Should Evolutionary Economists Embrace Libertarian Paternalism?

Martin Binder [a,b]

[a] Science and Technology Policy Research Unit, University of Sussex, Falmer, Brighton, BN1 9SL, UK

[b] Evolutionary Economics Group, Max Planck Institute of Economics, Jena, Germany

+44 (0) 1273 873582

M.Binder@sussex.ac.uk

Abstract

Libertarian paternalists hold that biases and distortions in human decision-making justify paternalistic interference into individuals' decisions. The aim of this paper is to analzye to what extent an evolutionary outlook supports libertarian paternalism. I will put forward three arguments in favour of libertarian paternalism and six objections that strongly oppose it. While evolutionary economists should take seriously the contention that our positive knowledge of real-world decision-making will have to influence our normative assessment of these decisions, the objections against libertarian paternalism brought forward in this paper serve as a cautionary note. Contrary to the claims of its proponents, libertarian paternalism is neither inevitable, nor does it provide an adequate measuring rod of normative rationality. It is prone to abuse by anchoring its standard of rationality pragmatically to norms and can thus promote conservative bias and stifle innovative exploration. It also presents the policy-maker with a compounded Hayekian knowledge problem. Finally, from a dynamic point of view, libertarian paternalism's manipulative shaping of preferences might lock-in individuals into heteronomous preference learning paths without individuals being even aware of it.

1. Introduction

People are stupid. In a nutshell, this statement seems to sum up not only the fundamental premise of paternalism but the insights of behavioural economics resulting from decades of work into the bounds of human rationality. Of course, the statement is an exaggeration. People are well able to act rationally in a wide variety of situations. But indeed, behavioural economics has shown us that there are limits to human rationality because our cognitive capacities are limited (Conlisk, 1996, Rabin, 1998, Gigerenzer et al., 1999, Kahneman, 2003) and the environment in which we live is increasingly complex. To come to decisions, we often rely on heuristics (Gigerenzer et al., 1999, Tversky and Kahneman, 1974), on habits (Pollak, 1976) or take other shortcuts that save on mental efforts (Simon, 1956). Often we have to decide in environments characterized by strong uncertainty (Gigerenzer, 2008, Witt 2009) or in fast-changing contexts resulting from the continuing innovativeness of society. This all puts constraints on rational decision-making and is bound to lead to suboptimal individual decisions. Thanks to the insights from behavioural economics, we are starting to understand the varieties of biases and distortions our decisions and actions underlie and the consequences this has for different economic situations, such as weakness of will and overconsumption (obesity being a case in point, Graham and Felton, 2005, Anand and Gray, 2009), present bias and under-saving (Thaler and Benartzi, 2004), status-seeking and zero-sum status races (Frank, 1999, 2012), falsely forecasting our future happiness (Gilbert, 2007), discrepancies in "wanting" and "liking" (Camerer, 2006, Witt and Binder, 2011) as well as the influence of institutional arrangements and the environment in shaping our preferences (on the problems of preference endogeneity and consumer welfare see extensively Binder, 2010, Binder and Witt 2011). In short, human preferences often are neither stable, nor consistent, nor independent of context (or for that matter: they are "incoherent", Sugden, 2008).

Maybe then people are best not characterized as stupid –at least not necessarily so–, but as "cognitive misers" (Fiske and Taylor, 1984), with limited time, information and cognitive capacity that force them to regularly take shortcuts in decision-making. While decision shortcuts may work reasonably well in many contexts, they have been shown to fall short of economic rationality in various contexts. It is no wonder then, that the findings of behavioural economics have prompted researchers to suggest that our cognitive limitations serve as a justification for intervention for our own good (e.g., Sunstein and Thaler, 2003, Camerer et al., 2003, Trout, 2005). If we, on the one hand, do not know what's in our best interests while behavioural economists, on the other hand, do know what's good for us, there seems to be scope for paternalistic intervention by debiasing our decisions, reframing decisions problems and so on. Benevolent paternalists, or in the less conspicuous formulation: "choice architects" (Thaler and Sunstein, 2008), then present us with reframed decision problems, default rules and other instruments that are bound to exploit our cognitive limitations and make us better off. Opposed to the hard paternalism of yore, these new soft (or libertarian) paternalists, however, aim to "nudge" (Thaler and Sunstein, 2008) us into decisions that are best for us, but leave open the door for rational individuals to reverse these nudges if they want to. Soft paternalists of this persuasion thus argue that their paternalistic

interventions do not limit the freedom of choice of those who do rationally reflect on their decisions (Sunstein and Thaler, 2003, p. 1170).

Libertarian paternalism has sparked a lively debate in recent years and has found a number of defenders (e.g., Camerer, 2006, Trout, 2005, Loewenstein and Haisley, 2008, Amir and Lobel, 2008, Desai, 2011, Sunstein, 2012) as well as receiving some pronounced critiques (e.g., Sugden, 2008, Rizzo and Whitman, 2009a,b). No matter where one stands on the issue, one has to acknowledge that libertarian paternalism is the most visible and fruitful *normative* outcome of behavioural economics so far as it prompts economists to take into account the positive findings of behavioural economics not only in their models but also with regard to their normative relevance for consumer welfare and autonomy, as well as their instrumental policy implications. The aim of the present paper is thus to develop these normative aspects further and contribute to the debate with a twist that has so far been rather neglected: it adds an evolutionary perspective.[1] Libertarian paternalism so far is most often analysed from a timeless perspective, with examples given often being one-shot situations that do not account for the more dynamic features that libertarian paternalism might have. In this paper I will add a temporal arrow to the analysis and ask whether an evolutionary perspective would change our assessment of the strengths and weaknesses of libertarian paternalism. I will argue that taking an evolutionary perspective does indeed have implications for the justification of libertarian paternalism as well as the toolbox of instruments one should legitimately use (or not) as a libertarian paternalist.

The paper is structured as follows. Section 2 lays the groundwork for the further discussion by shortly defining the necessary concepts. Section 3 discusses the justifications as well as problems associated with libertarian paternalism in the context of evolutionary economics and brings up a number of qualifications libertarian paternalists should be aware of and take into account. If neglecting those evolutionary insights, unwanted side effects of paternalistic intervention are bound to occur and severely decrease the applicability of libertarian paternalism in real-world contexts. Section 4 concludes.

2. Libertarian Paternalism

Economists tend to be wary of paternalism and often they brand disliked policy interventions as "paternalistic" to stifle discussion about their merits. But what is "paternalism"? Ronald Dworkin provides the canonical definition:

> "Paternalism is the interference of a state or an individual with another person, against their will, and defended or motivated by a claim that the person

---

[1] See also Cordes and Schubert (2010) as well as Schubert (2012b) who discuss how libertarian paternalism can be used to deal with dysfunctional preference learning. Lades (2012) analyses the underlying motivational mechanisms of impulsive consumption and the potential of self-nudges from an evolutionary perspective.

interfered with will be better off or protected from harm" (Dworkin, 2010, no page number given)[2]

It is thus paternalistic if the state intervenes into the affairs of an individual to improve that individual's welfare and the state does so without the consent of the individual and by curtailing the liberty or autonomy of said individual. At the heart of paternalism lies the idea that the interferer can improve the well-being (however defined) of an individual by instating a policy measure that forces the individual to do something beneficial, which the individual would otherwise not have done, maybe due to ignorance or weakness of will or for other reasons. Examples that come to mind are seatbelt or helmet laws, or pension schemes and health insurance.[3] In each case, the individual is forced to something it would otherwise not do (wear an inconvenient helmet, pay for pensions or health insurance) in order to improve its well-being (by avoiding poverty in old age, serious injuries in the case of an accident or being able to afford treatment in the case of sickness). Typically, thus, paternalism involves a trade-off between welfare and autonomy or liberty (e.g., Hausman and McPherson, 1994, p. 399).[4]

Economists of liberal persuasion oppose paternalism and subscribe to the value judgement of normative individualism (Buchanan, 1991, ch. 18, Hausman and McPherson, 2008, pp. 237-238), the principle that "assigns a higher moral value to the individual than to the community or society, and which consequently advocates leaving individuals free to act as they think most conducive to their self-interest" (Macpherson, 1991, p. 790). Normative (or ethical) individualism then less abstractly is defended by economists in the form of the principle of consumer sovereignty, i.e. the principle that individual preferences are given and not to be judged by some external force.

