

On Choice Inconsistency: The ‘Error’ Error in Behavioural Paternalism

*Adam Oliver*¹

*Department of Social Policy
London School of Economics and Political Science*

Abstract

A core normative assumption of welfare economics is that people ought to maximise utility and, as a corollary of that, they should be consistent in their choices. Behavioural economists have observed that people demonstrate systematic choice inconsistencies, but rather than relaxing the normative assumption of utility maximisation they tend to attribute these behaviours to individual error. I argue in this article that this, in itself, is an error – an ‘error error’. In reality, a planner cannot hope to understand the multifarious desires that drive a person’s choices. Consequently, she is not able to discern which choice in an inconsistent set is erroneous. Moreover, those who are inconsistent may view neither of their choices as erroneous if the context reacts meaningfully with their valuation of outcomes. Others are similarly opposed to planners intervening in the market mechanism to correct for behavioural inconsistencies, and advocate that the free market is the best means by which people can settle on mutually agreeable exchanges. However, I maintain that policy makers have a legitimate role in also enhancing people’s agentic capabilities. The most important way in which to achieve this is to invest in aspects of human capital and to create institutions that are broadly considered foundational for a person’s agency. However, there is also a role for so-called boosts to help to correct basic characterisation errors. I further contend that government regulations against self-interested acts of behavioural-informed manipulation by one party over another are legitimate, to protect the manipulated party from undesired inconsistency in their choices.

Acknowledgements:

I thank Malte Dold, Ralph Hertwig, Paul Lewis, Mario Rizzo and Robert Sugden for their comments on previous drafts of this article.

¹ a.j.oliver@lse.ac.uk

Introduction

In the early part of the 20th Century, following the abandoned project to measure interpersonal cardinal utility that had underpinned 19th Century classical utilitarianism, Paretian welfare economics was born. This form of welfare economics became the orthodoxy. In it, there is no attempt to define the composition of ‘utility’. Orthodox welfare economics merely assumes that an individual’s preference for one good over another is an indication that the preferred good will give her more of her personal conception of utility, which, over all her choices, she will seek to maximise. Although the constitution of utility is not imposed on people, the maximisation principle requires them to be consistent in their choices across contexts. Therefore, implicitly, the normative principle underlying orthodox welfare economics is context independence.

Yet since at least the 1950s, from Allais to Ellsberg to Slovic & Lichtenstein to Kahneman & Tversky to Thaler (to name but few), systematic context-dependent inconsistencies in individual choice have been reported and are the bedrock of behavioural economics. However, most behavioural economists, and all of those who believe that the findings of behavioural economics and psychology justify a paternalistic approach to behavioural public policy, retain the orthodox *normative* postulate of context independence. That is, they believe that people ought to maximise that which gives them utility and that they therefore should be consistent in their choices, even though, in observation, they are often not. In essence, those who align with this mantra believe that context often reacts with choice to produce errors that are, from a policy perspective, in need of correction.

All of those who are constructively engaged in the field of behavioural public policy accept the descriptive validity of many of the context-dependent choice inconsistencies and instabilities, but there is disagreement on the extent to which these are indicative of biases and errors. In this article, in agreement with a substantive claim in Rizzo and Whitman’s 2020 book, *Escaping Paternalism*, I will present a case in support of choice inconsistencies as legitimate aspects of individual choice, which runs counter to the normative underpinnings of orthodox welfare economics and standard behavioural economics. In so doing, I will propose a direction for behavioural public policy that departs from the paternalistic frameworks that have dominated the field to date.

An Overconfident Judgement

As noted above, it would be unusual to find anyone working in the field of behavioural public policy, whether they are of a paternalistic or a more liberal persuasion, who dismisses the notion that people are often inconsistent or, equivalently, incoherent in their choices. However, if a person does choose, say, Good A over Good B in one context, and Good B over Good A in another, how would a third person – including a policy maker – know which choice is the more accurate representation of that person’s preferences (see also Sugden, 2008)? The dominant paternalists in the field – the so-called libertarian paternalists – maintain that we must appeal to a person’s deliberative self, their inner rational agent (or in other words, the criterion of informed desire), rather than their automatic, reflexive choices, to determine what they really want (Thaler and Sunstein, 2008). Smokers, for instance, might say, after a period of reflection, that they really want to abstain from cigarettes, and thus the libertarian paternalist would assume that abstinence reflects their true preferences. However, this is a troubling conjecture, for several interrelated reasons.

For example, people who are asked for their deliberative preferences might simply respond in the way that they believe the interviewer wants to hear – smokers, say, may state on reflection that they wish to quit smoking because they believe that they are being prompted to do so. Moreover, we cannot be sure that the deliberative self is free of the behavioural affects against which the soft paternalists wish to mitigate. For instance, smokers, when asked to reflect on whether they wish to quit smoking, might ignore, or at least downplay, the benefits they reap from smoking. Finally, but most importantly, there is no philosophical or psychological basis on which to argue that an inner rational agent even exists; that the notion of there being an impartial deliberative self that resides within us is a myth (see, for example, Bernheim, 2021; Chater, 2019; Sugden, 2008). Thus, when people offer (deliberative) justifications for why they do, or do not do, something that they claim to prefer, they are perhaps merely offering reasons – motivated reasons – that do not really drive their choices.

