

Qizilbash, Mozaffar

Working Paper

Well-being, preference formation and the danger of paternalism

Papers on Economics and Evolution, No. 0918

Provided in Cooperation with:

Max Planck Institute of Economics

Suggested Citation: Qizilbash, Mozaffar (2009) : Well-being, preference formation and the danger of paternalism, Papers on Economics and Evolution, No. 0918, Max Planck Institute of Economics, Jena

This Version is available at:

<https://hdl.handle.net/10419/32663>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

PAPERS on Economics & Evolution



MAX-PLANCK-GESELLSCHAFT

0918

Well-Being, Preference Formation and the Danger of Paternalism

by

Mozaffar Qizilbash

The *Papers on Economics and Evolution* are edited by the
Evolutionary Economics Group, MPI Jena. For editorial correspondence,
please contact: evopapers@econ.mpg.de

ISSN 1430-4716

© by the author

Max Planck Institute of Economics
Evolutionary Economics Group
Kahlaische Str. 10
07745 Jena, Germany
Fax: ++49-3641-686868

Well-Being, Preference Formation and the Danger of Paternalism

Mozaffar Qizilbash,

School of Politics, Economics and Philosophy,

University of York,

York, YO10 5DD

E-Mail: mq500@york.ac.uk

Abstract Informed or rational desire, capability and prudential value list views of well-being - must accommodate human limitations, as well as address issues about adaptation and paternalism. They sometimes address adaptation by toughening the requirement(s) on those desires, satisfaction of which constitutes well-being. That exacerbates a concern that these accounts if adopted will encourage policies which override actual desires and enforce paternalistic restrictions. Sunstein, like Sen, invokes democratic deliberation to address the adaptation problem, and advocates autonomy promoting paternalistic restrictions. Sunstein and Thaler's 'libertarian paternalism' extends this flavour of argument to cover examples of irrationality from behavioural economics. Their variation of the informed desire account involves highly idealized preferences which cannot, in practical terms, guide a paternalistic social planner, but lead to a potentially large range of cases where paternalistic intervention might, in principle, be justified. I argue that the liberal paternalist policy agenda should as currently conceived be resisted.

This is a modified version of a preliminary draft prepared for circulation at the Freiburg workshop on 'New Frontiers in Normative Economics and Policy Analysis, 11-12 December 2009. I thank Robert Sugden for clarifying his position and others at the workshop for their comments on this version of the paper. Comments are welcome. Any error is mine.

Well-being, Preference Formation and the Danger of Paternalism.

1. Introduction

A variety of literatures have recently emerged on the subject of well-being, across economics, philosophy and psychology. One central tension between some strands within these literatures, I argue, relates to the approach they take to human limitations. On the one hand, an important criticism of a leading account of well-being in philosophy and economics – the informed or rational desire (or preference) account – favoured by Henry Sidgwick (1981) and, arguably, John Stuart Mill (1962) through to leading utilitarian thinkers in modern times, like Richard Brandt (1979 and 1992), James Griffin (1986) and John Harsanyi (1981) – is that it takes an unrealistic view of human beings. Human beings, it is argued, simply could not have the sorts of informed or rational desires that some of these accounts suppose that they might have (Rosati, 1995 and Sobel, 1995). But the criticism may apply, in different ways and to different degrees - to a wide range of views of well-being – including accounts which see it in terms of happiness, as well as capability views of the sort advanced by Amartya Sen and Martha Nussbaum. One way or the other, we need an account of *human well-being*, of the well-being of agents with the sorts of limitations and capacities which we humans have.

There is a different, if sometimes related, criticism which relates to the fact that our preferences and desires are malleable or can adapt to the circumstances in which they are formed. They are, to put it another way, ‘endogenous’. There are issues about how to respond to this possibility also, and different responses can lead people to favour one or other view of well-being (Qizilbash, 2006). One final criticism of some views of well-being is that they can undermine the authority of people’s actual desires or attitudes and indeed in so doing open the way to restrictions on liberty which may be seen as paternalistic. This line of criticism is popular with some economists, and variants of it are presented by Robert Sugden (2006 and 2008) and Richard Layard (2005). While this sort of critique has recently been leveled by these commentators at versions of Amartya Sen’s capability approach, it potentially applies to a number of the leading views of well-being. While Sen (2006 and 2009) himself has responded robustly to these commentators, Sen’s capability view may, or may not, be seen as paternalistic in this way, depending on how it is interpreted (Qizilbash, 2009a and b). Nonetheless,

Cass Sunstein's (1991) position which is similar to Sen's but explicitly advocates democratic controls in the market or government regulation when these promote autonomy in the formation of preferences shows how positions like Sen's can be developed in a paternalistic direction.

In a further development, Cass Sunstein and Richard Thaler (Sunstein and Thaler, 2005 and Thaler and Sunstein, 2006) have advanced what they provocatively term 'libertarian paternalism'. This is, on their view, a form of 'paternalism' since it involves interfering with people's choices with a view to improving their welfare, while respecting freedom of choice. When human beings fall short in terms of either informed or rational choice, the results of behavioural economics can help private institutions or the state improve their decision making and increase people's welfare, where this is – as I interpret Sunstein and Thaler - understood in terms of the satisfaction of informed or rational desires. This approach takes a markedly different view of human limitations as contrasted to some of the philosophical literature. It does not take such limitations to constitute grounds for questioning or refining the account of well-being chosen. And it takes the fact that preferences are ill-formed or malleable to be a reason to manipulate these so as to improve a person's welfare along lines not dissimilar to those Sunstein advanced in his discussion of endogenous preferences. In fact, this literature rather reinforces worries expressed by Robert Sugden that some accounts of well-being can encourage restrictions on liberty of a sort that some might consider paternalistic.

In this paper, libertarian paternalism is examined and evaluated in the light of discussions of accounts of well-being and of adaptive or endogenous preferences. In section 1, I consider respectively, certain informed desire (or preference), happiness and prudential value list views of well-being; I consider capability views and Sunstein's discussion of endogenous preferences and politics in section two; in section three I discuss and evaluate Sunstein and Thaler's libertarian paternalism in the light of the discussion and analysis of earlier sections; and section four concludes.

2. Desire, Happiness and Prudential Values List Views

The two most popular accounts of well-being in the utilitarian tradition are desire satisfaction and happiness views. The leading versions of the first of these views are rational and informed, or 'full information' desire views. These go back at least as far as Henry Sidgwick and John Stuart Mill.

Sidgwick's (1981, 111-2) formulation runs as follows: 'a man's future good on the whole is what he would now desire and seek on the whole if all the consequences of all the different lines of conduct open to him were accurately foreseen and adequately realised in imagination at the present point in time'. Mill's account, by contrast, is hedonistic. He defines happiness in terms of 'pleasure and the absence of pain' (Mill, 1962, 157). But he famously distinguished between higher and lower pleasures and argued that those who were acquainted with both – the 'competent judges' – 'showed a marked preference for the manner of existence which employs the higher faculties' (Mill, 1862, 159). His position can be seen as a form of informed desire or preference view (Crisp, 1997, 29) and 'informed preferences' would then be the preferences of 'competent judges'.