Anti-paternalism (or consumer sovereignty) features prominently in economics textbooks today and is also often praised as being ethically neutral (Gintis, 1974, p. 415), not imposing external ethical judgements on individuals. For the anti-paternalist, the only cases in which policy intervention is permissible are those of market failure, when

---

[2] A slightly less compact but more precise definition is: "X acts paternalistically towards Y by doing (omitting) Z iff (1) Z (or its omission) interferes with the liberty or autonomy of Y. (2) X does so without the consent of Y (3) X does so just because Z will improve the welfare of Y (where this includes preventing his welfare from diminishing), or in some way promote the interests, values, or good of Y." (Dworkin, 2010, no page number given). All three conditions could be subjected to further philosophical scrutiny (Dworkin, 2010) and are not as clear cut as they might prima facie seem: would hiding the sleeping pills of your suicidal partner limit the liberty or autonomy of your partner (1)? Or what about consent (2) if the "beneficiary" of the paternalist intervention is not even aware of the intervention? Or what about policies that have other goals in addition to the welfare of the targeted individuals (3)?

[3] Note that examples of policies that are purely paternalistic are difficult to come by (Dworkin, 1972, p. 65). Even in the above-mentioned examples, besides "protecting" the individual from itself, an additional motivation might lie in reducing externalities, e.g. when helmet or seatbelt laws reduce the cost of injuries to the social health care system.

[4] Note that one could also conceive of paternalism involving a trade-off in present liberty for future liberty, when one's liberties today are curtailed in order to have larger liberties in the future as a result (a case in point would be a definition of welfare that includes liberties). This case of trade-off might actually be easier to defend by the paternalist (see Dworkin, 1972).

markets malfunction and fail to allow consumers to satisfy their preferences (a staunch defender of anti-paternalism, even in the case of incoherent preferences is Robert Sugden, see e.g., Sugden 2004, 2008, 2011).[5] One can defend normative individualism by noting that "the individual chooses that which he chooses" and whether this choice is optimal is irrelevant as "[a]ll there is are individual choices…" (Buchanan, 1991, pp. 225-226). More often, however, normative individualism is defended on the grounds of "epistemic privilege" (ibid.), viz. the contention that the individual knows best about her own interests and no benevolent planner can know better. This lies at the heart of John Stuart Mill's classical (utilitarian) defence of anti-paternalism:

> "But neither one person, nor any number of persons, is warranted in saying to another human creature of ripe years, that he shall not do with his life for his own benefit what he chooses to do with it. He is the person most interested in his own well-being, the interest which any other person, except in cases of strong personal attachment, can have in it, is trifling, compared to that which he himself has; … with respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by any one else." (Mill, 1869[1991], p. 91)

Resulting from this insight, Mill then judges that the harm of interfering with an individual is far greater than having the individual commit errors.

> "All errors which he is likely to commit against advice and warning, are far outweighed by the evil of allowing others to constrain him to what they deem his good." (ibid., p. 92)

By the utilitarian principle, he then rejects paternalism and arrives at his "harm principle", i.e. the principle that only harm to others justifies the interference in a person's life and actions.[6] A similar anti-paternalist stance is taken by Jeremy Bentham, evidenced in his assertion that push-pin and poetry should be considered of equal value given they result in equal pleasure (Bentham, 1830, p. 206, Sugden, 2006, p. 34).

It becomes apparent now that with the advent of behavioural economics and a legion of findings on individual behaviour that falls short of full economic rationality, the idea that the individual actually knows best about his interests, has lost some credibility.[7] While

---

[5] Paternalists and anti-paternalists usually also tend to agree that paternalism is warranted in the case of minors or people that are mentally not competent to judge for themselves.

[6] "That principle is, that the sole end for which mankind are warranted, individually or collectively in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise, or even right." (Mill 1869[1991], p. 30.)

[7] Note that a defence of anti-paternalism along utilitarian lines seemed weak even before the advent of behavioural economics (see more extensively Dworkin, 1972, pp. 74-76): Mill also rejects paternalism in his treatise based on other more Kantian grounds, arguing that choice is

paternalism traditionally was meant to further the well-being of those who could not do so on their own (minors or mentally incompetent individuals are the traditional target group for paternalism), the line of argument of behavioural economists now increases the scope for paternalism and increases the category of "idiots" (Camerer et al., 2003, p. 1218) to include the typical human individual (to the extent it makes decision errors).[8] Paternalism becomes less focussed on groups that must be protected, but becomes more situational (ibid.): all of us might benefit from paternalistic intervention in certain situations.

Behavioural economists thus attack anti-paternalism via denying the plausibility of "epistemic privilege" and argue that experts on decision heuristics and biases actually know better than the individual, thus justifying paternalistic intervention. As such, proponents of libertarian paternalism subscribe to a similar definition of paternalism as the one offered above: "… a policy counts as paternalistic if it is selected with the goal of influencing the choices of affected parties in a way that will make those parties better off." (Thaler and Sunstein, 2003, p. 175). However, the idea of libertarian paternalism[9] is to improve people's well-being without limiting their autonomy:

> "...our emphasis here is not on blocking choices, but on strategies that move people in welfare-promoting directions while also allowing freedom of choice. Evidence of bounded rationality and problems of self-control is sufficient to suggest that such strategies are worth exploring." (Sunstein and Thaler, 2003, p. 1170)

Intervention should thus be designed so as to leave individuals' freedom of choice intact but "nudge" (Thaler and Sunstein, 2008) individuals towards a choice that will be considered preferable in terms of the nudgees' assessment. The libertarian paternalist tries to assess the costs and benefits of the policy in light of the *beneficiaries' best interest*, which should be "measured as objectively as possible" (Thaler and Sunstein, 2003, p. 175) and aims at a policy that reflects what individuals would have chosen for themselves if they were fully rational:

> "… we emphasize the possibility that in some cases individuals make inferior choices, choices that they would change if they had complete information,

---

valuable in itself and respect for a person (treating a person as end, not as a means) should prompt anti-paternalism (on Mill's view on paternalism, see also Arneson, 1989).

[8] By referring to "idiots", Camerer et al. 2003 quote an early legal case where the justification for paternalism was spelled out with respect to those who cannot make competent decisions for themselves, namely "idiots, minors or married women" (Rogers v. Higgins, 1868, quoted from Camerer et al., 2003, p. 1213 fn 5).

[9] Libertarian paternalism is sometimes also called: "soft" or "light paternalism" (see Loewenstein and Haisley, 2008, for a survey). It is very similar to "asymmetric paternalism" (Camerer et al., 2003), which is defined not in terms of not limiting choice options but through asymmetrically benefitting individuals making decision errors while imposing small costs on individuals who act completely rational. For the purpose of the present paper I'll focus on libertarian paternalism while noting that the two are rather similar in many cases and most of the discussion will also apply to asymmetric paternalism.

unlimited cognitive abilities, and no lack of willpower." (Thaler and Sunstein, 2003, p. 175)

An example of such a libertarian paternalist policy would be the setting of default rules, for example for pension schemes (Choi et al. 2004, Madrian and Shea, 2001): participation in these schemes tends to be low despite the fact that most individuals seem to want to save for their retirement (as e.g. evidenced by surveys) and despite the fact that employer matching mechanisms add "free" money for the individuals to the plan. Changing the default from opt-in to opt-out (individuals are enrolled in pension schemes automatically (by default) but are allowed to opt out of the scheme if they wish to do so) was shown to dramatically increase participation (Choi et al., 2004). Status-quo-bias was conjectured to be the root cause preventing individuals from choosing the preferred and welfare-increasing option. Similarly, automatically tying pension savings increases to salary increases can prevent under-saving by using this status-quo-bias (once the automatism is in place, individuals tend to not opt-out and loss aversion is circumvented since individuals never consider the portion of their salary increase as foregone earnings; see Thaler and Benartzi, 2004).