To illustrate the contention that one simply cannot discern which, if any, preference is the ‘right’ one when an inconsistency is observed, consider the following prominent inconsistencies in the behavioural literature: (i) classic preference reversals; (ii) actual pension savings levels versus *ex post* ‘desired’ pension savings levels; (iii) willingness to pay (WTP) versus willingness to accept (WTA) discrepancies. Classic preference reversals, most closely associated with Sarah Lichtenstein and Paul Slovic (e.g. Lichtenstein and Slovic, 1973), show that people have a tendency to prefer high probability low outcome bets over low probability high outcome bets in pairwise choice tasks, but equate the latter with relatively high money amounts when asked to separately value them. From these results, we cannot discern whether the preferences indicated in the pairwise choices or the direct valuations are definitively right or wrong; indeed, Lichtenstein, after explaining to respondents with reversed preferences exactly what they had done, found that most of them did not wish to change their answers.ⁱ

The alleged inconsistency regarding pension savings is discussed extensively in behavioural public policy (and elsewhere), and centres on the apparent regret that many of those of pensionable age feel in relation to not saving more for their retirements when they were younger. The inconsistency is thus temporal; i.e. it is between actual savings when young and expressed desired earlier savings when older, and is often attributed to present bias. However, one cannot definitively conclude that the later regret-fuelled preference is the most accurate reflection of an individual’s lifetime desires. The suggestion that the older self does not fully appreciate or remember the desires of the younger self is as plausible as the view that the younger self fails to appreciate or foresee the desires of the older self.ⁱⁱ In terms of an entire life, a third party – and possibly even the first party – cannot deduce whether more or less saving when young would best reflect an individual’s lifetime preference structure (see Oliver, 2023).

The substantial difference that is often observed between a person’s WTP – their buying price – for a particular good and their WTA – their selling price – for that same good is not accounted for in standard economic theory, and is generally attributed to loss aversion. That is, once a person owns a particular good, their reluctance to give it up – to ‘lose’ the good – is not reflected in their *a priori* desire to own the good. This endowment effect was perhaps most famously reported by Kahneman *et al.* (1990), who found that the median respondent selling price for coffee mugs was twice that of buying prices, an implicitly inconsistent preference but again one in which it is impossible to discern which is the more accurate reflection of ‘true’ preference. To contend that one particular preference is more likely to be

the correct one when an inconsistency in choice is observed is a judgement that is often driven by an overconfidence that one's own preferences are shared by others. To steer people in the direction of that preference from a policy perspective (perhaps without the target population's permission or even knowledge), even if one maintains that the intervention is non-coercive, is inevitably paternalistic. Given that preferences are apparently highly malleable, it might be that neither preference is a good reflection of what people really value. Or it may be that both offer a good reflection, given the context and way in which they were uncovered (see also Rizzo and Whitman, 2020, p.75).

Justifiable Inconsistency

In the previous section, I argued that when one observes a choice inconsistency, one cannot generally determine which, if any, of the choices is the 'right' one, which counters the paternalistic urge to steer, or even push, people in particular directions. According to Gigerenzer (2018), the contention that inconsistencies are caused by behavioural biases that are in need of correction has become a truism in much of the behavioural public policy discourse, and yet, he argues, the fast and frugal heuristics that drive much of our behaviour and choices evolved to help us to navigate the uncertainty that is an inevitable feature of almost all of the decisions that we face.ⁱⁱⁱ Gigerenzer thus suggests that to view choice inconsistencies as necessarily caused by biases is in itself a bias; beware, he warns, of the 'bias bias' (see also Gigerenzer, forthcoming).

The 'error error' that forms part of the title of this article is something akin to Gigerenzer's bias bias in that I suggest that it is often erroneous to assume that either of the choices that form an inconsistency are necessarily the result of individual error.^{iv} Rather, it is plausible that we have evolved to be flexible in different contexts because this helps us to achieve our objectives. A person with rigid preferences, living a context-independent life, may, at the very least, suffer a highly regimented, uninteresting existence; at worst, they may find themselves at an evolutionary disadvantage. The argument is thus taken one step further than that outlined in the previous section in that it is postulated here that choice inconsistencies often occur when considerations of outcomes interact with those of context in meaningful ways, such that different values are legitimately placed on the same outcomes in different contexts. In short, the contention is that context dependence is a strength, not a weakness, of the human psyche.