Richard Brandt's modern version of this sort of account associates well-being with the satisfaction of rational desires, desires which are not *irrational*. An *irrational* desire is 'one which would *not* survive, in a given person, in the presence of vivid awareness of knowable propositions' (Brandt, 1992, 40). On a fuller formulation of this view, Brandt describes the process of 'confronting desires with relevant information, by repeatedly representing it, in an ideally vivid way, and at an appropriate time' as *cognitive psychotherapy* (Brandt, 1979, 113). A person's desire is then 'rational' if it would survive or be produced by careful cognitive psychotherapy. Brandt characterises the process of confronting desires with relevant information in this way because 'it relies on available information, without influence by prestige of someone, use of evaluative language, extrinsic reward or punishment, or use of artificially induced feeling-states such as relaxation' (Brandt, 1979, 113). His concern is to rule out 'mistaken' desires of various sorts, whether these be generated by ignorance or social conditioning. But there are well-known problems with Brandt's formulation of the desire account. Someone may have had years of analysis and may recognise that his desire to be the centre of attention is bad for him, and yet the desire may persist and indeed may survive cognitive psychotherapy. If so, the satisfaction of his rational desire does not constitute well-being.

James Griffin's version of the 'informed desire' or 'full information' account (Griffin, 1986) attempts to deal with this problem. Griffin has articulated several different formulations of the requirement for a desire to count as 'informed' or 'fully informed'. On one of these 'an "informed" desire is one formed by appreciation of the nature of its object, and it includes anything necessary to

achieve it' (Griffin, 1986, p. 14). However, Griffin identifies 'the technical sense' of 'informed desires' with desires which avoid all the faults he finds with actual desires (Griffin, 1986, 12-14) and these go beyond those faults which relate purely to a lack of information and rationality in any ordinary sense. Griffin then relates 'information' to 'what advances plans of life' and 'information is full' on his view 'when more, even when there is more, will not advance them further' (Griffin, 1986, 14). Griffin's criterion for a desire to be 'informed' - at least in the 'technical sense' of 'informed' - is clearly demanding. It is certainly more demanding than Brandt's criterion for a desire to be rational. But exactly how demanding this criterion is depends on which formulation of it one focuses on. The key to the informed desire account is that it connects desires with the constituents of well-being – the things that make a life go better – which Griffin calls 'prudential values'. His list of such values includes: the constituents of a characteristically human life - freedom from great anxiety, basic capabilities, minimal material provision, liberty and autonomy - accomplishment (the sort of achievement that gives a life point and weight), deep personal relations and enjoyment (Griffin, 1996, 29-30). If informed desires are simply desires which involve a proper appreciation of their objects, which in turn are prudential values, or realisations of such values - such as specific forms of accomplishment or enjoyment - then it may not require a *great deal* of information to appreciate the relevant sorts of thing properly. So if the key requirement is that the desire is formed with a proper appreciation of its object, then it may not be so demanding.

Nonetheless, it is unsurprising that a critical literature has grown surrounding the requirements on desires imposed by rational and informed desire accounts. These accounts – as Connie Rosati (1995) and David Sobel (1995) argue - sometimes require superhuman levels of rationality and indeed capacities for information acquisition and retention which are well beyond human beings. It does not matter if the information requirement is merely counterfactual. That is, it does not matter if the desires which connect to well-being are ones one would have *if* one had such information or such levels of rationality. Human beings cannot, it might be argued, have the requisite sorts and levels of capacities as regards rationality and information acquisition and retention. And when these accounts take the form of 'ideal observer' accounts – such as the account developed by Peter Railton – the idealisation of desires is even more explicit. On Railton's view we are to imagine an idealised version of the

person. Suppose we are concerned with A. We must give him ‘unqualified cognitive and imaginative powers, and full factual and nomological information about his physical and psychological constitution, capacities, circumstances, history, and so on.’ (Railton, 1986, 173). A has then become A+, an idealised version of A with ‘complete and vivid knowledge of himself and his environment, and whose instrumental rationality is in no way defective’ (Railton, 1986, 174). A+ is to advise A about what he would want A to want. A+ can be seen as A’s ‘ideal advisor’. It is the satisfaction of these desires – the desires A+ would want A to want - which on Railton’s account constitute well-being. On a formulation like Railton’s it is plausible that the desires which are required by informed desire or ideal advisor accounts are well beyond humans, so that the relevant accounts need to show how they can fit the limited human beings whose well-being we are concerned with.

What of the other two criticisms of accounts of well-being which I mentioned earlier? One is that people – such as underdogs - can adapt to their restricted conditions and can cut their desires to be more easily satisfied, or find happiness in small mercies. But the satisfaction of their desires may not, in these circumstances, constitute well-being. Put another way, adaptation of this sort might make the metric of desire satisfaction misleading. This argument is associated with Amartya Sen and the problem it poses for accounts of well-being (or normative evaluation more broadly) is the ‘adaptation problem’ (Qizilbash, 2006 and 2008). One response to this problem – associated with James Griffin - is simply to argue that *any* gap between an actual desire and the proper appreciation of its object is captured by the information requirement – so that if people are fully informed their desires would not take the adaptive form they do in Sen’s examples. Instead they would be desires formed with ‘legitimate expectations’ (Griffin, 1986, 88). On this view, informed desires would be *defined* in such a way that they would never reflect diminished expectations. This might seem like an easy, formal, solution to the problem and as we will see Sen’s own discussion of how to deal with the problem of adaptive desires has a similar flavour. Nonetheless, some, like Richard Arneson do not believe that adaptive desires can be addressed by an information requirement – after all it is precisely the information people have about their life situations which leads them to adapt. Making that information fuller or more vivid may not address the problem (Arneson, 1999, 133).

An alternative approach supposes that adaptive desires might involve a failure of autonomy so that an autonomy requirement might be added to the requirement that desires are informed. Jon Elster's discussion of 'adaptive preferences' attempted to define an autonomy requirement of this sort on preferences (Elster, 1981, 229). As we shall see this strategy can also be pursued in the context of happiness views. In both strategies addressing the adaptation problem – whether it strengthens the information requirement as Griffin does or adds an autonomy requirement - leads to the conditions which a desire must meet to connect with well-being becoming stricter, and the desires themselves potentially even more remote from the human beings whose well-being we are concerned with. So attempting to address the problem seems to heighten, rather than lessen, the worry that the requirement for desires to connect to what is good for people is rather strong. That exacerbates, rather than alleviates, the problem posed by the criticism that the criterion for treating desires as informed and rational is too strong given human limitations.

What about the second criticism which suggests that these views are paternalistic? It is certainly true that they disconnect what is good for people from the satisfaction of their actual desires, and associate it with the satisfaction of informed or rational desires. But who decides on which desires are rational or informed, so that their satisfaction connects with what is good for one? or connects better with what makes one's life better? Robert Sugden's worry – which is more fully developed in the context of capability views and those more explicitly paternalistic views which build on results in behavioural economics are discussed in sections 2 and 3 - is that philosophers, experts or society may decide this, and impose restrictions on liberty as a consequence. The argument might run as follows. If people knew the consequences of eating unhealthy food or smoking, they would realise that it is not good for them. Given that people are limited in rationality and imperfectly informed, society might then impose restrictions on smoking or on eating some sorts of food. That would be paternalistic. But notice that it is not at all obvious that, even following the logic of the informed desire account, one would necessarily do this. One might instead simply provide people with better information (e.g. about the effects of smoking or eating unhealthy food) or help them in some way (e.g. help them to give up when they are addicted to smoking certain substances or provide information on diets and nutrition). Nonetheless, this sort of account of welfare might be adopted by those who have recently

taken explicitly paternalistic stances: as we will see in section 3, on my reading, Sunstein and Thaler arguably subscribe to an informed desire view of welfare and do so in advocating their ‘paternalistic’ position. So it is hard to dismiss the worry that informed and rational desire accounts do *encourage* paternalism, even if they do not in themselves necessarily *imply* a paternalistic position.