Ill-informed, weak-willed individuals are thus "nudged" towards a welfare-improving behaviour (saving for their pensions) while any rational individual could opt out if with no or just small costs. Other instruments in the toolkit of the libertarian paternalist are "required active choosing" (mandated choice), which can accompany default rules. Required active choosing works well with yes/no questions (e.g. as regards organ donations) but it might work much less well when individuals face many choices (Sunstein and Thaler, 2003, p. 1173), such as between thousands of mutual funds to put their 401(k) contributions into (Iyengar et al., 2003). Additional instruments are "cooling off periods" (individuals can impulsively purchase things at will but can later on reverse their decisions for a limited period of time) or the "debiasing" and "reframing" of consumer choices (Trout, 2005). While debiasing would aim to present relevant information in the format that is most conducive to a well-informed rational decision and aims to help people avoid making errors, reframing can go a step further in exploiting framing so as to manipulate people in choosing that which is in their (alleged) best interest. Similarly, it has been suggested to use default options as "anchors": if individuals have no well-formed preference, giving out defaults or specifying ranges in mandated choice exercises can give individuals useful hints about where their preference should lie from the point of view of the paternalist (Sunstein and Thaler, 2003, p. 1177). Finally, Camerer (2006) suggests issuing learner's permits in a wide range of fields where individual decisions are error-prone, such as borrowing and credit card use.[10]

In sum, behavioural economists justify paternalistic intervention with bias and irrational decisions of individuals (we could call this the "weak form" of the behavioural

---

[10] These instruments can be combined with other procedural or even substantive constraints, thus posing more severe restrictions of individuals' liberties. For example, opt-out can be limited to specified waiting periods or other substantive constraints, e.g. conditions under which opt-out will be possible or not (Thaler and Sunstein, 2003, pp. 1188-1190).

economists' argument). Defenders of libertarian paternalism can go one step further and argue that many policy measures by the state rely on framing no matter what and thus paternalism is "inevitable" (Thaler and Sunstein, 2003, p. 177): if individual choices depend on the way they are being framed, then there is no neutral way of presenting individuals with choices and the state cannot avoid being paternalistic. In this "strong form" of the argument, defenders of libertarian paternalism claim that the inevitability of paternalism makes the anti-paternalist stance of standard economics "incoherent" (ibid.). Whether these claims are plausible will be the focus of the following section.

3. Libertarian paternalism from an evolutionary perspective

Libertarian paternalism has sparked a lively debate in recent years having offered scholars the opportunity to explore the normative implications of behavioural economics. Typical lines of defence centre on the idea that large welfare gains can be realized by helping less rationally calculating individuals to overcome their bias. Compared to these welfare gains, reductions in autonomy or freedom of choice are allegedly small (e.g., Trout, 2005, Thaler and Sunstein, 2008, Camerer et al., 2003). It is usually argued that paternalism is justified in situations that are complex and where people lack experience in decision-making. Experts here are assumed to be more knowledgeable and the potential for improving people's choice is stressed (Loewenstein and Haisley, 2008, p. 8). Moreover, libertarian paternalists often defend intervention with the assumed consent, i.e. people allegedly want to be nudged.[11] Objections against libertarian paternalism come from a variety of directions, most viciously from libertarians who deny that libertarian paternalism is libertarian or in any way liberty-preserving (Mitchell, 2005, White, 2008). It has also been disputed whether libertarian paternalism is liberal and respects individual subjectivity and heterogeneity (Gruene-Yanoff, 2012). In this tradition fall also arguments that question the inevitability of paternalism, even when preferences are incoherent (Sugden, 2008, 2011). Similarly, critiques from a liberal tradition have focussed on rejecting the idea that autonomy and liberty should be equated with freedom of choice (Qizilbash, 2009). Other definitional objections have been made with regard to the informational campaign tool from the libertarian paternalist's toolbox (Hausman and Welch, 2010): informing individuals about their biases and framing information in a way that does not distort decision-making do not count as paternalism under the canonical definition of paternalism (see above). More traditional critiques in a public choice theoretic spirit doubt the will and ability of government to correct bias (Glaeser, 2006). It is not the purpose of the present paper to survey the whole debate (see Rebonato, 2012, for a rather comprehensive overview). Rather, in the present paper I want to analyse to what extent subscribing to an evolutionary methodology and world-view makes libertarian paternalism a plausible

---

[11] People might want to be nudged in some cases. However, according to our definition of paternalism, nudging people who consent to this is no longer paternalism (Sugden, 2011, p. 32). Moreover, people enjoy making decisions and in many cases, one could as easily argue that they want to be offered the possibility to choose and resent being manipulated (see Rebonato, 2012).

policy tool or not. Such an analysis is necessary because of the normative nature of the debate. Depending on the methodological outlook scholars of libertarian paternalism hold, different arguments to the viability of libertarian paternalism can be given different weights. For example, libertarians, on the one hand, put strong emphasis on arguments relating to the autonomy- and liberty-decreasing nature of some of the proposed interventions. Behavioural economists, on the other hand, stress the fallibility of human decision-making and the loss of welfare as trumping libertarian concerns. In a similar vein, I will argue that evolutionary economists, qua their focus on the dynamics of economic phenomena, should pay particular attention to a number of features of libertarian paternalism that are usually neglected in a static mode of analysis, viz. the preference dynamics entailed by paternalistic intervention. Before discussing these arguments in more detail, it is useful to spell out what I hold to be a plausible characterization of the key two features of evolutionary economics (Witt, 1987, 2008, Binder, 2010). The first feature of evolutionary economics is its focus on evolutionary change. Opposed to being merely dynamic, evolutionary change exhibits historicity, is thus irreversible. Opposed to dealing with closed systems, evolutionary economics' focus is on open systems, where the economy is understood as driven by change resulting from the creation and dissemination of novelty (Witt, 2003). In these continuously evolving systems, the notion of equilibrium becomes untenable (Metcalfe, 2001, Nelson and Winter, 1982, p. 356) and individual behaviour cannot be well characterized by optimization. It rather is conceptualized as the adaptation to new environments via learning. Assuming preferences to be (*a priori*) fixed and stable is no longer viable, shifting the focus to the analysis of preference change dynamics (Witt, 2001, Binder, 2010, Witt and Schubert, 2013). The second feature is closely related to the strong uncertainty characterizing endogenously changing systems: Human rationality and decision-making has evolved to deal with specific problems posed by the environment and cannot easily be conceptualized with the rational choice model of Olympic rationality. This prompts a "naturalistic methodology" (Witt, 1987) or a behavioural economics understanding of human rationality. Taking recourse to findings of biology, psychology and other behavioural sciences gives us a better-founded understanding of human rationality. This second feature of evolutionary economics *prima facie* prompts the hypothesis that libertarian paternalism would be as welcome to evolutionary scholars as it is to behavioural economists.[12]

To what extent can an evolutionary outlook support libertarian paternalism? In the following I will discuss three arguments in favour of libertarian paternalism and a number of objections against it. The arguments pro libertarian paternalism I will call the *fairness argument*, the *autonomy argument* and the *faulty preference learning argument*. The *fairness argument* in favour of libertarian paternalism rests on the empirical fact that cognitive abilities are not equally distributed and that a subset of people makes decisions that are not furthering their interest as much as well-informed decisions could do. Compared to other individuals, this subset of people thus achieves

---

[12] In my understanding evolutionary economics thus extends behavioural economics with a dynamic (evolutionary) perspective and a focus on the motivational underpinnings of human decision-making (Witt, 2012).

lower levels of welfare as a result of decisions errors and biases. This empirical fact then can give support to the value judgement that fairness (or justice) would obligate the paternalist to help the less able in reaching higher levels of welfare (see, e.g., Arneson, 1989, p. 412). Essentially, this argument also draws on the conviction that welfare is more important than autonomy and paternalism is justified. Ultimately, however, the justifying value judgement is fairness, not welfare. In a strong form, this argument can even dictate that paternalism that imposes costs (how minimal they might be in the case of libertarian paternalism) on the better abled individuals (who might have to reverse default rule decisions) is justified for fairness reasons if the less well-able profit from it. It should be noted that these redistributive considerations cannot be considered libertarian under any usual definition of libertarianism (see Valentyne, 2012, Mitchell, 2005).

