

PAPERS on Economics & Evolution



MAX-PLANCK-GESELLSCHAFT

1208

Opportunity and Preference Learning

by

Christian Schubert

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

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

© by the author

Opportunity and Preference Learning

Christian Schubert¹
Max Planck Institute of Economics
Kahlaische Str. 10, 07745 Jena, Germany
schubert@econ.mpg.de
phone: +49-3641-686 810
fax: +49-3641-686 868

Abstract

Robert Sugden has recently elaborated upon the case for a normative standard of freedom as “opportunity” that is supposed to cope with the problem of how to realign normative economics – with its traditional rational choice orientation – with behavioral economics. His standard, though, presupposes that people respond to uncertainty about their own future preferences by dismissing any kind of self-commitment. We argue that the approach lacks psychological substance: Sugden’s normative benchmark – the “responsible person” – is a purely artificial construct that can hardly serve as a convincing role model in a contractarian setting. An alternative concept is introduced, and some policy implications are briefly discussed.

Keywords: Opportunity Criterion, Preference Change, Reconciliation Problem

JEL: D51, D63

¹ For helpful comments and discussions I would like to thank Martin Binder, Leonhard Lades, Maurizio Pugno, and Ulrich Witt. All remaining errors and omissions are mine.

1. INTRODUCTION

In a series of papers, Robert Sugden has recently elaborated upon the case for a criterion of freedom as “opportunity” that is supposed to cope with the problem of how to realign normative economics – with its traditional grounding in orthodox rational choice methodology – with behavioral economics (Sugden 2004a; 2006b; 2007; 2008a; 2010; McQuillin and Sugden 2012a).² Among the various attempts that have been made in the literature to come to terms with this “reconciliation problem” (McQuillin and Sugden 2012b) and to explore the normative implications of behavioral economics – such as “libertarian paternalism”³, “behavioral welfare economics”⁴, and “happiness politics”⁵ –, his stands out in two respects.⁶

First, Sugden wishes to refute the popular claim that acceptance of behavioral economics insights makes anti-paternalism unattractive or even “incoherent”.⁷ His point of departure is the failure of the standard welfarist or well-being approach to normative economics to work properly, once it is clear that individuals often do not reveal coherent (i.e. consistent, stable, and context-independent) preference orderings through their observable choices. The opportunity criterion that he proposes admits the possibility of incoherent preferences, while retaining the substance of liberal welfare economics – i.e., the principle of consumer sovereignty. His underlying normative intuition is that “it is good that each person is free to get what she wants”, provided she does not harm others (Sugden 2004a: 1016). As McQuillin and Sugden (2012b: 563) put it, “[v]alue is assigned to an individual’s being free to choose from a wide range of options, independently of whether her choices reveal coherent preferences.” The microfoundation is provided by the concept of a “responsible” (rather than “rational”) person that fully endorses the decisions made by all her – past and future – selves (Sugden, 2004a: 1018). Sugden characterizes the responsible agent as follows: “She values opportunity for herself because (looking ahead to future decision problems) she wants it to be the case that she will then

² See also Sugden (1998a; 1998b; 2003; 2004b; 2006a; 2008b; 2009, 2011). In general, there is a growing emphasis on opportunity (rather than preference satisfaction) as a “currency of welfare” among normative economists, see, e.g., Sen (1992), Arrow (1995) and Roemer (1998).

³ See, e.g., Thaler and Sunstein (2008), Camerer et al. (2003).

⁴ See, e.g., Bernheim and Rangel (2009).

⁵ See, e.g., Schubert (2012).

⁶ The capability approach (e.g. Sen 1988) might also be interpreted as an attempt to tackle the reconciliation problem, but since it is rarely explicitly justified on this ground, we abstract from it here.

⁷ See, e.g., Sunstein and Thaler (2003: 1164-65, 1182).

be able to do as she then desires” (Sugden 2010: 55). According to Sugden, such a person can value opportunity even if her preferences are incoherent (ibid.).

Second, and apart from the substantive suggestion to adopt the criterion of opportunity, Sugden proposes to redirect normative economics more generally. Rather than seeing themselves in the role of policy advisors that assess the value of social states from some allegedly neutral and encompassing (“synoptic”) perspective, normative economists should take a rule-based approach, addressing the citizens affected by their advice directly: The task would then be to help negotiating “a fair agreement between individuals, each of whom is looking at the world from his own viewpoint” (Sugden 2008a: 229-230).⁸ His concept of opportunity is, thus, taken to be offered to the citizens themselves as a currency of advantage in which to assess concrete institutions and policies.