There is an even easier way to respond to the paternalism criticism within James Griffin’s version of the informed desire account. One can argue that autonomy – making one’s own way in life – and liberty are prudential values – and so the objects of informed desire. If they are, the objects of informed desire then there is indeed no problem – the informed desire account would allow people to make their own mistakes – that is part of what the value of autonomy requires. On Griffin’s account restrictions on autonomy or liberty actually would (other things being equal) lower well-being. So the paternalism criticism may have limited bite when it is applied to this version of the informed desire account.

It should be clear from this discussion that a large part of the work being done in Griffin’s version of the informed desire account is done by the actual list of prudential values he proposes. These are the constituents of well-being – the fact that they are the objects of informed desire does not do much work when it comes to identifying them or in working out what constitutes progress in prudential deliberation – progress in deliberation about prudential values. For example, what makes accomplishment prudentially value has very little to do with the fact that it is the object of a desire – rather it is a matter of what fulfils a human life. Griffin (1986, 65) argues that what fulfils a human life is not just any achievement but one of the sort that gives a life point and weight. So the fact that a desire is fulfilled does very little work on this account – in fact it is not obvious that it is a desire account at all. Griffin (1986, 34) himself hesitates before calling it an ‘informed desire’ account. Perhaps then, it is best to think of Griffin’s account as a version of what Derek Parfit terms an ‘objective list theory’, or - given that Griffin (1996, 35-6) sees the opposition between ‘objective’ and ‘subjective’ as misleading - what one might call a ‘list theory’ (see Qizilbash, 2006, 96) – a theory of ‘what is good or bad for us, whether or not we want the good things, or want to avoid the bad things’ (Parfit, 1984, 493). If so, the three criticisms of accounts of well-being being considered here may take a different form. The problem of human limitations might have less force, since the rationality or

information requirement is of lesser importance. The kinds of prudential deliberation involved in identifying the constituents of well-being do not seem superhuman, nor do they seem to make excessive informational demands, though we do need a picture of a human life and what makes such a life go well (Qizilbash, 2006, 95-8). How about the adaptation problem? It also seems to have less force since its critical bite derives from the fact that well-being is, on the informed desire account, connected with desire satisfaction. If what makes a life good is not the mere fact that a desire is fulfilled, it is not clear that the critique has the same bite. The lives of the various disadvantaged people Sen mentions in explaining the adaptation problem are clearly short on prudential values (see Qizilbash, 1998), so that Griffin's account looks less vulnerable on that score as well (see Qizilbash, 2006, 98-101). What about the paternalism problem? Surely this applies – one might argue – because, by articulating a list of prudential values, Griffin is telling us what is good for us, or makes our lives go better. And it provides a basis for government policy to promote those values, whatever people actually desire. However, Griffin's account is more nuanced. Since autonomy is a prudential value, Griffin allows people to advance their own lists – and allows for the possibility that people may have different lists, and indeed disagree with his proposed list (Griffin, 1996, 30). So including autonomy on the list and not insisting on his particular list helps block that criticism. So it may be that a list view is more defensible in relation to the three criticisms discussed here than desire views. I shall return to list views in section 2 when discussing Nussbaum's version of the capability view.

When one turns to happiness views it becomes clear that some of the problems which arise for desire views can also apply to these. Consider, for example, Wayne Sumner's version of the happiness view (Sumner, 1996). Sumner suggests that welfare 'consists in authentic happiness, the happiness of an informed and autonomous subject' (Sumner, 1996, 172). 'Happiness' here refers to a positive evaluation of the conditions of one's life, 'a judgement that, at least on balance, it measures up favourably to your standards and expectations' (Sumner, 1996, 145). The information requirement used in this account relates to the 'relevance' of information, which in turn relates to 'whether it would make a difference to a subject's affective response to her life, given her priorities' (Sumner, 1996, 160). Sumner thinks that if someone's endorsement of her life is factually uninformed, or misinformed, 'that gives us one reason for doubting its authority' (Sumner, 1996, 161). While this

information requirement is perhaps less demanding than those adopted in some informed desire views, it clearly does some of the same work. Sumner argues, furthermore, that an information requirement cannot adequately deal with the adaptation problem. He takes the issue seriously because, according to his account, the evaluation of one's life relates to one's 'standards and expectations'. If these standards and expectations have been seriously affected by disadvantage, that must distort the metric of happiness. He writes that: 'the requirement that endorsement be empirically informed will not suffice to exclude ... social influences on the standards by which people judge how well their lives are going; the problem here is rooted not in the adequacy of people's factual information but in the malleability of their personal values' (Sumner, 1996, 162). He addresses the problem by introducing a requirement that the subject be autonomous in a specific way. Without going any further (see Qizilbash, 2006, 90-95 for a fuller discussion) it should be clear that the same criticisms about idealization and paternalism which potentially apply to desire accounts can also be leveled at this version of the happiness view when it genuinely attempts to respond to the adaptation problem and issues relating to evaluations which are ill-informed.

3. *Capability Views, Democracy and Preference Formation*

The capability approach was initially developed by Amartya Sen in his Tanner lecture, 'Equality of What?' (Sen, 1980) in the context of egalitarian justice. It has also been applied to well-being. The capability approach sees well-being in terms of an evaluation of 'functionings', which are beings and doings, or states of a person, so that lives are constituted by combinations of such functionings. The capability of a person is, in a technical sense, the range of lives made up of combinations (or *n-tuples*) of functionings from which one can choose one (Sen, 1993, 31). In this sense it reflects the *opportunities* open to her. However, capability can also be understood at a more intuitive level in terms of the ability to do and be various things – where capability typically refers to a particular *power* or *ability* – such as the ability to achieve minimal nutrition or to appear in public without shame. These definitions are different, and Sen (2009, 233) has recently admitted that they are both at work in his writings. In practice the difference between these definitions is not important to any claims made in this paper. I shall for the most part treat this 'technical sense' of capability as the

‘official sense’ of capability in Sen’s work. The quality of a person’s life can then be seen in terms of her opportunities or what she can be and do in living a life.

Invariably the question arises: which functionings are valuable and constitute good lives on Sen’s account? On this subject, Sen initially had very little to say, beyond listing some uncontroversial functionings which might be relevant to different sorts of analysis (see, for example, Sen, 1993, 31). He intentionally left his account open to different views of the good life or of evaluation more broadly. By not explicitly stating what is good or bad for people, or providing a list of valuable functionings, in this ‘thin’ version of the capability approach – that is, a version which does not say much about evaluation and application of the approach (Qizilbash, 2007) - he appeared to avoid the paternalism criticism. There was simply no basis for restricting liberty on the grounds that there are things which are good or bad for people, since Sen’s capability approach does not offer any definitive list – and for that reason cannot be classified as an objective list theory. Nor does this approach appear to take a particularly idealized view of people – it is motivated specifically to focus on *human* lives, and the evaluation of what people can do and be in living a human life. So – at a first glance - the approach seems well constructed to deal with issues relating to idealization and paternalism. But at the same time it can be seen – at least in its initial formulations – as somewhat underspecified or empty of content.

While Sen himself posed the adaptation problem for desire and happiness views, it also potentially raises a problem for his own account if it remains largely silent on the evaluation of functionings or does not offer a list of these. In her very first discussion of capability – in the context of a discussion of Aristotle’s view of political distribution – Martha Nussbaum (1988) raises the point that if Sen leaves the evaluation of functionings to people and does not propose an objective list of his own, the capability approach can itself fall prey to the adaptation problem. In particular, it is not clear why people’s own lists of functionings may not reflect adaptation if they are left to themselves. A deprived person’s list may reflect his situation, or the list of a woman who has been brought up in a society where gender bias is entrenched may reflect cultural indoctrination or be shaped in some other way by gender bias. The mere fact that the approach focuses not on desire or happiness but requires

an evaluation of functionings is of little significance in this context: valuations can themselves be influenced and shaped by the situation or culture one finds oneself in.