Let us consider this assertion in relation to the Allais paradox, an empirical finding that was foundational to the development of behavioural economics. The paradox shows that people tend to demonstrate inconsistent implied risk attitudes when choosing between two risky options compared to choosing between a risky option and an option where risk is entirely absent (Allais, 1953). In the latter choice, many people place great weight on certainty, resulting in relatively strong apparent risk aversion, an attitude that can in itself be attributed to loss aversion. Although considered a paradox in that it is a direct violation of the independence axiom of von Neumann-Morgenstern expected utility theory and thus descriptively challenges the notion that people maximise their valuation of final outcomes across different choices, one may reasonably ask, what's wrong with placing a heavy weight upon certainty – with being relatively risk averse – when certainty is an option? Indeed, similar to respondents demonstrating classic preference reversals, Slovic and Tversky (1974) observed that Allais paradox-committing respondents generally did not wish to alter their choices after the reasoning of the paradox had been fully explained to them. Certainty brings

the security that one may value in and of itself, because it frees one from the worry and/or realisation of a potential catastrophic outcome of a risky option; when risk is unavoidable, on the other hand, one may quite legitimately choose the riskier but potentially more rewarding option, particularly when the differential risk between the options available is almost indiscernible.

For a simple hypothetical example of context apparently affecting a preference ordering, consider the archetypal example to which advocates of nudge interventions often refer: the differential placement of fruit and cake on cafeteria shelves. If fruit is placed at the front and cake at the back of the counter, so the argument goes, there will be a greater likelihood that people will choose fruit over cake, and vice-versa if the positionings are reversed. Soft paternalists argue that this choice inconsistency is caused by salience; if one were to care about maximising welfare – which, according to nudge advocates if we remember, can be deduced by asking people for their deliberative preferences – then most people will prefer the healthier option, which justifies the placement of fruit rather than cake in the most salient position. That is, it is assumed that, for many people, the choice of cake, when cake is made most salient, is an error in need of correction.

However, as noted earlier, Sugden (2008) and others maintain that the notion that there is a deliberative self is a fallacy, and thus there is no mode of reasoning that the customer could have used to determine whether they prefer fruit or cake. In these terms, the inconsistency is unproblematic because neither choice can be rationalised. However, an alternative reason why the inconsistency is unproblematic is because the different contexts – i.e. fruit or cake at the front – may react with the value that consumers place on either desert in meaningful ways, ways that a third party cannot discern, such that neither choice can be viewed as erroneous. For example, placing an item at the front of a shelf, whether it be fruit or cake, may imbue it with a sense of freshness that the customer values, or may incentivise them to focus on how it complements the rest of their meal, etc. With all these features influencing the customers' choices, we simply cannot say which choice would maximise their welfare, and therefore if planners were to insist that the fruit be placed at the front we can only conclude that this reflects their preferences and not those whom they are aiming to influence.

As an aside, Sugden (2008) also contends that any dessert ordering that is imposed by a planner does not take into account the preferences or interests of the cafeteria owners. It is likely that the primary objective of the cafeteria owners is to make a profit in a business where profit margins are notoriously tight, and so long as they are not engaged in obfuscation or deception, Sugden maintains that they should be allowed to do so without being subjected to interferences by third parties. If it is more profitable for the cafeteria owners to guide people in the direction of cake rather than fruit, perhaps because the cakes they are selling have a shorter shelf life, then in Sugden's schema their placing of cake at the front of the counter is a legitimate business practice. However, one can question whether this practice undermines the notion of a free and fair exchange between the owner and the customer.^v If we, as a society, judge the practice as an unreasonable manipulation of customer buying patterns, then there may be non-paternalistic grounds to regulate against the deliberate salient placement of cake for profit-making purposes (assuming such a regulation is enforceable). This would be a regulation against actions by the cafeteria owners that impose harms on their customers – an externality concern.

Leaving the argument for regulating against externality-generating manipulations to one side (for now), I have so far argued that there are few grounds for correcting choice

inconsistencies by appealing to the possibility that one or the other of the choices best reflects the underlying preferences of those targeted for behaviour change. However, this does not mean that there are no grounds at all. Bernheim (2021), for instance, who, as noted earlier and like Sugden, questions the notion of an inner rational agent, has in a series of articles postulated a model of behavioural welfare economics where some choice inconsistencies can be classified as errors (e.g. see Bernheim and Rangel, 2009). The Bernheim-Rangel framework classifies a choice as a mistake if it is predicated on a clear misunderstanding of the available options and consequences, and thus an unallowable inconsistency arises if such a choice conflicts with at least one choice where no such misunderstanding is evident. For example, if an individual expresses a preference for a box of bananas over a box of apples but was mistaken in believing that the box of bananas contained not bananas but peaches, and yet normally, in the absence of what Bernheim and Rangel call a characterisation error, prefers apples over bananas, a third party may legitimately correct the former choice. If there are no characterisation errors, the Bernheim-Rangel model enforces no consistency requirements and thus the model, in such circumstances, adjudges that each of us is the best arbiter of, and will choose that which is most conducive to, our own wellbeing.

Sugden remains opposed to the notion that policy makers can discern and correct characterisation errors, maintaining instead that they ought to encourage the provision of the most extensive possible range of opportunities from which individuals can choose, and allow individuals free rein to pursue their wants via mutually agreeable exchanges between providers and consumers (Sugden, 2018; see also Sugden, 2021). Bernheim's (2021) view that Sugden's opportunities framework should offer more scope for considering whether people understand the choices that they face is, to me, convincing. With respect to Bernheim's specific contention, I also find the argument for correcting characterisation errors appealing, at least with respect to mistakes that seem very obvious (e.g. when people mistake a box of peaches for a box of bananas), if they can be identified and broadly agreed upon. As another hypothetical example, a UK shoe store orders a collection of shoes that it estimates are suitable for its clientele from an overseas manufacturer, not realising that the manufacturer standardises its shoe sizing in US rather than UK measurements. The shoe store's 'true' preference was for shoes in UK sizing, and thus we can conclude the inconsistency between that preference and the order is a genuine error. Ways in which policy intervention might reduce the possibility of such errors occurring will be discussed later in this article.