The *autonomy argument* takes the anti-paternalist objection that paternalism decreases individuals' autonomy and liberty and puts it on its head. The very fact that individuals' choices are influenced by the environment and institutional setups is taken as evidence that not interfering with individuals' choices might threaten their autonomy (Sunstein and Thaler, 2008, p. 1170): the true danger thus lies in biased decisions an individual is not aware of (Trout, 2005, p. 408-414). In this line of defence of libertarian paternalism, it is argued that just increasing individuals' numbers of choice options or making them choose ("required active choosing") might be inimical to their freedom and autonomy, if people are overwhelmed by their choice options. Information and choice overload is indeed a plausible reason for people to avoid coming to decisions (Camerer et al., 2003, p. 1235; Sunstein and Thaler, 2003, p. 1196). For example requiring a person to choose from thousands of mutual funds for their pension schemes has been empirically be shown to decrease choice rates as opposed to choice situations with only few mutual funds (Iyengar et al., 2003). Moreover, providing all choice options might still not solve the problem of biased or distorted choices, which are made on this expanded choice set. In this respect, it might indeed be autonomy-promoting (autonomy here understood not in its narrow definition of increasing the number of choices/preserving all choice options) to limit choice if people actually desire choice sets that are cognitively manageable for them. Blindly arguing for maximal liberties (choices) then might be running against individuals' interests more than a modest libertarian choice architecture. From this point of view, one would question why a maximum in freedom of choice should be promoted if individuals are actually opposed to this (Arneson, 1989, p. 435). This argument gains force if one considers that cooling-off periods and required active choosing might indeed be paternalist instruments that can promote autonomy and liberties: decisions made in the spur of the moment can be reversed after critical (re-) consideration and decisions that would have not been obvious and probably neglected can be brought to the individuals' attention through mandating reflection and subsequent choice (Hausman and Welch, 2010, p. 134).[13]

---

[13] An interesting recent twist on this argument has been given by Duflo (2012) who argues that libertarian paternalism might be justified in the context of poverty alleviation in the sense that

The third *argument from faulty preference learning* basically puts the previous argument into a dynamic context. From an evolutionary point of view, many decision situations do not take place in limbo but are part of a series of decisions that can ultimately lead to preference learning dynamics that are not in the interests of the individuals who experience them. Preference learning in this case "'fail[s] to work' effectively" (Cordes and Schubert, 2010, p. 23) in the sense that it continually fails to increase the person's well-being (judging from the title of the paper: this makes individuals "unhappy" or leads to the "systematical frustration on the part of the agent" in achieving well-being through preference learning, ibid., p. 24). The example the authors give for these dysfunctional preference dynamics are social conditioning processes where people imitate role models and, through social learning, engage in status races that are by definition bound to leave everyone worse off (e.g., Frank, 1999). These status-oriented consumption paths then would frustrate individuals' abilities to effectively experiment with new preferences and should be subjected to libertarian paternalist nudges: dysfunctional preference learning is replaced with a more functional, effective form of preference learning (see also Schubert, 2012a, 2012b, p. 21).[14]

While the arguments discussed above support (libertarian) paternalism, an evolutionary perspective also prompts a number of serious reservations against libertarian paternalism. I will call these arguments the *inevitability objection*, *rationality objection*, *compounded knowledge objection*, *conservatism objection*, *slippery slopes objection* and the *manipulative preference learning objection*. The *inevitability objection* first purports to show that one central claim of libertarian paternalists, namely the inevitability of paternalism when human decisions are fraught with error, is a *non sequitur*. While there are indeed choice situations where a planner cannot avoid framing the options and influencing decision-makers, the claim that this is always the case is just plain wrong. Moreover, it has been shown that in many cases, even if individuals make non-rational decisions, the aggregate outcome of these decisions can be in the (material) interests of those involved. The idea that consistent micro-behaviour can have diverging macro-consequences is well-known for quite a while (Schelling, 1978). An example of which is when individuals fail to react in what rational choice theory predicts and contribute to public goods or to partaking in the costs of reducing externalities. Given the desirable outcome (for the individuals!), it is hard to argue that such failure in rational choice should be corrected through paternalism; such examples rather point to the inadequacy of rational choice theory as normative standard of rationality.[15] Similarly, it has been established that many heuristics or decision-shortcuts fail to comply with normative rationality demands of the libertarian paternalist but lead to actually quite adaptive and

<hr>

nudging people with respect to basic goods such as health and nurture are the prerequisites for these individuals to be able to enjoy other more valuable positive freedoms.

[14] One has to add that Cordes and Schubert (2010) are clear in that they would also see these dynamics as legitimizing hard paternalism (p. 25).

[15] A different example would be entrepreneurship: excessive risk-taking or overconfidence might actually be a good thing for society in the case of innovativeness. While overconfidence of entrepreneurs might be detrimental for the individual entrepreneur who fails with a business, overall it does benefit society because, once in a while, a successful entrepreneur markets an innovation (Rizzo and Whitman, 2009b, p. 704).

effective well-being outcomes (Gigerenzer et al., 1999). It is thus far from clear whether every bias and distortion from rational choice decision-making runs counter to individuals' interests. Whether heuristics work or not often depends on the decision environment and for a given policy-goal of the paternalist, failure to optimize (and, say, "satisfice" instead) can either imply paternalism or laissez-faire/anti-paternalism (Berg and Gigerenzer, 2007). This context-dependent applicability of corrective paternalism is a farcry away from the alleged "inevitability" of libertarian paternalism.[16] We here see that the debate on libertarian paternalism has been based on a distorted picture of human rationality, where every deviation from the standard of rational choice theory is interpreted as a deficiency, having clearly negative connotations. It can be questioned whether this interpretation is apt. Consider the human tendency to value the present stronger than the future, which is evidenced in hyperbolic discounting. Is this an error in decision-making or rather a feature of human nature? We live in the present and thus give present experiences higher weight, so what is the optimal discount rate (is there only one for everybody, or one for each of us, or many discount rates for each of us, especially once we depart from the abstraction of *quasi*-hyperbolic discounting)? A different example is status-quo-bias (or decisional inertia). Why should this be considered a bias rather than an aspect of human nature (e.g., Sugden, 2008, p. 232)?[17] Proponents of libertarian paternalism tend to assign priority always to reasoned and cold-blooded long-term decisions (the "true" interests of the individual). But isn't affect also an indelible part of human nature? Why should affective decisions thus be (completely) discounted? Here libertarian paternalists gloss over important questions that need to be addressed rather than sticking to normative rational choice theory as measuring rod for human rationality. Opposed to what libertarian paternalists do, such a measuring rod would need to be explicitly defended (and the one used by libertarian paternalists seems rather hard to defend, see the arguments below).

The next set of counter-arguments all centre on the "naturalistic methodology" of evolutionary economics and I will argue that these arguments make clear that libertarian paternalists in a very important way fail to account for a realistic picture of human rationality (this extends the above inevitability objection in a different rationality-based direction). While libertarian paternalists accurately depict human decision-making as prone to biases, defenders of libertarian paternalism ignore these findings when it comes to their *normative* view of what constitutes rationality. In essence, given our positive knowledge of how humans behave and decide, there seems to be no point in demanding that humans behave like perfectly informed, controlled and

---

[16] On the debate about effectiveness and adaptiveness of biases and heuristics see also more extensively Rebonato (2012), pp. 47-8. It is far from settled that deviations from normative rational choice standards are suboptimal.