Nussbaum's own initial response to this problem was to develop a list of valuable capabilities, which were derived from an Aristotelian picture of the good human life (see Nussbaum, 1990, 1992, 1995a and 1995b). This was a version of the objective list view: the components of the good human life are derived from a view of what makes us humans as opposed to gods or non-human animals. The approach seems to deal well with adaptation and also – in its initial articulation – with idealization. There is no assumption that people are superhuman – in fact the approach is actually very focused on what makes a *human* life good. The approach can nonetheless be criticized on the grounds that it is paternalistic and tells us what is and is not good for people on an Aristotelian account and might restrict liberty on that basis. Indeed Sugden's critique has been leveled at Nussbaum's view which Sugden sees as a potential development of Sen's view, given Sen's reluctance to advance a list of functionings (Sugden, 2006, 41). One potential criticism, as we earlier saw, of any objective list view is that it must be paternalistic because it simply lists what is good for people or makes their lives better. Will it, in practice be paternalistic by restricting liberty and by imposing an Aristotelian account of the good life on people? One reason Nussbaum's approach can dispel this worry is that it makes capability not functioning the goal for responsible adults. So the approach only aims to give people the opportunity to act in certain ways – it does not matter if people do not actually take advantage of the opportunity. For example, even if the capability to live a healthy life is guaranteed or promoted as an item on the list in Nussbaum's approach, she is not concerned if people fall short in terms of functioning – i.e. if they smoke or eat junk food and are unhealthy as a consequence – as long as they had the capability or freedom to do otherwise. Other ways in which Nussbaum avoids telling people what is and is not good for them is by allowing them to further specify the list, which is left incomplete and vague so that it can be multiply specified by different people in different contexts. Finally, Nussbaum lists capabilities relating to practical reason and choice in her capability list – so that the approach recognizes the importance of autonomy (in the way that Griffin's list does). These aspects of Nussbaum's approach seem to address the paternalism worry to some degree, though in her later work various other issues arise.

In these later writings (see Nussbaum, 1998, 1999, 2000, 2006) Nussbaum argues that a modified version of the list she earlier advanced can be the basis of an overlapping consensus between people with different views of the good life. It is now seen as part of what John Rawls called a ‘political conception of justice’, i.e. one which is ‘free standing’ in not being based on any metaphysical views or views of the good life (Rawls, 1993, 9-13). The approach is also more fully developed as an alternative to Rawls’ theory of justice in these recent writings. But there are at least two areas where issues relating to informed desires and the list are important. Firstly, there is the question of whether political principles based on the list will be stable if items on the list have no connection with desires. Secondly, there is the question of how the list can be justified to people whose desires are not connected to items on the list. In this context, Nussbaum considers issues relating to the list, preference formation and informed desires. She thinks that informed desires can have a role and if there is a convergence between the objects of informed desires and items on her list, then informed desire views can have a procedural role. The list can be checked against the informed desires of particular women in the context of international feminism to check that the two are not too far apart. If there is some convergence then that is good for stability and justification. When she adopts an informational criterion, the problem that the criterion may be too demanding arises. She offers us some hints about the sort of criterion she has in mind. She writes that:

When people are respected as equals, and free from intimidation, and able to learn about the world, and secure against desperate wants, their judgements about the core of a political conception are likely to be more reliable than judgements formed under the pressure of ignorance and fear and desperate need.) Informed desire plays a large role in finding a good substantive list, for epistemic reasons (Nussbaum, 2000, 152)

Nussbaum treats some of those she has spoken to in revising the list as informed in the relevant sense – so that having an informed desire is not in her approach something which is beyond people. Nonetheless, by borrowing the language of informed desire, she does potentially open herself up to the question: which women are, and are not, to be treated as informed? Who decides? The above discussion however points to the absence of various things (ignorance, fear and intimidation) and does not suggest a very ambitious requirement on human capacities or informed desires.

Having given informed desire this procedural role, Nussbaum (2000, 155-165) goes on to worry that informed desires may not address the adaptation problem. Where there is entrenched gender bias, preferences based on the status quo may be resilient, especially when those who dominate society may not want to give up their privileges. Moral education might potentially help to bring people's informed desires into line with the list of capabilities which she has advanced. But Nussbaum guesses that this may take some time – possibly generations - and in the absence of convergence one should, on her view, stick to the list. Nonetheless, Nussbaum would not stop people acting on adaptive preferences – as long as this does not harm others. Here again her position remains libertarian and anti-paternalistic, and indeed invokes Mill's harm principle. So on this view, Nussbaum's later version of the approach addresses worries about paternalism fairly well. It does not in general see adaptive desires as a reason to interfere in people's lives.

In the light of Nussbaum's and other criticisms, Sen's later writings (Sen, 1999, 2004, 2005 and 2009) have gone further in addressing evaluative issues and the version of the capability approach advanced in these writings can be seen as constituting a 'thick view' in as much as they have more to say about evaluation and application (Qizilbash, 2007). A crucial dimension he adds to his earlier discussions of the capability approach is an emphasis on public reasoning and democracy. He gives such reasoning and discussion a central importance arguing that evaluative decisions for social decision making should be made in a democratic way, however messy this might get (Sen, 1999, 78-9). Another difference between Sen's earlier and later writings is that the functionings which are valued or relevant to the evaluation of development or the quality of life or egalitarian justice are not merely functionings which people *value*, but ones they also have *reason to value* (see, for example, Sen, 2009, 231). The introduction of these further elements can help address problems of ignorance and adaptation. Firstly, if people's valuations are ill-informed, they will not necessarily value what they have reason to value. Secondly, if they have adapted to their living conditions, they may have values which do not adequately reflect what they have reason to value. In fact, here the 'value and reason to value' formulation plays a role very similar to that played by information and rationality criteria in informed and rational desire accounts.

How does Sen ensure that people's desires do reflect what they have reason to value? Public reasoning is important here. First, if people's perspectives are narrowed by their living conditions and they intentionally cut their desires to reflect what is possible within their culture, public discussion may help broaden their perspectives and may lead to more informed and reasoned judgements. Secondly, if the form public reasoning takes allows in voices from outside people's own nation or culture and is open to 'distant voices' that may also address the issue to some degree. For example, if the voices of feminists outside some country are introduced into public discussion in a context where there is a culture characterized by entrenched gender inequality, that might help to raise the consciousness or broaden the perspectives of people who may have begun with a narrow perspective reflecting an adjustment to their living conditions or indeed a parochial view which can be broadened in the face of discussion. It should be clear here that public reasoning, on this picture, has an epistemic role in shaping better informed desires. There may, nonetheless, be little in the way of idealization of the desires people have: no claim is made that people's desires will fully adjust in the light of public discussion so that they are fully informed, rational and non-parochial.

Once it is clear that this approach uses public reasoning to achieve some of what informed and rational desire accounts do through an idealized information requirement, it is also clear that it is open to the criticism that it is paternalistic. In fact, in some of his writings Sugden takes Sen's approach to be nothing other than a version of the rational or informed or rational preference view (Sugden, forthcoming). The criticisms which, I earlier suggested, Sugden would level at the informed or rational desire accounts now apply to Sen's account in a modified form. Public reasoning may yield a consensus, or judgement, about what it is best for us, or of the functionings which should be promoted or equalized. The collective judgement may, in turn, run counter to our individual desires as they are. Restrictions on liberty may again result where an individual's actual desires may be overridden because they do not reflect what she has reason to value, as this is judged by society.