Unfortunately, cases in which there is general agreement that inconsistencies are caused by characterisation errors might be the less common occurrence, and those where there is little reason to overturn either choice in an inconsistent set may be rife.^{vi} Take classic preference reversals again, for example. To recap, people have a tendency to prefer high probability bets with modest payoffs over low probability bets with high payoffs in pairwise choice tasks, but place a higher money value on the latter than on the former in independent WTA/WTP tasks. It is possible that people demonstrate this inconsistency because we are ill-adept as humans to process these kinds of decision tasks (Gigerenzer, 2018). If people cannot really understand very well the options that they face we might expect a degree of inconsistency; if so, designing the presentation of the options in ways that people can mentally better process may reduce these inconsistencies. However, if that were the case, the inconsistencies that have been observed are likely to be random rather than systematic in their direction. That is, apparent preference patterns in the direction of choosing the low probability high payoff bet but valuing higher the high probability modest payoff bet would be as prevalent as choosing the latter but valuing higher the former, but this is evidently not the case.^{vii} The systematic

direction of the inconsistency implies intentionality, compounded by the fact that most respondents do not wish to alter their decisions after their choices, valuations and inconsistencies have been carefully explained to them.^{viii}

If we are sceptical that the notion of a ‘deliberative preference’ exists, as Bernheim and Sugden both are, then we ought to be cautious of any rationale for why people demonstrate these quite complex inconsistencies, whether it be offered by a third party or by the respondents themselves. The most likely explanation for classic preference reversals *may* be that people tend to anchor on the payoffs when asked to independently value bets, which would favour the bet with the highest payoff, and yet compare each bet’s various aspects in pairwise choice tasks, which may provoke a greater degree of risk aversion. However, even if this is what is happening there is no obvious error in either type of response, and thus, assuming that the respondents are not being manipulated to serve the interests of another party, there is no reason to intervene. It may be difficult to offer explanations for many types of choice inconsistency; it is often difficult to articulate, or even fully comprehend, why we do the things that we do. We might just choose, act and behave differently across different contexts because that is simply what we want to do, and if there is no obvious error in doing so then there is little justification for direct intervention to counter the observed inconsistencies.

Dealing With Inconsistency

The conclusion reached so far is that intrapersonal inconsistencies in choice across time or context should not be judged as unreasonable, other than when they are caused by clear and obvious mistakes in individual choice.^{ix} The question, then, is what are the lessons for policy?

Sugden (2008), who, as noted earlier, even rejects the concern for characterisation errors, draws on the work of Buchanan (1968), who in turn drew on Wicksell’s (1896/1958) voluntary exchange theory, to reject the notion that an ideal planner can ‘clean’ the choice inconsistencies that are often evident in free exchange. Sugden (2008, p.242) writes that “... it is to each person’s advantage that he is able to get what he wants and is willing to pay for, when he wants it and is willing to pay for it.” That is, we ought to look to forms of exchange that individuals recognise as mutually advantageous, each one on their own terms and given their own individual objectives for their lives. Following this line of argument, there is no need for individual preferences to be fully consistent or coherent, and thus the findings uncovered by behavioural economists and psychologists do not challenge the market mechanism. From a policy perspective, we deal with choice inconsistencies by ignoring them. Therefore, aside from regulating against cases of fraud and obfuscation, Sugden’s principal policy recommendation is to maximise the opportunity set and to let people choose and exchange as they desire (Sugden, 2018).

One way to attempt to deal with inconsistencies caused by the characterisation errors identified as important by Bernheim (2021) is to improve agentic capabilities, an aspect that Dold and Rizzo (2021) and Dold and Lewis (2023) have also, like Bernheim, convincingly (to my mind) identified as an important, missing, part of Sugden’s thesis.^x Agentic capability refers to the ability for people to reflect on their choices in a meaningful way, so that they are more likely to be ‘their own’. There are very general ways in which agency may be enhanced if it is held that there are some ‘goods’ – say, education, health or a basic level of income – that are considered in some sense ‘primary’; primary, that is, if having access to these is

considered foundational to a person's ability to take effective advantage of their own agency. This line of argument offers a classical liberal justification for governments to invest in the provision of (at least) basic universal levels of education, health care and income, mirroring the arguments made by Mill (1859/1969), at least with respect to education, and to some degree Sen (1999) in his capabilities approach.

