-08-11.
- OIRA 2010/6= *Disclosure and Simplification as Regulatory Tools*. Office of Information and Regulatory Affairs’ memorandum for the heads of executive departments and agencies. 2010-06-18. Washington: Executive office of the President.
- OIRA 2011/2= *Executive Order 13563, “Improving Regulation and Regulatory Review”*. Office of Information and Regulatory Affairs’ memorandum for the heads of executive departments and agencies, and of independent regulatory agencies. 2011-02-02. Washington: Executive office of the President.
- OIRA 2011/09= *Informing Consumers through Smmi Disclosure*. Office of Information and Regulatory Affairs’ memorandum for the heads of executive departments and agencies. 2010-09-08. Washington: Executive office of the President.
- Ostrom, Elinor, 2007. ”Institutional Rational Choice: An Assessment of the Institutional Analysis and Development Framework” in Sabatier, Paul, A., (ed.), 2007. *Theories of the Policy Process*. Boulder: Westview Press. pp. 21-64.
- Pronin, Emily, Daniel Y. Lin & Lee Ross 2002. “The Bias Blind Spot: Perceptions of Bias in Self Versus Others”. *Personality and Social Psychology Bulletin*, Vol. 28, No. 3, pp. 369-381.
- Sabatier, Paul A., 2007. *Theories of the Policy Process*. Boulder (Colorado): Westview Press. (2nd ed.).
- Samuelson, William & Richard Zeckhauser, 1988. “Status Quo Bias in Decision Making”. *Journal of Risk and Uncertainty*, Vol. 1, pp 7-59.
- Seawright, Jason & John Gerring, 2008. “Case Selection Techniques in Case Study Research – A Menu of Qualitative and Quantitative Options”. *Political Research Quarterly*, Vol.61, No. 2, 294-308.
- Stanovich, Keith E., 2011. *Rationality & the Reflective Mind*. New York: Oxford University Press.

- Stanovich, Keith E. & Richard F. West, 2000. "Individual differences in reasoning: Implications for the rationality debate?". *Behavioral and Brain Sciences*, Vol. 23, pp. 645-726.
- Sunstein, Cass R., 2013. *Simpler: The Future of Government*. New York: Simon & Shuster.
- SVT 2014= Sveriges Television 2014-07-22. "Ingen bötfälls för skippat bälte på Buss". Newspaper article. [Electronic] Available: <http://www.svt.se/nyheter/regionalt/vastnytt/>. Downloaded: 2014-07-23.
- Thaler, Richard H. & Cass R. Sunstein, 2003. "Libertarian Paternalism". *The American Economic Review*, Vol. 93. No. 2, pp. 175-179.
- Thaler, Richard H. & Cass R. Sunstein, 2009 [2008]. *Nudge – Improving Decisions About Health, Wealth, and Happiness*. London: Penguin Books.
- Tversky, Amos & Daniel Kahneman, 1974. "Judgment under Uncertainty: Heuristics and Biases". *Science*, Vol. 185, No. 4157. pp. 1124-1131.
- Tversky, Amos & Daniel Kahneman, 1981. "The Framing of Decisions and the Psychology of Choice". *Science*, Vol. 211, No. 4481. pp. 453-458.
- Tversky, Amos & Daniel Kahneman, 1991. "Loss Aversion in Riskless Choice: A Reference-dependent Model". *The Quarterly Journal of Economics*, Vol. 106, No. 4, pp. 1039-1061.
- Whitman, Douglas Glen & Mario J. Rizzo, 2007. "Paternalist Slopes". *NYU Journal of Law & Liberty*, Vol. 2, pp. 411-443.
- Yin, Robert K., 2003. *Case Study Research – Design and Methods*. (3rd ed.). London: Sage Publications.

10 Appendix 1