How can or does Sen respond to this criticism? There is not space here to deal fully with this issue, and Sen's brief but robust response has been published (Sen, 2006). A great deal depends on whether Sugden's criticism is leveled at Sen's discussions of well-being and freedom more broadly or just his capability view (see Sugden, 2008 and forthcoming as well as Qizilbash, 2009 b). But if it is

simply the capability approach understood as a thick view, then it can be criticized along the following lines. First, Sen does not specify the weight to be given to capability which captures what he calls the ‘opportunity aspect’ of freedom, as compared to other aspects of freedom – such as what he calls the ‘process aspect’ of freedom – which, for him covers freedom from interference, and freedom of choice. But he accepts that opportunity and process freedom can overlap. If that is so, the promotion of opportunity or capability – say to lead more healthy lives – may involve restrictions on liberty – e.g. on smoking in public or on the use of hard drugs or indeed on consumption of certain sorts of unhealthy foods - even within the capability approach. Suppose then that there is a conflict between these values - opportunity and process - or aspects of freedom. The capability approach – as a thick view - would presumably see the resolution of this conflict as a matter of social choice – and leave society to decide whether these restrictions are imposed through a process of democratic discussion or public reasoning. That seems to be a standard case of dealing with conflicts of value within a democratic system, and may not best be seen as ‘paternalistic’. An alternative response, however, is simply to accept that the capability approach may be paternalistic in this sort of case, but to argue that it accepts paternalistic restrictions only in certain circumstances, that is, when they promote freedom or autonomy. That idea can be traced back to John Stuart Mill himself according to Gerald Dworkin (1971). One can argue that if one takes this line, the specific restrictions the approach will impose may not be particularly objectionable since they are freedom promoting. I shall call this the claim that the relevant kind of paternalism which promotes freedom is ‘unobjectionable’.

A similar, if not identical, argument to Sen’s is advanced by Cass Sunstein in his paper ‘Preferences and Politics’. In fact, Sunstein follows through an argument similar to Sen’s, but to a conclusion which Sugden would see as definitely paternalistic. Sunstein is in part addressing the problem of adaptation as it applies to preference views or what he terms ‘subjective welfarism’. For Sunstein (1991, 7) on this view ‘government, even or perhaps especially in a democracy, should attend exclusively to conceptions of welfare as subjectively held by its citizens’. Sunstein’s discussion is concerned with ‘adaptation’ and sees preferences as ‘endogenous’ in the sense that ‘they are adaptive to a wide range of factors – including the context in which the preference is expressed, the existence of legal rules, past consumption choices, and culture in general’ (Sunstein, 1991, 5). For

Sunstein, because laws and entitlements in part shape preferences, it is impossible to defend those laws or entitlements in terms of our preferences as they are (Sunstein, 1991, 9). In specific cases, Sunstein argues that there is a case for government interference. He writes that ‘respect for preferences that have resulted from unjust background conditions and that will lead to human deprivation or misery hardly appears the proper course for liberal democracy’ (Sunstein, 1991, 10). Given the malleability of preferences, Sunstein thinks that there is a range of cases where government interference and controls – such as controls on addictive substances (cigarettes and heroin) and support for public broadcasting of high quality programmes – can increase welfare (Sunstein, 1991, 11). He associates the notion of autonomy with ‘decisions reached with a full and vivid awareness of available opportunities, with reference to all relevant information, and without illegitimate or excessive constraints on the process of preference formation’ (Sunstein, 1991, 11). When these conditions are not met, he argues that promotion of preference satisfaction and autonomy do not connect. Rather one goal of democracy is to ‘ensure autonomy’ in the ‘processes of preference formation’ (Sunstein, 1991, 12).

On Sunstein’s model of politics ‘political choices will reflect a kind of deliberation and reasoning, transforming values and perceptions of interests, that is often inadequately captured in the marketplace’ (Sunstein, 1991, 16). Like Sen, Sunstein argues that democracy can, through deliberation bring ‘additional information and perspectives to bear, [and] may affect preferences as expressed through government processes. A principal function of democracy’ he argues ‘is to ensure that through representative or participatory processes, new and submerged voices, or novel depictions of where interests lie and what they in fact are, are heard and understood ...’ and he adds that ‘[i]t should hardly be surprising if preferences, values, and perceptions of both individual and collective welfare are changed as a result’ (Sunstein, 1991, 18). The argument clearly leads to unproblematic interventions when there is unanimity about collective judgements. But if a minority disagree with the majority, there is more likely to be a problem (whether the judgement relates to controls on violent pornography or smoking or some other intervention). Finally, Sunstein argues that there are important cases where the collective judgement cannot be decisive in this way. Nonetheless, it is clear that on Sugden’s criticism this approach would be problematic, in the way that Sugden thinks Sen’s approach

would be if fully developed, even if for Sunstein ‘[t]he argument from democratic controls in the face of endogenous preferences must rely on a belief that welfare or autonomy will thereby be promoted’ and he adds ‘[u]sually, governmental interference should be avoided’ (Sunstein, 1991, 24).

Sunstein does not adopt a capability approach in developing this argument, but his prominent discussion of adaptation shows that it is powerfully motivated by the adaptation problem. When there is adaptation to unjust conditions he thinks that ‘citizens in a democracy might override existing preferences in order to foster and promote diverse experiences, with a view to providing broad opportunities for the formation of preferences and beliefs and for distance on the scrutiny of current desires’ (Sunstein, 1991, 19). He suggests that sometimes government controls are needed to ‘promote divergent conceptions of the good and to ensure a degree of reflection on those conceptions’ (Sunstein, 1991, 20). He finds foundations for the pursuit of this sort of goal in some views including an Aristotelian position – and he cites Nussbaum’s work (Nussbaum, 1990) – arguing that if this approach were adopted in the context of his claims, the ‘governing goal would be to ensure that individual capacities and capabilities are promoted and not thwarted by government arrangements’ (Sunstein, 1991, 20). This discussion suggests that Sunstein’s position – which is certainly of the sort which Sugden finds objectionable – is a close relation to Sen’s and can also find a basis in Nussbaum’s earlier Aristotelian version of the capability approach. The paternalism involved is one which can override people’s actual desires, and restrict their liberty, precisely in order to promote autonomy and welfare. So it takes the form of what I earlier termed an potentially ‘unobjectionable’ form of paternalism. Nonetheless, Sunstein sees ‘direct coercion’ as ‘[t]he most intrusive option, to be used rarely’ (Sunstein, 1991, p. 24). So even if it is hard to show conclusively that capability views are paternalistic, Sunstein’s argument shows how one line of argument in Sen’s work on capability – involving a reliance on democracy and public reasoning - can lead to an unambiguously paternalistic position even if Sen himself would not himself follow that line himself to such a position.

4. Libertarian Paternalism

The idea that there are forms of paternalism which nonetheless can promote, or respect, freedom, as well as the idea in Sunstein’s work that preferences are never given but are always a product of the

environment people find themselves in are further developed in Cass Sunstein and Richard Thaler's 'libertarian paternalism', which contrasts with views discussed thus far in as much as they focus almost exclusively on errors in rationality which are investigated in economics and psychology. Similar claims are advanced in the work of Colin Camerer, Samuel Issacharoff, George Lowenstein, Ted O'Donohue and Matthew Rabin (2003) which advances a different form of paternalism ('asymmetrical paternalism') which makes no claims to being libertarian. Two aspects of Sunstein's earlier work are relevant: one is that preferences are always the product of the environment and so to the degree that that environment is a factor which forms preferences, and is 'chosen' (if only as a default) by society, 'interference' in people's actual preferences is 'inevitable'. And if interference in people's preferences is 'inevitable', then we ought to favour forms of interference which promote welfare. That makes their position 'paternalistic'. The second aspect is the claim that, interference in people's choices is only advocated when it does not block freedom to choose. This makes the paternalism they are discussing 'libertarian' in the sense that it is a form of paternalism which should not be objectionable from the point of view of libertarians. Their aim is to improve people's decision making by taking into account the findings of behavioural economics and psychology, with a view to improving welfare.