A more specific set of interventions aimed at identifying characterisation errors fall under the label of boosting, an approach principally associated with Gerd Gigerenzer, Till Grüne-Yanoff and Ralph Hertwig (e.g. see Gigerenzer, forthcoming; Hertwig, 2017; Hertwig and Grüne-Yanoff, 2017). Hertwig (2017, p.143) writes that the “goal of boosts is to make it easier for people to exercise their own agency in making choices. For instance, when people are at risk of making poor health, medical or financial choices, the policy-maker ... can take action to foster or boost individuals' own decision-making competencies.” Boosts can be of many different kinds, but those commonly considered include teaching people simple accounting rules aimed at improving their financial literacy, and instructing people on the implications of false negatives and false positives to give them a better understanding of what proportion of those who test positive for medical conditions actually have the illness under investigation. Visual aids can also be used to perhaps help people better comprehend risky options, as might presenting probabilities as natural frequencies (e.g. ‘1 in 100 people are at risk of...’, rather than ‘there is a 0.01 risk of...’). Moreover, boosts can focus on teaching people about the behavioural phenomena such as so-called present bias, loss aversion and the like, so that they can in principle guard against the impact of these phenomena in their own decision-making if they autonomously wish to do so (e.g. they may feel as though salience is tempting them to eat too many sweets, and might therefore choose to place their store of chocolate at the back rather than the front of their food cupboard, and example of what has been termed self-nudging – see Reijula and Hertwig, 2022).

However, there are several possible issues with boosting. For instance, since they generally – although not always – target individual decision contexts, they might not improve agentic capabilities in the kind of blanket way that is potentially offered by public investment in the basic autonomy-enabling goods and services discussed earlier.^{xi} Boosts are also labour intensive, and rely on the assumption that people are receptive to, and capable of, absorbing the forms of educational training that are embedded within them. One could also make the claim that boosts change the ‘context’ of choice such that decisions are altered in ways that are unnatural when one considers most real-world decision-making scenarios, and yet the real-world scenarios reveal preferences that are not necessarily illegitimate given their framing. If so, boosts may remove a choice inconsistency when, at least in some circumstances, the inconsistency may offer a better reflection of a person's preferences given the choice context faced.^{xii} That said, boosts are not intended specifically to correct choice inconsistencies; the advocates of the approach, including Gigerenzer (2018) as discussed earlier, recognise that such inconsistencies are often a reflection of legitimate preferences. But it may well be that in some circumstances boosts help to correct characterisation errors, thus addressing, at least in part, the concern raised by Bernheim (2021). Thus, the careful use of boosts, as a complement rather than as substitute for the broader institutional interventions that are required to enhance agentic capabilities (e.g. educational and health care services, and a decent minimum income), might be a useful, if marginal, weapon in the policy maker's armoury.^{xiii}

In addition to basic characterisation errors, illegitimate choice inconsistencies can also occur when one party to an exchange manipulates another party through an act of self-interest,

which places a further potential qualifier on Sugden's (2018) opportunities framework. Countless real and hypothetical examples could be given to illustrate this point, but let us imagine that a shopkeeper notices that her customers are more likely to buy a particular brand of discounted biscuits if a selection of the same – but undiscounted – biscuits are placed next to the discounted biscuits on her shelves. That is, a change in the decision context to some extent causes customers who would not otherwise have bought that particular brand of biscuits to buy them because the introduction of the undiscounted biscuits affects people's perceptions of the relative 'value' of discounted ones – an example of what is known as the 'decoy effect'. We may conclude from the previous discussion in this article that there is no obvious error in individual choice when preference inconsistencies are observed via the decoy effect. Context can affect people's preferences for goods in meaningful ways.^{xiv} However, what is different here is that the shopkeeper has deliberately altered her customers' choices through an act of self-interested behavioural-informed manipulation. We can therefore surmise that the behavioural affects are steering choices not in a way that guides customers towards meaningful desires that they wish to pursue, but are instead being deliberately used to cause choices that conflict with their unmanipulated preferences. Due to the very nature of manipulation, customers find it difficult to avoid these influences.^{xv}

One may therefore conclude that such acts of self-serving manipulation cause unwarranted choice inconsistencies, the mitigation of which calls for government intervention. One form of intervention would be to boost in the manner mentioned above – i.e. to educate people about the behavioural affects in the hope that they may then be better able to protect themselves against unwanted interferences. However, even when they are educated as such, boost advocates are perhaps overoptimistic in terms of the capacity and time that people have to notice manipulations that are, by their very definition, well concealed. Thus, the contention here is that there will be occasions when policy makers are justified in openly regulating against particular behavioural manipulations in the exchange relationship. The 'openness' clause here is important, since there must be public scrutiny of the legitimacy of any potential form of regulation so as to abide by the rules of deliberative democracy. Not all manipulations of this type will warrant mitigation and thus there cannot be a 'blanket' policy response to them. They need to be considered by policy makers on a case-by-case basis, who are likely to conclude that some are relatively harmful infringements designed to attract notice of a firm's products in crowded marketplaces, and that profit margins in some sectors are so tight that companies may legitimately entice customers to buy their products with the help of these techniques. Yet, as I have noted elsewhere, there are cases where such manipulations are so egregious that policy makers will do well to conclude that some specific, otherwise egoistically-fuelled, freedoms have to be constrained if one wants to protect freedom in general (Lippmann, 1937/2017; Oliver, 2022). I have in other writings defined government regulations against interferences in a free and fair exchange as 'budges' (e.g. see Oliver, 2013).^{xvi}