[17] Camerer et al. (2003), p. 1254, discuss this general question shortly when it comes to individuals' irrational demand for extended warranties that are extremely expensive compared to their actual expected value. In this calculation, the "peace of mind" extended warranties provide for individuals is typically ignored. What if people actually are willing to pay a lot of money for not having to fear a breakdown of their appliances? Similarly, state lotteries pay much less than a dollar on the dollar "invested" and yet it has not to be irrational if people spend the money to purchase some thrill or a shred of hope of becoming rich (see Gruene-Yanoff, 2012, p. 643).

unboundedly cognitively able beings that conform to the expected-utility-theory axioms of optimal decision-making under uncertainty (see Qizilbash, 2012, Rebonato, 2012, Ch. 7). Libertarian paternalism's principle of making an individual's hypothetical set of "informed preferences" the measuring rod for rational behaviour suffers several shortcomings. First, the informed preferences to which the paternalist aspires to are often only vaguely defined (Camerer et al., 2003, p. 1214). They moreover tend to be so demanding and strict that failing to conform to those can hardly be called a fault of human decision-making (but rather a fault of these standards; see Sugden, 2011, pp. 24-25). And, perhaps most importantly, this standard is not operable due to its vagueness: at what point is a preference informed *enough*? How much will-power is needed and what is the exact degree of self-control for a preference to count as informed (ibid.)? All these questions tend to be not answered in informed-preference accounts and if they are, they are answered in a normative fashion. One paternalist's definition of informed preference of the individual might be distinct from another paternalist's, removing the judgement of rationality from its original ideal of being in the interested of the individual as judged by itself. It is thus no wonder that the literature on informed preferences has defended a multitude of different definitions of when a preference should count as informed and scholars of normative economics are disenchanted with informed preference views of well-being (see, e.g., Sobel, 1995, Qizilbash, 2011, 2012, Binder and Schubert, 2011, for critical discussion). Given that informed preferences seem so different from individual's actual preferences, it is not even clear and comprehensible whether the informed individual would be the same individual as the individual having its actual preferences. Finally, as normative measuring rod for human rationality, from an evolutionary point of view, postulating Herculean rationality requirements seems to be at odds with the empirical fact that individuals make decisions that regularly and systematically differ from those judgements. Taking behavioural economics seriously would require positing a much more coherent and less demanding standard of normative rationality than the one rational choice theory has developed. In this sense, libertarian paternalists take behavioural economics seriously in their description of human behaviour but ignore it normatively by adhering to the Olympic rationality principles endorsed in traditional economics.

Besides the problem this standard of rationality has in being made operational, additional problems arise: considering that not many people will actually rise to the Olympian heights of rationality posited for decisions to count as informed, it can be conjectured that legislators will have a ready excuse to interfere with individual decisions in a wider range of cases than we see at the moment. Setting an unattainable standard of rationality thus amounts to giving a blanket justification for paternalistic intervention (Qizilbash, 2011, p. 37). This in itself is problematic (see the slippery slopes objection below). But this becomes even more dubious when combined with the manipulative character of many libertarian paternalist interventions. Consider the exploitation of decisions biases through default rules: in theory, individuals are free to opt out of these arrangements if they dislike the default option. If individuals, however, are prone to decision biases and not aware of their being manipulated via a default option, this freedom to opt out is not real. In this sense, only individuals with full rationality can opt out and enjoy the "libertarian" feature of libertarian paternalism. If

you are not endowed with Olympic rationality, you will not know about the manipulation and only have nominal as opposed to real freedom to opt out (see also Rebonato, 2012, p. 132). In the end, a large number of individuals will thus fall short of the rationality ideal and will be constrained in their freedom of choice through libertarian paternalist intervention. Moreover, being manipulated also decreases those individual's autonomy: since they are not aware of the manipulation, critical reflection is impossible. Libertarian paternalism here contradicts liberal values by curtailing liberties and autonomy much more than libertarian paternalists admit.

In practice, libertarian paternalists have tried to evade these problems with a more "pragmatic approach" (Loewenstein and Haisley, 2008, pp. 19-21) of their standard of rationality and the currency of welfare to measure it. One way of operationalizing "informed preferences" is to distinguish between two selves an individual consists of, namely the System 1-self (the "automatic system") which is fast, uncontrolled, unconscious, intuitive, affective and the System 2-self (the "reflective system") which is slow, deductive, self-aware, cognitive (e.g., Thaler and Sunstein, 2008, Ch. 2, Sunstein 2012, pp. 9-12). As in typical multiple selves views, these selves compete with each other in decision-making. To simplify things, most often short-term preferences are distinguished from long-term preferences and neatly mapped onto the two Systems mentioned above. Informed preferences are then the long-term rationally considered preferences the individual would express (for example) on a "New Year's resolution test" (this test is actually suggested in Thaler and Sunstein, 2008, p. 73). While this way of characterizing informed preferences is indeed less fraught with conceptual difficulties, it brings other deficits to the table. Giving long-term preferences priority seems only *prima facie* a viable operationalization of informed preferences. Why, one can ask, should long-term preferences be actually more informed and more reasoned than short-term preferences (Schnellenbach, 2012, p. 270-1)? While many people seem to have long-term preferences such as "I should eat more healthily", it is not clear that these are actually based on better information than the corresponding short-term preferences about an upcoming culinary indulgence. We actually might know much more about the pleasure (and maybe pain or regret) associated with indulging now than we know about the pleasures and pains associated with leading a healthy lifestyle (many benefits of which are far removed and probably rather unclear in their extent to the individual; Rizzo and Whitman, 2009a, pp. 120-22). Also, long-term preferences such as the ones that would pass a New Year's Eve test have a much stronger expressive character (see Brennan and Lomasky, 1983, Schnellenbach, 2012) than short-term preferences and it is unclear whether an individual would actually want to turn these into action. Given the lack of information individuals might actually have about the content of their long-term preferences, it can be conjectured that these preferences are strongly influenced by norms ("one should eat healthily"), some of which might also not be in the actual interest of an individual (ibid.). The libertarian paternalist here could try and salvage the distinction by pointing out that individuals might regret their indulgence of short-term preferences, but then again, isn't regret an affective notion (of System-1) that the libertarian paternalist does not accord any weight? Isn't in that sense regret the mistake that should be discounted (and which, as any other emotion, can fade with time; see Hill, 2007, p. 446)? In this sense, prioritizing System-2 preferences (or some

other sort of long-term, maybe stable, meta-preference) will not rescue the libertarian paternalist's normative notion of rationality. While the description of multiple selves with different planning horizons and different amounts of reasoned deliberation might actually capture human behaviour, it is not clear why this should only be two selves and moreover, a credible measure of normative rationality should not altogether discount the affective decisions (System 1) in favour of purely cognitive decisions (System 2) but rather strike at aiming a balance of sorts between both of them (Rizzo and Whitman, 2009a, p. 133).[18]

Before concluding the discussion about the rationality objection, note that the "pragmatic definition" of what is in an individual's best interests as judged by the individual opens the door to *ad hoc* standards of rationality, e.g. in the case of obesity. Instead of taking into account individual preferences, health propensities (genetic makeup, age, etc.), the criterion that measures what's in the individual's best interests becomes the aggregate obesity rate and libertarian intervention then aims to target everyone with a one-size-fits-all intervention (Schnellenbach, 2012). Again, a very important and central tenet of behavioural economics, viz. the heterogeneity of individuals, is neglected in this *modus operandi* when libertarian paternalists assume a representative individual that benefits from healthier diets or saving more. The fact that the rationality of eating healthier or saving more is rather a function of individuals' preferences and life situation is then ignored (Schnellenbach, 2012, p. 270). Ignoring the subjectivity and heterogeneity of individuals' experiences also seems strongly at odds with liberal (not to speak of libertarian) values (Gruene-Yanoff, 2012).