One additional feature of the approach is that Sunstein and Thaler argue that some anti-paternalistic claims are actually mistaken. So while many of the views discussed above – including those of Sen and Nussbaum – see the paternalism issue as a problem which their approaches must address, Sunstein and Thaler provocatively accept and embrace it, seeing it as a consequence of the endogeneity of preferences. This makes their position a straightforward extension of Sunstein's earlier work, though, as we shall see, the kinds of example which are used in this literature are quite different. Another important aspect of Sunstein and Thaler's view is the implicit definition of welfare. Sunstein and Thaler are most concerned about poor decision making. They write that 'in some cases individuals make inferior decisions in terms of their own welfare – decisions they would change if they had complete information, unlimited cognitive abilities, and no lack of self control' (Sunstein and Thaler, 2005, 176). On my reading, it is clear that – while they do not explicitly endorse this –

they implicitly adopt a version of the informed or rational desire or preference view of welfare. Given no obvious alternative, I assume this for the remainder of the paper.

One way in which Sunstein and Thaler's position differs from Sunstein's earlier work is that the examples used are typically cases of 'irrationality' which are discussed in the economics literature, usually where agents systematically violate the axioms of expected utility theory. The key issue is not that preferences are formed in the light of, or adapted to, poor opportunities. Nonetheless, as with Sunstein's earlier work, they give prominence to cases where the existing state of affairs – especially in the form of 'default rules' – can influence decisions. In particular, they focus on well-documented cases of status quo bias, framing, myopia, procrastination and addiction *inter alia*. One or two examples should suffice to give a flavor of their analysis. In the case of savings for retirement plans, one default rule is that employees must 'opt-in' to savings plans and are sent some information about these with a form to fill in. Alternatively, they can be automatically enrolled on the plan as the default rule while being given the option of 'opting-out'. It turns out that the latter default rule results in much higher levels of savings. The default rule and indeed the status quo matter to the decisions people make in a way that standard economic theory (i.e. expected utility theory) does not predict. To take another example, the way in which food is presented in a cafeteria where employees queue up to be served food will influence people's choices. Food that is presented earlier in the queue and at eye level is more likely to be chosen. That is, people's choices are based on how the options are presented, or framed. In both cases, libertarian paternalists will recommend that private or government organizations should, in these circumstances, adopt procedures which improve decision making – in the one case by adopting a default rule which will counter employee myopia, while in the other presenting food in such a way as to help people make healthier choices. In both cases, the relevant organization would be paternalistic – in the sense that it would be adopting procedures with a view to promoting welfare (over a lifetime), given potential irrationality in people's preferences and choices – while not blocking any choices – by allowing people to opt-out of the savings plan in one case, and not actually banning various forms of junk food in the other. Since these interventions do not restrict liberty, they may not even be treated as paternalistic on some definitions (where paternalism involves some sort of restriction on liberty).

Sunstein and Thaler also argue that anti-paternalist arguments involve one false assumption and two misconceptions. The false assumption is that ‘almost all people, almost all the time, make choices that are in their best interest or at the very least are better, by their own lights, than the choices that would be made by third parties’ (Sunstein and Thaler, 2005, 178). They argue that this is a claim that is both testable and false. Thaler and Sunstein (2006, 259) suggest that people are better at making decisions about subjects on which they are well-informed and experienced (such as choosing between ice cream flavours) than in cases where they are not (such as choosing between financial or savings plans). This looks like a variation of Mill’s claim about ‘competent judges’, where ‘competence’ in effect involves having certain sorts of experience of, and information about, the objects of choice. Sunstein and Thaler suggest that data on obesity as well as on smoking and drinking demonstrates that people are not necessarily making good decisions, or decisions which are in their interests, all or most of the time. But behavioural economics provides their main examples of irrationality. There are two ‘misconceptions’ relating to paternalism that they reject. The first is that, given that preferences depend on the existing state of affairs, paternalism is avoidable. For them, by contrast, ‘paternalism is inevitable’. The second ‘misconception’ is that ‘paternalism’ necessarily involves coercion – so that it would of necessity be incompatible with a libertarian concern with freedom of choice.

While it is not clear that the anti-paternalist would necessarily argue that ‘almost all people, almost all the time’ are making good decisions – they may make a much weaker claim - the evidence they cite does indeed suggest that, if health is a component of well-being, people are often not choosing to eat or drink in a healthy way. However, it is not clear that Sunstein and Thaler establish that a third party would necessarily make better choices for people. How would one conceive of the third party? One way would be to think of the third party as a kind of ‘ideal advisor’ of the sort Railton describes in his version of the informed desire account. We have earlier seen that it is not clear that such an advisor really can be thought of as a human being at all given the limitations human beings have and the extraordinary capacities and knowledge such an advisor would need to give advice on an informed desire view. On the other hand, one might think of the third party as simply being better informed about the relevant choice, or more experienced in the relevant area, than the chooser: the model may be a version of the advisor as a ‘competent judge’. In this second case, the

obvious thing for the chooser - who is aware of her own limitations - to do is to seek advice from a competent expert or a group of experts. But on that model, it is the chooser's decision to make, as is the decision about how much or little advice she wishes to take, or about how much information she wishes to acquire, before making a final decision. Taking Sunstein and Thaler's approach a step further, the suggestion may be that people will not make these decisions well either. Someone must decide for them how much advice to take or how much information to acquire before making a decision which will be in their interests or promote their welfare. Furthermore, some third party may also need to decide on how limited the chooser's capacities for rationality are, and when they need to defer decisions to better informed others in various contexts. Presumably, she will not be informed or rational enough to make that decision well either. So there is a danger of an infinite regress. It seems likely that people may need a great deal of help with decision making, not just a few 'nudges'.

How about the capacities of the third party who (or which) must advise poor decision makers in various circumstances? Aside from the fictional 'ideal advisors' in some informed desire accounts, the sorts of advisors needed to help poor choosers act in their best interests – as conceived by Sunstein and Thaler - i.e. the way they would 'if they had complete information, unlimited cognitive abilities, and no lack of self control' are not obviously at hand. Even if ordinary people make mistakes and are seen as 'irrational' on this criterion, furthermore, it is not clear that third parties will do any better. The argument against the anti-paternalist seems weak if it is combined with the informed or rational desire account of welfare they implicitly adopt, which does not allow for human limitations in setting the criteria for people to be informed and rational. Given human limitations, and the earlier discussion of informed desire accounts, it seems likely that on Sunstein and Thaler's account *nobody* could or would ever necessarily act in their best interests. One does not need much evidence to make that claim. One simply needs to accept that human beings have limitations and to set the standard of rationality or for being well-informed – as they do - at so high a level that no human being (whether it is the person whose choice affects her own life or a third party) can meet it.