Conclusion

Welfare economists, rightly I believe, attach huge importance to individual autonomy, but together with rational choice theorists, they tend to believe that people ought to, and by and large will, choose so as to maximise their own utility. Embedded within these assumptions is that people will be consistent in their choices – that context does not matter. Economics in its dominant form is thus a context independent social science. Behavioural economics and psychologists have observed over many decades that people's choices are often inconsistent,

and systematically so, according to the context in which they make their decisions. That is, descriptively, individual choice is often context dependent. However, rather than questioning the normative assumption of utility maximisation, most behavioural economists have tended to attribute the choice inconsistencies to individual error. This, in itself, is an error: an error error.

This problem has been compounded in recent years, both in relation to the increased prominence given to behavioural paternalistic and subjective welfarist/happiness approaches to public policy. Both of these approaches assume that a benevolent planner can correct or move people's choices in a utility enhancing direction. However, economists themselves have not even reached a consensus on what 'utility' means (e.g. does it mean hedonism, life satisfaction, eudemonia, or anything and everything?), but even if they had, the maximisation of some outcomes-defined concept of utility is unlikely to be the objective to which most people, much of the time, adhere (e.g. see Oliver, 2021, but there is a voluminous and age-old literature on this topic). Rather, most of the time, desires precede any consideration of utility (however utility may be defined) in driving our decisions, and desires vary across people, and – importantly for the theme of this article – within people across context and time.^{xvii} A planner, for the most part, cannot hope to understand the multifarious desires that drive a person's choices. Consequently and generally, the planner is not able to discern which choice(s) is/are an error(s) in an inconsistent choice set. Indeed, the individual who demonstrates an inconsistency may view neither choice as an error if the context reacts meaningfully with their valuation of the outcome. In such circumstances, it is legitimate to value the same outcome differently across different contexts.

I have argued that the policy response to the above is to eschew, for the most part, arguments for a planner to intervene to correct for the inconsistencies in individual choice, a contention that challenges the forms of paternalism that have thus far been dominant in the field of behavioural public policy. Sugden (2018) is similarly opposed to planners intervening to correct for these so-called behavioural irregularities, and maintains that the free market is the best means by which people can settle on mutually agreeable exchanges, irrespective of what their individual objectives might be. His general recommendation is to extend the range of opportunities that people can engage with in these exchanges, so that their own personal desires may be more fully realised.

In line with other scholars, however, I maintain that policy makers have a legitimate – highly important – role to play in enhancing people's agentic capabilities, a line of reasoning that stretches through at least some arms of the classical liberal tradition (e.g. Dold and Rizzo, 2021; Mill, 1859/1969). The most important ways to achieve this is to identify and invest in aspects of human capital that are broadly considered foundational to a person's ability to make fully informed decisions, and yet may not be readily available to everyone through the functioning of the private market (e.g. basic education, health care and income services and levels). However, aside from these and the broad institutional and legal features that protect and nurture individual agency, there is a role for boosting interventions and strategies, which if applied judiciously may have an effective, if relatively marginal, role to play, and may help to correct for the basic characterisation errors identified by Bernheim (2021) as the cause of some illegitimate inconsistencies in choice. Finally, I also contend that government regulation is sometimes warranted, regulation that delegitimises self-interested acts of behavioural-informed manipulation by one party to an exchange over another party, that can cause inconsistent choices and behaviours by the manipulated party. Decisions on when and where it is appropriate to introduce regulations of this kind require, I maintain, discussion via

open public forums on a case-by-case basis, to comply with the rules of deliberative democracy.^{xviii} We should never risk abrogating our own responsibility to hold politicians and policy makers to account.

References

- Allais, M. 1953. Le comportement de l'homme rationnel devant le risqué: critique des postulats et axiomes de l'école Américaine. *Econometrica* 21: 503-546.
- Arkes, H. R., G. Gigerenzer, and R. Hertwig. 2016. How bad is incoherence? *Decision* 3: 20-39.
- Buchanan, J. M. 1968. *The Demand and Supply of Public Goods*. Chicago: Rand McNally.
- Bernheim, B. D. 2021. In defense of behavioral welfare economics. *Journal of Economic Methodology* 28: 385-400.
- Bernheim, B. D., and A. Rangel. 2009. Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics* 124: 51-104.
- Chater, N. 2019. *The Mind is Flat: The Illusion of Mental Depth and the Improvised Mind*. London: Penguin Random House.
- Delmotte, C., and M. Dold. 2022. Dynamic preferences and the behavioral case against sin taxes. *Constitutional Political Economy* 33: 80-99.
- Dold, M. F., and M. J. Rizzo. 2021. The limits of opportunity-only: context-dependence and agency in behavioral welfare economics. *Journal of Economic Methodology* 28: 364-373.
- Dold, M., and P. Lewis. 2022. F. A. Hayek on the political economy of endogenous preferences: An historical overview and contemporary assessment. *Journal of Economic Behavior and Organization* 196: 104-119.
- Dold, M., and P. Lewis. 2023. A neglected topos is behavioural normative economics: the opportunity and process aspect of freedom. *Behavioural Public Policy* 7: 943-953.
- Gigerenzer, G. 2018. The bias bias in behavioral economics. *Review of Behavioral Economics* 5: 303-316.
- Gigerenzer, G. Forthcoming. From bounded rationality to ecological rationality. In: Gigerenzer, G, S. Mousavi, and R. Viale. Eds. *The Herbert Simon Companion*. Cheltenham: Elgar Publishing.
- Grüne-Yanoff, T. 2021. Boosts: a remedy for Rizzo and Whitman's Panglossian fatalism. *Review of Behavioral Economics* 8: 285-303.
- Hayek, F. A. 1961. The non sequitur of the "dependence effect". *Southern Economic Journal* 27: 346-348.
- Hertwig, R. 2017. When to consider boosting: some rules for policy-makers. *Behavioural Public Policy* 1: 143-161.
- Hertwig, R., and T. Grüne-Yanoff. 2017. Nudging and boosting: steering or empowering good decisions. *Perspectives on Psychological Science* 12: 973-986.