The objection from rationality makes yet another appearance in what I call the *compounded knowledge objection* of libertarian paternalism (Rizzo and Whitman, 2009a). In this objection, it is not the normative rationality postulated for the individual that is criticized as deficient in libertarian paternalism, but rather the view of rationality that is credited to the choice architect. From the point of view of evolutionary economics, one should also have a reasonable idea of the cognitive capacities of the policy-maker. While the notion of epistemic privilege of the individual might be doubted from the point of view of behavioural economics, the present line of criticism asks why any social planner or choice architect should possess the knowledge about human preferences and biases requisite to design the appropriate choice architecture (Rizzo and Whitman, 2009b, p. 737). While acknowledging the actual fallibility of human decision-making, the libertarian paternalist still assumes that the social planner acts benevolently, informationally unconstrained, without bias and free of negotiation constraints (Sugden, 2011, pp. 6-7). Compared to the standard model of state intervention, where the social planner already suffers from the Hayekian knowledge problem (Hayek, 1945), the libertarian paternalist then does need to know a large number of additional things, compounding the knowledge problem (Rizzo and Whitman, 2009a, p. 101): assuming that the informed or true preferences of heterogeneous

---

[18] See more extensively Rebonato, 2012, on the problems of a multiple selves view in this context and to what extent the metaphor of an "internality" that the warring selves create for the individual is actually problematic.

individuals can be discerned, the choice architect further needs to know individuals' extent of bias, the types of bias individuals succumb to, the extent of interdependence of biases (something at the moment ill-researched in behavioural economics), the extent of self-debiasing an individuals undertake (we know that we sometimes are weak-willed and sometimes we try and self-commit to varying degrees), and so on. Much of this information is highly local, person-specific and tacit (Rizzo and Whitman, 2009a, pp. 159-160). Even if experts or choice architects could solve the problem to define a suitable normative measuring rod of rational behaviour (i.e., how much will do we need not to be weak-willed; how much affect may go into a decision, how much regret is permissible, how much inertia, how much present bias, etc.), it remains unclear how they will measure and operationalize the degree of deviance from this normative ideal and structure choice accordingly. Note that this is not a theoretical concern but of much practical relevance for the implementation of paternalist policies and has proven to be difficult in real-world contexts: a sin-tax implemented on alcohol for example aims at heavy drinkers and their alcohol-related health problems, however, empirically, it was shown that such a tax actually mostly affects moderate drinkers and their behaviour. Even in such a rather clear-cut case, it has proven impossible to accurately predict how the tax affects behaviour (Rizzo and Whitman, 2009a, p. 158; Mast et al., 1999).

While the previous objections were predominantly taking insights of behavioural economics seriously, the remaining three objections centre on the dynamic features of libertarian paternalism and hence stress an evolutionary perspective. The *objection from conservatism* objects that the "pragmatic approach" to libertarian paternalism tends to favour the status quo and introduce a conservative bias into paternalist policies. If the paternalist policy-maker cannot reliably uncover individuals' informed preferences and what is in their best interest, a fall-back approach is to rely on more pragmatic considerations of what benefits individuals. While it has been argued above that this *ad hoc* approach to formulating the goals of paternalist intervention ignores the heterogeneity of individuals, it also can promote conservatism by relying on social norms as guiding principle to individuals' welfare (Schnellenbach, 2012, pp. 273-275): social norms tend usually not to be embodied in and enforced through legal rules but rather through social sanction and peer pressure. While these norms might play useful roles in society, it is by no means guaranteed that they are welfare-maximizing or have come to pass with reference to being beneficial for individuals' welfare. It can be conjectured that some norms are outright detrimental to individual welfare but remain in place through habit and conservative bias (individuals who deviate from norms tend to get punished, increasing their self-stabilizing potential, ibid.). If policy-makers use these social norms to formulate paternalist intervention, this can mean that the policy is not in the interest of the individuals involved. Moreover, as norms are self-stabilizing, this introduces disincentives to innovation and experimentation in society. Taking this conservative bias of libertarian paternalism to its extreme, both social norms and paternalist legislation mutually reinforce each other over time in a "conservative circle" where norms stabilize behaviour and get additional legitimacy through policies that enforce them, thus prompting further stabilizing policies for these norms and so on. In the end, this interplay can cement existing social norms even though these are not efficient or in the interests of the majority of individuals. In this respect, libertarian

paternalism might stifle experimentation and innovation and seems at odds with value judgements of evolutionary economics that favour diversity and innovation (see also the objection regarding manipulative preference learning).

A second strong objection from a dynamic point of view is the *slippery slopes objection*. Many libertarian paternalist proposals have a strong intuitive appeal, however, they all are described as one-shot situations without regard for any inter-temporal dynamics. Slippery slope objections analyse libertarian paternalism with respect to the fact that policy interventions never occur atemporally but are embedded in a dynamic context and shape the future decision environment as much as they can shape preferences and preference learning (see below). In this dynamic context, a paternalist intervention can thus increase the probability of further paternalist interventions, thus creating a slippery slope with regard to the number of interventions, but also with regard to the type of intervention. The latter has been criticized extensively (e.g. Rizzo and Whitman, 2009b) and is indeed present in the descriptions of libertarian paternalist toolboxes, which in their clearest expositions form a continuum from harmless information provision up to outright bans and coercive measures (e.g. Camerer et al., 2003). Drawing a line of which tool would be still permissible as libertarian paternalist becomes thus a blurry exercise and while emphasis by proponents is given to informational measures and default rules, repeated application might lead policy-makers to expand their focus to the stronger variants in the toolbox. The line between soft and hard paternalism becomes blurry and libertarian paternalism quietly loses its libertarian stance. A similarly problematic expansion of scope can be conjectured to appear over time considering that the applicability criterion in libertarian paternalism has been expanded from clearly defined groups (minors, mentally-challenged, …) to situational idiocy. This is another example of blurred categories that facilitate the extension of scope of libertarian paternalism. As Rizzo and Whitman (2009b) point out, libertarian paternalism no longer provides clear-cut boundaries and distinctions but gradients: neither is the extent of bias clearly defined, nor what constitutes lack of will, inertia etc. With the abandoning of a clear-cut rationality measure and a pragmatic method of defining the interest of individuals in different cases, it is easy to imagine how a once drawn line (say, a given acceptable hyperbolic discount rate) is progressively shifted over time. Given the complex interactions of decision anomalies, framing and biases, the libertarian paternalist can also exploit justificatory chains to expand the scope of intervention: if one bias is targeted and removed this might give rise to other biases that now come to the fore in full force and legislators can point to the original aim of the intervention and argue to extend measures to cover the adjusted situation. Similar chains can accrue if interventions crowd-out individual self-regulation. Rizzo and Whitman (2009b, p. 729) rightly criticize that libertarian paternalism is presented as a widely applicable all-encompassing paternalism-favouring public policy framework that encourages these intervention in all walks of life (see the many and diverse examples in Thaler and Sunstein, 2008). At the same time Thaler and Sunstein are remarkably vague about who is supposed to implement these measures (infamous Carolyn who owns the cafeteria; companies who structure their employees' pension scheme options, the state who regulates pollution and organ donation?) and what the actual costs and benefits are (in the example of pension schemes, the vague references to the "fact" the Americans save

not enough seems to sparse a justification). This potential for widening the scope of paternalist intervention is exacerbated by the fact that many of these interventions are conducted in a decidedly manipulative fashion: if individuals are not aware that they are being nudged, it becomes increasingly difficult to oppose these paternalistic measures. One positive thing to say in favour of hard paternalism is that individuals are at least aware that they are being curtailed in their liberty or autonomy and can try and oppose this. The hidden character of some of the tools of the libertarian paternalist make this extremely difficult and by this disable another safeguard against this particular slippery slope (Rebonato, 2012, p. 132).[19] These slippery slope dynamics work in tandem with social norms (as argued above) and lead to mutually reinforcing.