An alternative argument against paternalism which they cite – and ascribe to John Stuart Mill in a footnote (Sunstein and Thaler, 2005, 181) – rests on autonomy and suggests that people should be able to make their own mistakes. Sunstein and Thaler argue that their position allows for freedom of

choice within the libertarian ‘arm’ of their proposal. However, if the worry about paternalism is conceived of in these terms, nothing in the mistakes which people make actually undermines the anti-paternalistic argument. Irrationality of the sorts they mention, furthermore, do not necessarily support paternalism understood as interference in choice with a view to improving welfare. What is true, of course, is that the practical proposals Sunstein and Thaler advance do not block choice, and to that degree do not undermine autonomy. That is what they presumably mean when they say that autonomy is respected by the libertarian aspect of their proposal. But that is not what is denied by those whose anti-paternalistic position depends on the argument from autonomy. To take an example of an account of autonomy, consider that advanced in the context of Griffin’s prudential value theory – ‘[o]ne element of agency is deciding for oneself. Even if I constantly made a mess of my life, even if you could do better if you took charge, I would not let you do it. Autonomy has a value of its own’ (Griffin, 1986, 67). That sentiment does run contrary to even the weak form of paternalism adopted in Sunstein and Thaler’s view. Indeed this variety of libertarian sentiment remains anti-paternalistic so that the idea of ‘libertarian paternalism’ as they define it remains potentially incoherent.

According to Sunstein and Thaler, their proposal remains paternalistic even though it does not block choice. The idea that paternalism requires coercion is, as mentioned earlier, a ‘misconception’ on their view (Sunstein and Thaler, 2005, 180). Its paternalism lies in the fact that it aims to promote people’s welfare by helping them make better decisions. Earlier, when discussing Sugden’s variant of the paternalism criticism of various accounts of well-being, I took restrictions on liberty to be a necessary condition for some intervention to count as paternalistic. Sunstein and Thaler argue that in the examples they focus on, such as the cafeteria example - what makes intervention paternalistic is the reorganization of the presentation of products with a view to making people better off, and since there is no restriction on liberty, libertarians should not object. In the terms used earlier one can either see the intervention as ‘unobjectionably’ paternalistic or actually not treat it as paternalistic at all – since no restriction on liberty is involved. But the thought that ‘paternalism does not require coercion’ does not seem to be a ‘misconception’ about paternalism, but merely a decision about how to define it. Sunstein and Thaler make that decision, in part, to be able to advance the idea of ‘libertarian

paternalism' as a coherent concept. But that does not in itself imply that other definitions which require some form of coercion are 'misconceived'.

The other 'misconception' about paternalism, on their view, is that 'there are viable alternatives to paternalism' (Sunstein and Thaler, 2005, 178). Sunstein and Thaler claim that some form of paternalism is inevitable (Thaler and Sunstein, 2006, 239-253). Given that preferences are endogenous and influenced by the existing state of things and the way in which choices are presented, the relevant state of affairs and presentation are a matter of choice (by the state or private organizations). In as much as they are decision variables, choices which influence our preferences are always in the background. But even if people or organizations make (or have to make) those choices, they do not always do so with a view to improving our lives or welfare. They may do so with a view to pursuing their own interests or profits. Clearly, paternalism is not inevitable. What Sunstein and Thaler show is that the idea of 'given preferences' so commonly adopted in economics is a myth, but that is not a contribution in itself.

What we seem to have learnt so far is that: (1) the anti-paternalist's assumption which Thaler and Sunstein take to be testable and false, is not obviously false since it is not generally true that third parties will, in general, make decisions which will be better for welfare than the people whose lives are involved; and (2) that the two misconceptions about paternalism which they seek to disabuse us of are not obviously misconceptions at all. Paternalism is not inevitable and we may choose to define it in such a way that it requires coercion.

In the light of the earlier discussion of informed desire accounts, a deeper worry about libertarian paternalism is whether it is actually possible or practical, given their definitions. The libertarian paternalist must improve our welfare using the informed preference account of welfare. Yet given that nobody has the relevant preferences, the libertarian paternalist cannot rely on people's actual preferences. So Thaler and Sunstein (2006, 253) recognize that the use of cost-benefit analysis which libertarian paternalists would recommend 'cannot be based on the economists' measure of willingness to pay (WTP).' They add that what is needed is 'a more open ended (and inevitably more subjective) assessment of the welfare consequences.' But there is no reason to suppose that this – if it is a standard form of cost-benefit analysis - will connect with welfare either given their strong definition

of it. So they recognize that '[s]ome readers might think that our reliance on behavior as an indication of welfare is inconsistent with one of our central claims – that choices do not necessarily coincide with welfare' (Thaler and Sunstein, 2006, 254). They try to allay this concern by arguing that when relatively few people choose over time, and on reflection, to opt out of savings plans this means that they believe that the plans benefit them. But, of course, Thaler and Sunstein cannot claim that these agents are perfectly rational. So this attempt to allay the basic concern that cost-benefit analysis will not accurately capture benefits and costs in terms of people's welfare seems to have limited force in the light of their demanding criteria of rationality and welfare.

Thaler and Sunstein also consider two rules of thumb which might be adopted when cost-benefit analysis is not possible. On one, they suggest that 'the libertarian paternalist might select the approach that *the majority would choose if explicit choices were required and revealed*' (Thaler and Sunstein, 2006, 257). Here again they must admit that: '[p]erhaps the majority's choices would be insufficiently informed, or a reflection of bounded rationality or bounded self-control. Perhaps those choices would not, in fact, promote the majority's welfare' (Thaler and Sunstein, 2006, 257). On the second rule of thumb, 'the libertarian paternalist might *force people to make their choices explicit*' (Thaler and Sunstein, 2006, 257). Unsurprisingly again they must admit that '[h]ere too, however, there is a risk that the choices that are actually elicited will be inadequately informed or will not promote welfare' (Thaler and Sunstein, 2006, 258). While they make these qualifications or admissions on a case by case basis while discussing how libertarian paternalism might be implemented, the problem is a deep one and relates in part to the definition of welfare they adopt.

The main point is that if we, being limited human beings, do not know what perfect rationality, on their account, demands, paternalism of the sort they advance is not possible: 'experts' cannot tell people what is good for them, or in their interests or indeed confidently rearrange default options or reframe decisions with a view to making them better off. What seems more plausible in reading their discussion is that they – or those engaged in some of the studies they cite – think they know what is better for us than we do, or better for the subjects of their studies than those subjects themselves do. And, the background assumption must be that those objects which are good for us are not necessarily only knowable to people who are fully rational in having perfect cognitive abilities, self-control and

so on – since none of us are rational in that sense or have perfect cognitive abilities or self-control. Rather they are familiar goods such as health – in the cafeteria and smoking examples - and security in old age - in the case of the savings plan.

The effect of Sunstein and Thaler’s adopting such a strong definition of welfare or rationality is to ensure that a wide variety of violations of expected utility theory *do* count as cases where paternalistic intervention is justified, so that the findings of behavioural economics can be the basis for interventions not dissimilar to those which are advanced by approaches of the sort Sunstein earlier advocated on the grounds that preferences are not necessarily autonomous. But if people are potentially made ‘better off’ simply on the basis that they are not acting according to expected utility theory, and choices may be reframed, or default rules changed, to make them ‘better off’ in some more ordinary sense - which would have to be less ambitious than the informed or rational preference account which they adopt - rather more widespread measures may be needed. The sort of information the social planner (or the various organizations, private and public) would need to bring about these measures would also be potentially beyond any human being whatever account of well-being were adopted. And it is not clear that a world with that range of ‘well-meaning’ interventions would be better – in consequentialist terms - than the one we live in. Even if one thinks that libertarian paternalism is possible, or that its programme could be refined, or redefined, so as to make it possible, one certainly has many grounds to hesitate before pursuing that programme or making reforms on the basis of it. Certainly more thought would need to go into defining the conception of welfare which underlies the proposal.