- Kahneman, D., J. L. Knetsch, and R. Thaler. 1990. Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy* 98: 1325-1348.
- Lewis, P., and M. Dold. 2020. James Buchanan on the nature of choice: ontology, artifactual man and the constitutional moment in political economy. *Cambridge Journal of Economics* 44: 1159-1179.
- Lichtenstein, S., and P. Slovic. 1973. Response-induced reversals of preference in gambling: an extended replication in Las Vegas. *Journal of Experimental Psychology* 101: 16-20.
- Lippmann, W. 1937/2017. *The Good Society*. Oxon: Routledge.
- Lyons, B., and R. Sugden. Forthcoming. Transactional fairness in consumer markets. *Behavioural Public Policy*.
- Mill, J. S. 1859/1969. On liberty. In: *On Liberty and Utilitarianism*. Oxford: Oxford University Press.
- Oliver, A. 2013. From nudging to budging: using behavioural economics to inform public sector policy. *Journal of Social Policy* 42: 685-700.
- Oliver A. 2021. An unhappy pursuit of happiness. *LSE Public Policy Review* 1: 1-8.
- Oliver, A. 2022. Curtailing freedoms to protect freedom: regulating against behavioural-informed infringements on a fair exchange. *Journal of European Public Policy* 12: 1982-1993.
- Oliver, A. 2023. The authors of our own lives: the limitations of the behavioural justification for paternalism. *Behavioural Public Policy* 7: 924-932.
- Raz, J. 1986. *The Morality of Freedom*. Oxford: Oxford University Press.
- Reijula, S. and R. Hertwig. 2022. Self-nudging and citizen choice architecture. *Behavioural Public Policy* 6: 119-149.
- Rizzo, M. J., and Whitman, G. 2020. *Escaping Paternalism: Rationality, Behavioural Economics, and Public Policy*. Cambridge: Cambridge University Press.
- Sen, A. 1999. *Development as Freedom*. New York: Random House.
- Slovic, P., and A. Tversky. 1974. Who accepts Savage's axiom? *Behavioural Science* 19: 368-373.
- Sugden, R. 2008. Why incoherent preferences do not justify paternalism. *Constitutional Political Economy* 19: 226-248.
- Sugden, R. 2018. *The Community of Advantage: A Behavioural Defence of the Liberal Tradition of Economics*. Oxford: Oxford University Press.

Sugden, R. 2021. A response to six comments on *A Community of Advantage*. *Journal of Economic Methodology* 28: 419-430.

Thaler, R. H., and Sunstein, C. R. 2008. *Nudge: Improving Decisions About Health, Wealth and Happiness*. New Haven: Yale University Press.

Wicksell, K. 1896/1958. A new principle of just taxation. In: Musgrave, R., and A. Peacock. Eds. *Classics in the Theory of Public Finance*. London: Macmillan.

Notes

ⁱ See <https://www.decisionresearch.org/interview-conducted-by-dr-sarah-lichtenstein> This said, in some circumstances – for example, following financial feedback – the rate of preference reversal has been observed to diminish (see Arkes *et al.*, 2016).

ⁱⁱ In discussing the legitimacy of so-called sin taxes, Delmotte and Dold (2022) also question whether the future self is a better judge than the present self of one's current decisions.

ⁱⁱⁱ Dold and Lewis (2022) suggest that Hayek also believed that the use of heuristics is a reasonable context-dependent adaptation that allows individuals to process complex choices; in short, that Hayek saw heuristics as often ecologically rational.

^{iv} Gigerenzer and I arrived at these terms independently of one another. He defines bias bias as the "... tendency to see systematic biases in behavior even when there is only unsystematic error or no verifiable error at all" (Gigerenzer, 2018, p.307). My contention in this article is that systematic choice inconsistencies are genuine and quite common, but even in those circumstances we cannot necessarily attribute them to human error. Since I make arguments in this article that differ a little from his, I will stick to with my error error term here. See also Arkes *et al.* (2016) for arguments against coherence as being the bedrock of rationality.