As such, the slippery slope objection is closely related to the *manipulative preference learning objection*. Here again, we analyse libertarian paternalism through the lens of an evolutionary world-view, where human preferences are not given and stable but shaped by economic processes as well as the libertarian paternalist's policies. There are two sides to the preference learning objection, the first of which has been discussed above as the argument from faulty preference learning. The other side, however, deals with the fact that intervention into individuals' preference learning can have serious dynamic side-effects that put into doubt the desirability of trying to engineer preference learning paths. Libertarian paternalistic intervention, from this dynamic point of view, runs the risk of nudging individuals into preference learning paths that were originally not their own preferences and that get reinforced through repeated interaction. Intervention might not only create a preference in the individual, but lead to sustain and/or increase it over time. Combined with the problematic ethos of libertarian paternalists of changing choice environments without so much as making individuals aware of this fact, this leads to associative preference learning that happens unconsciously to the individual. The behavioural sciences have extensively researched associative learning (Hergenhahn and Olson, 1997, see also Witt 2001, Binder and Schubert, 2011, pp. 12-17), where individuals are positively reinforced for choices, and which can be characterized of involving not much conscious effort. The resulting preferences tend to be highly stable and difficult to unlearn, in parts because of the low conscious involvement but also because of repeated reinforcement over long time horizons. Many of our childhood preferences are acquired via associative learning and tend to be not easily reversible (Zajonc and Markus, 1982). If these preference learning mechanisms are (inadvertently) exploited by libertarian paternalists, people can get locked-in into preference learning paths they might not even recognize.[20] Given the reservations discussed above,

---

[19] If policy-makers are boundedly rational, the problem also gets worse: hyperbolic discounting then leads to "policy temptations" of shortsighted policy-makers (Rizzo and Whitman, 2009b, p. 724). Learning of policy-makers can also be "pathological" (Freytag and Renaud, 2007).

[20] In the least, libertarian paternalism needs to be developed into a more dynamically minded theory that takes into account these features of human learning. For example, if one wanted to propagate such intervention, a distinction should be made whether decision-making errors can be overcome through experiential learning or not. If individuals can learn from their mistakes, the libertarian paternalist should aim at providing information in helping people to overcome the mistake. If a decision bias cannot overcome even in repeated situations and through learning (or if a decision is indeed a one-shot problem without the possibility to learn from the experience),

libertarian paternalism can create preference learning trajectories where individuals are nudged into preferences one can doubt are in their actual interests, but rather reflect the norms society adheres to at a given moment (or they reflect goals of a boundedly rational policy maker). Moreover, they are impeded in their autonomy given that they are manipulated into these preference paths and not aware of this fact. And finally, due to the nature of learning processes, even if individuals were to become aware of this manipulation, the stickiness of preference learning can actually negate the possibility of opting-out and trap individuals with their acquired preferences (whether preferences can be unlearned depends *inter alia* on the reinforcement schedule, see more extensively Binder 2010, Sec. 6.4.3; even if preferences can ultimately be unlearned, the costs associated with this are much higher than libertarian paternalists claim they would be in the case of opt-outs).

Even if individuals are not locked-in permanently into preference learning trajectories, the manipulative character of nudges might hinder learning in repeated interactions. (Additionally, setting of defaults might not be optimal as opposed to letting people learn from others through social learning, as Carlin et al. 2009 show. This depends *inter alia* on the cost of information gathering through social learning which is probably rather case-specific.) If individuals learn to rely on the government (or other choice architects) making decisions for them, libertarian paternalism might oppose the ideal of autonomous citizens and lead to a society of heteronomous individuals waiting for others to decide for them. Individuals would then, in general, learn to be dependent and inactive, and libertarian paternalism promotes a society where policy-makers trick and manipulate individuals. The manipulation of consumers by firms is then replaced by a manipulation through the benevolent social planner instead of helping individuals to overcome their decision biases through learning and critical reflection (compare Rebonato, 2012, p. 195).[21]

4. Conclusion

To what extent can an evolutionary outlook support libertarian paternalism? I have put forward three arguments in favour of libertarian paternalism and six objections that I submit will considerably dampen the enthusiasm for it. While evolutionary economists should take seriously the contention that our positive knowledge of real-world decision-making will have to influence our normative assessment of these decisions (see Nelson and Winter, 1982, p. 356), the objections against libertarian paternalism brought forward in this paper will serve to caution evolutionary economists to wholeheartedly embrace libertarian paternalism. Contrary to the claims of its proponents, libertarian paternalism is neither inevitable, nor does it provide an adequate measuring rod of

---

then the setting of a default might have more legitimacy. These kinds of considerations are, as of yet, largely absent in the libertarian paternalist debate.

[21] This is not to say that we should ignore human fallibility in decision-making, but instead of exploiting it in government decision-making, we should aim at increasing our faculty for critical thinking and deliberation. The libertarian paternalist should not exploit decision-making deficiencies but help in overcoming them (Rebonato, 2012, pp. 218-220).

normative rationality. It is prone to abuse by anchoring its standard of rationality pragmatically to norms and can thus promote conservative bias. It presents the policy-maker with a compounded Hayekian knowledge problem and from a dynamic point of view, its manipulative shaping of preferences might lock-in individuals into heteronomous preference learning paths, of which individuals not even know about.

In light of these features, in the end, I hope to have convinced the reader that from the point of evolutionary economics, libertarian paternalism is a rather dangerous policy tool.

References

Amir, O. and Lobel, O. (2008). Stumble, predict, nudge: How behavioral economics informs law and policy. Columbia Law Review, 108(8):2098–2137.

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

Arneson, R. J. (1989). Paternalism, utility and fairness. Revue Internationale de Philosophie, 43(170):409–437.

Bentham, J. (1830). The Rationale of Reward. Robert Heward, London.

Berg, N. and Gigerenzer, G. (2007). Psychology implies paternalism? Bounded rationality may reduce the rationale to regulate risk-taking. Social Choice and Welfare, 28(2):337–359.

Binder, M. (2010). Elements of an Evolutionary Theory of Welfare. Routledge, London.

Binder, M. and Schubert, C. (2011). Reconciling normative and behavioral economics: applying the 'naturalistic approach' to the adaptation problem. Mimeo.

Binder, M. and Witt, U. (2011). As innovations drive economic growth, do they also raise well-being? Papers on Economics & Evolution #1105.

Brennan, G. and Lomasky, L. (1983). Institutional aspects of "merit goods" analysis. FinanzArchiv, 41(2):183–206.

Buchanan, J. M. (1991). The foundations for normative individualism. In: The Economics and the Ethics of Constitutional Order, chapter 18, pages 221–229. The University of Michigan Press, Ann Arbor.

Camerer, C. F. (2006). Wanting, liking, and learning: Neuroscience and paternalism. University of Chicago Law Revue, 73:87–110.

Camerer, C., Issacharoff, S., Loewenstein, G. F., O'Donoghue, T., and Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". University of Pennsylvania Law Review, 151:1211--1254.

Carlin, B. I., Gervais, S., and Manso, G. (2009). When does libertarian paternalism work? NBER Working Paper No. 15139.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2004). For better or for worse: Default effects and 401(k) savings behavior. In Wise, D. A., editor, Perspective on the Economics of Aging, chapter 2, pages 81–125. University of Chicago Press.

Conlisk, J. (1996). Why bounded rationality? Journal of Economic Literature, 34(2):669–700.

Cordes, C. and Schubert, C. (2010). Role models that make you unhappy: Light paternalism, social learning and welfare. Papers on Economics & Evolution #1022.

Desai, A. C. (2011). Libertarian paternalism, externalities, and the "spirit of liberty": How Thaler and Sunstein are nudging us toward an "overlapping consensus". Law & Social Inquiry, 36(1):263–295.

Duflo, E. (2012). Human values and the design of the fight against poverty. Tanner Lecture 2012.

Dworkin, G. (1972). Paternalism. The Monist, 56:64–84.

Dworkin, G. (2010). Paternalism. Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/entries/paternalism. (accessed: 11.11.2012)

Fiske, S. T. and Taylor, S. E. (1984). Social cognition. Addison-Wesley, Reading/MA.

Frank, R. H. (1999). Luxury Fever: Why Money Fails to Satisfy in an Era of Excess. The Free Press, New York.

Frank, R. H. (2012). The Darwin economy - liberty, competition, and the common good. Princeton University Press, Princeton/New Jersey.