5. Conclusions

Philosophical accounts of well-being, notably versions of the informed or rational desire (or preference) account, sometimes place the criterion for a desire to be informed or rational at a very high level. Attempts to deal with the adaptation problem within these accounts thus sometimes exacerbate rather than alleviate the concern that these accounts must take into account people’s limitations. The paternalism criticism of informed desire views draws its force from the possibility that these views undermine the authority of actual preferences. Problems of the sort which arise for

the informed desire account also arise for happiness and capability views, though I argue they pose lesser difficulties for some list views (such as Nussbaum's earlier version of the capability view and Griffin's prudential value list view). Sen adopts a democratic, public reasoning approach to addressing ignorance and adaptation in his work on capability, but denies Sugden's claims that his adoption of such an approach would lead to restrictions on liberty. Whatever view one takes of that debate, a line of argument similar to Sen's, involving democratic deliberation, articulated by Sunstein, does imply paternalistic restrictions in some cases, where preferences are endogenous and potentially not autonomous, with a view to promoting autonomy. In such cases, it can be argued that because paternalistic restrictions promote freedom, paternalism is unobjectionable. An extension of this line of thought, in the 'libertarian paternalism' adopted in Sunstein's work with Thaler, sets the standards of rationality and information adopted in their variation on the informed or rational desire account of welfare very high. These standards maximize the range of forms of 'irrational' behavior investigated by behavioural economists which would justify 'libertarian paternalistic' interventions on their approach. But at the same time, they make the libertarian paternalist programme potentially impossible, since it is not clear that any human being would be able to say when welfare would be promoted on the relevant account of rational or informed preference. The programme would need to be refined or redefined to be coherent. At the same time, a number of claims that Sunstein and Thaler make about anti-paternalist positions are not obviously tenable. There are good reasons to resist the libertarian paternalist policy programme in its current form.

References

- Arneson, R 1999 'Human Flourishing versus Desire Satisfaction' E.F. Paul, F. D. Miller Jr and Paul, J (ed) *Human Flourishing* Cambridge: Cambridge University Press.
- Brandt, R 1979 *A Theory of the Right and Good* Oxford: Oxford University Press.
- Brandt, R 1992 *Morality, Utilitarianism and Rights* Cambridge: Cambridge University Press.
- Camerer C , S Issacharoff, G Lowenstein, T O'Donohue and M Rabin 2003 'Regulation for Conservatives: Behavioural Economics and the Case for Asymmetric Paternalism' *University of Pennsylvania Law Review* 151: 1211-1254.

- Crisp, R 1997 *Mill on Utilitarianism* London and New York.
- Dworkin, G 1971 'Paternalism' in Richard A. Wasserstrom *Morality and the Law* Belmont, California: Wadsworth.
- Elster, J 1981 'Sour Grapes – Utilitarianism and the Genesis of Wants' in A Sen and B Williams *Utilitarianism and Beyond* Cambridge: Cambridge University Press.
- Griffin J 1986 *Well-Being: Its Meaning, Measurement and Moral Importance* Oxford: Clarendon Press.
- Griffin, J 1996 *Value Judgement: Improving Our Ethical Beliefs* Oxford: Oxford University Press.
- Harsanyi J 1981 'Morality and the Theory of Rational Behavior' in A Sen and B Williams (eds) *Utilitarianism and Beyond* Cambridge: Cambridge University Press.
- Layard, R 2005 *Happiness: Lessons from a New Science* London: Allen Lane.
- Mill, J S 1962 *Utilitarianism. On Liberty and Essay on Bentham. Together with Selected Writings by Jeremy Bentham and John Austen* ed. M Warnock, Glasgow.
- Nussbaum, M C 1988 'Nature, Function and Capability: Aristotle on Political Distribution', *Oxford Studies in Ancient Philosophy*, 6 Supplementary Vol., 1988: 145-184;
- Nussbaum, M C 1990 'Aristotelian Social Democracy', *Liberalism and the Good*, B. Douglass, G. Mara and H. Richardson (eds.) London: Routledge.
- Nussbaum, M C 1992 'Human Functioning and Social Justice. In Defence of Aristotelian Essentialism', *Political Theory*, 20: 202-246.
- Nussbaum, M C 1995 'Aristotle on Human Nature and the Foundation of Ethics', *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams* J.E.J. Altham and R. Harrison (eds.), Cambridge: Cambridge University Press.
- Nussbaum, M C 1995 'Human Capabilities, Female Human Beings' *Women, Culture and Development: A Study of Human Capabilities* M C Nussbaum and J Glover (eds) Oxford: Oxford University Press.
- Nussbaum, M C 1998 'The Good as Discipline, the Good as Freedom' in *Ethics of Consumption: the Good Life, Justice and Global Stewardship* David A. Crocker and Toby Linden (eds) Rowan and Littlefield.

- Nussbaum, M C 1999 *Sex and Social Justice* Oxford: Oxford University Press.
- Nussbaum, M C 2000 *Women and Human Development* Cambridge: Cambridge University Press.
- Nussbaum, M C 2006 *Frontiers of Justice: Disability, Nationality, Species Membership*, Cambridge, Mass.: The Belknap Press of Harvard University Press.
- Qizilbash, 1998 'The Concept of Well-Being' *Economics and Philosophy* 14(1) 51-73.
- Qizilbash, M 2007 'Social Choice and Individual Capabilities' *Politics, Philosophy and Economics* 169-192.
- Qizilbash, M 2009a The Adaptation Problem, Evolution and Normative Economics in K.Basu and R. Kanbur (eds) *Arguments for a Better World, Vol. 1 Ethics, Welfare and Measurement* Oxford: Oxford University Press.
- Qizilbash, M 2009b, 'The Paternalism Critique of the Capability Approach', Paper Presented at the Conference of the Human Development and Capability Association, Lima, September 2009
- Railton, P, 1986 'Moral Realism' *The Philosophical Review* 95(2): 163-2-7.
- Rawls, J 1993 *Political Liberalism* New York: Columbia University Press.
- Rosati, C 2005 'Persons, Perspectives and Full Information Accounts of Well-Being,' *Ethics* 105(2).
- Sen, A 1980. 'Equality of What?' in S M McMurrin (ed.) *Tanner Lectures in Human Values* Cambridge, UK: Cambridge University Press.
- Sen, A 1993 'Capability and Well-Being' in M C Nussbaum and A Sen (eds) *The Quality of Life* Oxford: Oxford University Press.
- Sen, A 1999 *Development as Freedom* Oxford: Oxford University Press.
- Sen, A 2004 Amartya K Sen, 'Dialogue. Capabilities, Lists and Public Reason: Continuing the Conversation' *Feminist Economics* 10(3); 77-80.
- Sen, A 2005 'Human Rights and Capabilities' *Journal of Human Development* 6,(2): 151-166 .
- Sen, A 2006 'Reason, Freedom and Well-Being,' *Utilitas* 18(1)80-96.
- Sen, A 2009 *The Idea of Justice* Penguin: Allen Lane
- Sidgwick, H 1981 *The Method of Ethics* Indianapolis: Hackett.
- Sobel, D 1995 'Full Information Accounts of Well-Being' *Ethics* 104:784-810.

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 33-51.

Sugden, R 2008 'Capability, Happiness, and Opportunity' in Luigino Bruni, Favio Comim and Maurizio Pugno (eds) *Capabilities and Happiness* Oxford University Press.

Sugden, R forthcoming, 'Opportunity as Mutual Advantage,' *Economics and Philosophy*.

Sumner, L W *Welfare, Happiness and Ethics* Oxford: Clarendon Press, 1996

Sunstein, C 1991 'Preferences and Politics,' *Philosophy and Public Affairs* 20 (1) 3-34.

Sunstein C and Thaler R 2005 'Libertarian Paternalism' in C Sunstein *Laws of Fear: Beyond the Precautionary Principle* Cambridge: Cambridge University Press.

Thaler, R and Sunstein C 2006 'Preferences, Paternalism, and Liberty' *Philosophy* (59) Supplement: 233-264.