^v In my schema, a free and fair exchange has been compromised if it is adjudged that one party has used the behavioural affects *in an unacceptable manner* to manipulate another party to act in ways that he would not otherwise have done. Whether or not the use is acceptable would need to be determined on a case-by-case basis through some form of public discussion. It is not for me to specify if any particular use of the behavioural affects is acceptable or not. Sugden, with Lyons, has developed a concept of transactional fairness that forbids deception and hindrance in the exchange relationship – e.g. that forbids a seller from obstructing transactions between his potential customers and other trading partners (see Lyons and Sugden, forthcoming). On that, Sugden and I agree. But Sugden does not disallow the use of psychological cues to attract customers. I maintain that the use of psychological cues for some such purposes *may* be deemed illegitimate, and, as such, are legitimate targets of regulation. That is the difference between Sugden's and my approaches.

^{vi} Sugden (2021) dislikes the tendency to draw general lessons from examples that he considers to be trite – e.g. on a person's misunderstanding of which fruit is contained in a box to justify inconsistency corrections (or the reordering of desserts on counters to justify the nudge approach). I sympathise with his concern.

^{vii} Admittedly, it may, for example, be the case that people understand the pairwise choice task better than they understand the money valuation task; if so, this could cause them to overvalue the low probability high payoff bet in the valuation task. A different – for them, more easily understandable – presentation of the options may quell this overvaluation and thus largely eliminate the inconsistency. But to conclude as such without supporting evidence requires a lot of speculation, and even if evidence were provided that appears to reduce these common inconsistencies, who is to say that the conventional way of valuing the bets is not offering a valid indication of the extent to which individuals value the bets in the context of how the questions are being asked? That is to say, an invalid inconsistency has not necessarily been corrected; a valid one has simply been concealed by an invalid consistency.

^{viii} That said, if it is the case that the preference reversals are caused by the respondents' failure to comprehend the options, and they still cannot understand the options after they have been explained to them, then they might not choose to revise their answers after hearing the explanations. In Lichtenstein's interviews, however, her respondents appeared to understand the options (see endnote i).

^{ix} Of course, this framework allows people to be consistent if that is their desire.

^x Joseph Raz (1986, p.372) maintained that "...the conditions of autonomy are complex and consist of three distinct components: appropriate mental abilities, an adequate range of options, and independence." By "independence", he meant freedom from coercion or manipulation by others. These three components are consistent with the arguments that I make in this article.

^{xi} Advocates of boosts might contest this assertion. For instance, they might contend that improving statistical skills among a population, or educating people about the phenomena uncovered by behavioural economists and psychologists, would enable individuals to make more informed decisions across a wide variety of decisions.

^{xii} It is also possible that boosts that focus people's attention more intensively on particular decision tasks will exhaust their attentiveness and consequently reduce their agentic capabilities over tasks where boosts are not used.

^{xiii} In short, boosts are intended to improve decision-making competences while retaining individual agency. Grüne-Yanoff (2021) argues that boosts are paternalistic in that they diagnose and target individual mistakes, and he disagrees with Rizzo and Whitman's (2020) assertion that boosts are not paternalistic. It is a fine line, but by maintaining agency they strike me as forms of education that people can use or otherwise as they see fit. Boosts do not manipulate or coerce anybody to do anything 'for their own good'. I side with Rizzo and Whitman on this point.

^{xiv} For example, if I, as a guest at someone's house, were offered a choice between a chocolate biscuit and a plain biscuit, I may choose the plain biscuit even though I have a liking for chocolate, because I might wish to leave what I perceive to be the better biscuit for others. However, if I was instead offered a choice between a plain biscuit and two chocolate biscuits, I will likely choose one of the latter. The additional chocolate biscuit is essentially a decoy that could cause an inconsistency in my choices, but since a desire to be perceived as being polite is intrinsic to the first (in leaving the better biscuit) and perhaps the second (in wanting to leave for others a choice between two different types of biscuit) choice in different ways in this example, the context has changed and the inconsistency is perfectly understandable. As stated by Arkes *et al.* (2016, p.28): "If one analyses observable choice only, without a theory of personal values or strategies underlying choice, it is impossible to determine the rationality of the choice."

^{xv} Hayek (1961), to name but one scholar, somewhat optimistically believed that market competition irons out the possibility that any one producer can substantively influence consumer buying patterns.

^{xvi} To me, a free and fair exchange in its purest form is one where behavioural-informed manipulations are absent (as are, of course, acts of deception and hindrance referred to by Lyons and Sugden, forthcoming).

^{xvii} James Buchanan, among others, believed that we, as humans, are constantly striving to become the persons we wish to be and, as such, we are creative in constructing our choices and goals. This inevitably leads to inconsistencies in choice and we do not therefore have stable utility functions, but this, so the argument goes, is what it means to be human (see Lewis and Dold, 2020).

^{xviii} It is acknowledged that it is probable that public forums will be heavily impacted by behavioural affects, such as herd effects, group think, the confirmation heuristic and the like. It is beyond the scope of this article to consider whether and how these affects ought to be ameliorated in such discussions.