Freytag, A. and Renaud, S. (2007). From short-term to long-term orientation - political economy of the policy reform process. Journal of Evolutionary Economics, 17:433–449.

Gigerenzer, G. (2008). Rationality for Mortals - How People Cope with Uncertainty. Oxford University Press, New York.

Gigerenzer, G., Todd, P. M., and the ABC Research Group (1999). Simple Heuristics That Make Us Smart. Oxford University Press, New York/Oxford.

Gilbert, D. T. (2007). Stumbling on happiness. Harper Perennial, London.

Gintis, H. (1974). Welfare criteria with endogenous preferences: The economics of education. International Economic Review, 15(2):415–430.

Glaeser, E. L. (2005). Paternalism and psychology. NBER Working Paper No. 11789.

Graham, C. and Felton, A. (2005). Variance in obesity across cohorts and countries: A norms-based explanation using happiness surveys. Mimeo.

Gruene-Yanoff, T. (2012). Old wine in new casks: Libertarian paternalism still violates liberal principles. Social Choice and Welfare, 38(4):635–645.

Hausman, D. M. and McPherson, M. S. (1994). Preference, belief, and welfare. The American Economic Review, 84(2):396–400.

Hausman, D. M. and McPherson, M. S. (2008). The philosophical foundations of mainstream normative economics. In Hausman, D. M., editor, The Philosophy of Economics - An Anthology, chapter 13, pages 226–251. Cambridge University Press, Cambridge, 3rd edition.

Hausman, D. M. and Welch, B. (2010). Debate: To nudge or not to nudge. Journal of Political Philosophy, 18(1):123–136.

Hayek, F. A. (1945). The use of knowledge in society. The American Economic Review, 35(4):519–530.

Hergenhahn, B. R. and Olson, M. H. (1997). An Introduction to Theories of Learning. Prentice Hall, Upper Saddle River/New Jersey, 5th edition.

Iyengar, S. S., Jiang, W., and Huberman, G. (2003). How much choice is too much?: Contributions to 401(k) retirement plans. Mimeo.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. The American Economic Review, 93(5):1449–1475.

Lades, L. K. (2012). Behavioral and evolutionary economics on impulsive consumption and reflexive thought: Nudging ethical consumer behavior. Mimeo.

Loewenstein, G. and Haisley, E. (2008). The economist as therapist: Methodological ramifications of 'light' paternalism. In A. Caplin and A. Schotter (Eds.), "Perspectives on the Future of Economics: Positive and Normative Foundations", volume 1 in the Handbook of Economic Methodologies, Oxford, England: Oxford University Press.

Macpherson, C. (1991). Individualism. In Eatwell, J., Milgate, M., and Newman, P., editors, The New Palgrave: A Dictionary of Economics, Vol.II, pages 790–793. Macmillan, London.

Madrian, B. C. and Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. The Quarterly Journal of Economics, 116(4):1149–1187.

Mast, B. D., Benson, B. L., and Rasmussen, D. W. (1999). Beer taxation and alcohol-related traffic fatalities. Southern Economic Journal, 66(2):214–249.

Metcalfe, J. S. (2001). Institutions and progress. Industrial and Corporate Change, 10(3):561–586.

Mill, J. S. (1991[1869]). On Liberty. Routledge, London/New York.

Mitchell, G. (2005). Libertarian paternalism is an oxymoron. Northwestern University Law Review, 99(3):1245-1277.

Nelson, R. R. and Winter, S. G. (1982). An Evolutionary Theory of Economic Change. The Belknap Press, Cambridge/Mass.

Pollak, R. A. (1976). Interdependent preferences. The American Economic Review, 66(3):309–320.

Qizilbash, M. (2009). Well-being, preference formation and the danger of paternalism. Mimeo.

Rabin, M. (1998). Psychology and economics. Journal of Economic Literature, 36(1):11–46.

Rebonato, R. (2012). Taking Liberties - A Critical Examination of Libertarian Paternalism. Palgrave-Macmillan, Basingstoke.

Rizzo, M. J. and Whitman, D. G. (2009a). The knowledge problem of the new paternalism. Brigham Young University Law Review, pages 103–161.

Rizzo, M. J. and Whitman, D. G. (2009b). Little brother is watching you: New paternalism on the slippery slopes. Arizona Law Review, 51(3):685–739.

Schelling, T. C. (1978). Micromotives and Macrobehavior. W.W. Norton, New York.

Schnellenbach, J. (2012). Nudges and norms: On the political economy of soft paternalism. European Journal of Political Economy, 28(2):266–277.

Schubert, C. (2012a). Is novelty always a good thing? Towards an evolutionary welfare economics. Journal of Evolutionary Economics, 22:585-619

Schubert, C. (2012b). Opportunity and preference learning. Papers on Economics & Evolution #1208.

Simon, H. A. (1956). Rational choice and the structure of the environment. Psychological Review, 63(2):129–138.

Sugden, R. (2004). The opportunity criterion: Consumer sovereignty without the assumption of coherent preferences. The American Economic Review, 94(4):1014–1033.

Sugden, R. (2006). What we desire, what we have reason to desire, whatever we might desire: Mill and Sen on the value of opportunity. Utilitas, 18(1):33–51.

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

Sugden, R. (2011). The behavioural economist and the social planner: to whom should behavioural welfare economics be addressed? Papers on Economics & Evolution #1121.

Sunstein, C. R. (2012). The Storrs lectures: Behavioral economics and paternalism. Forthcoming in Yale Law Journal.

Sunstein, C. R. and Thaler, R. H. (2003). Liberaterian paternalism is not an oxymoron. The University of Chicago Law Review, 70(4):1159–1202.

Thaler, R. and Sunstein, C. R. (2003). Libertarian paternalism. The American Economic Review, 93(2):175–179.

Thaler, R. H. and Benartzi, S. (2004). Save more tomorrow[TM]: Using behavioral economics to increase employee saving. Journal of Political Economy, 112(S1):S164–S187.

Thaler, R. H. and Sunstein, C. R. (2008). Nudge - Improving Decisions about Health, Wealth and Happiness. Penguin Books, London.

Trout, J. D. (2005). Paternalism and cognitive bias. Law and Philosophy, 24:393–434.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157):1124–1131.

Vallentyne, P. (2012). Libertarianism. The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2012/entries/ libertarianism/ (accessed: 09.12.2012)

White, M. D. (2008). Behavioral law and economics: The assault on consent, will, and dignity. In: NEW ESSAYS ON PHILOSOPHY, POLITICS AND ECONOMICS: INTEGRATION AND COMMON RESEARCH PROJECTS, Gerald Gaus, Christi Favor, Julian Lamont, eds., Stanford University Press, Forthcoming. Available at SSRN: http://ssrn.com/abstract=1274444

Witt, U. (1987). Individualistische Grundlagen der evolutorischen Oekonomik. J.C.B. Mohr (Paul Siebeck), Tuebingen.

Witt, U. (2001). Learning to consume - a theory of wants and the growth of demand. Journal of Evolutionary Economics, 11:23–36.

Witt, U. (2003). Economic policy making in evolutionary perspective. Journal of Evolutionary Economics, 13:77–94.

Witt, U. (2008). What is specific about evolutionary economics? Journal of Evolutionary Economics.

Witt, U. (2009). Propositions about novelty. Journal of Economic Behavior & Organization.

Witt, U. (2012). Economic behavior—evolutionary versus behavioral perspectives. Biological Theory, pages 1–11.

Witt, U. and Binder, M. (2011). Disentangling motivational and experiential aspects of "utility" – a neuroeconomics perspective. Mimeo.

Witt, U. and Schubert, C. (2013). Verhaltensökonomische und Evolutorische Perspektiven auf Konsumentensouveränität und Wohlfahrt", *Jahrbuch Normative und Institutionelle Grundfragen der Ökonomik*, in press.

Zajonc, R. B. and Markus, H. (1982). Affective and cognitive factors in preferences. Journal of Consumer Research, 9(2):123–131.