As we will argue in this paper, despite all its merits, Sugden’s concept of opportunity suffers from being based upon an incomplete model of the attitude real-world individuals are likely to assume toward the fact that their own preferences will change in future periods (thereby most probably violating coherence axioms). As a consequence, this concept will most likely fail to convince his addressees of the claim that it provides a possible basis for a “fair agreement”. Specifically, we argue that Sugden overlooks the potentially wide range of attitudes underlying behaviors with which people actively strive to achieve *subjective* coherence among their preferences, by managing the way they learn new tastes in the future. Sugden narrows this range of attitudes down to just one: Unconditional endorsement. Crucially, he denies the practical relevance of personal precommitment: “Most of us, most of the time, recognise that the best way to cope with ... uncertainty [about one’s future preferences] is to avoid making unnecessary commitments about future consumption” (Sugden 2008a: 243). Only by identifying with (and privileging as authoritative) the “acting” or “impulsive self” (in his role as consumer, say), rather than the self as “maker of plans or source of reflective judgement about the well-being of the continuing person” might one be able to appreciate the value of the market in providing for opportunity (ibid.). This view implicitly assumes that the mere absence of interferences (i.e., negative liberty) is sufficient to give rise to the quality of agency necessary for the normative role model of the “responsible” person to be empirically plausible. This, however, is at odds with basic insights from behavioral economics.

⁸ His general contractarian approach is elaborated upon in Sugden (1989). See also Sugden (2011).

The argument proceeds as follows. Section 2 presents, in a nutshell, Sugden's concept of opportunity, as he has developed it so far. Section 3 discusses what are, in our view, its main shortcomings. Section 4 then introduces an alternative notion of "opportunity to learn" and discusses some policy implications with respect to the use of "nudges". Section 5 concludes.

2. SUGDEN'S CONCEPT OF OPPORTUNITY

Sugden aims at developing a "standard of value" – to be used in assessing alternative institutional arrangements – that does not require individuals to hold stable and coherent (or "considered") preferences. It is supposed to allow for preferences to be formed in the very process of interaction on markets for private goods (Sugden 2008a: 230).⁹ Hence, such a standard would be applicable in the world as described by behavioral economics. At the same time, Sugden wishes to preserve, as much as possible, the liberal emphasis on personal freedom – epitomized in the principle of consumers' sovereignty¹⁰ – and the particular (Millian) focus on individual self-development. Specifically, he claims that although individual preferences that are being satisfied may be unconsidered, and, therefore, not suitable as a basis for an account of subjective value, people's "valuing the *opportunity to satisfy them* is considered" (Sugden 2006b: 217).¹¹ In other words, while the opportunity to satisfy preferences may be *used* in an unconsidered way (e.g., leading to dynamic inconsistencies), the valuation of this opportunity itself can be assumed to be perfectly coherent and stable. Thus, Sugden suggests substituting an opportunity criterion for the traditional preference satisfaction standard: Opportunities, rather than outcomes or consumption bundles, are the ultimate carriers of value. Given his contractarian framework, he has to show (i) that his claim is "credible as moral psychology" and (ii) that the corresponding standard of value can be formalized in a coherent way (Sugden 2006b: 218). In elaborating upon his approach, Sugden has so far focused on question (ii), answering question (i), affirmatively, basically by introspection.¹²

⁹ He applies his criterion to the provision of public goods in Sugden (2009).

¹⁰ The defense of this principle by Sugden is somewhat puzzling, given that it tends to be firmly embedded in a welfarist framework of preference-satisfaction that Sugden actually rejects. We submit that what he really means to defend is the principle of *freedom of choice*. Both principles can have conflicting implications (Rothenberg 1962).

¹¹ Italics partly omitted. Generally, in social choice theory, opportunity for a given individual is defined as "something that he has the power to bring about, if he so chooses" (Sugden 2010: 49).

¹² See, e.g., Sugden (2006b: 217), Sugden (2007: 671), Sugden (2008a: 243, 247), McQuillin and Sugden (2012a).

Sugden demonstrates that competitive markets respond in a coherent way to the fact that people's preferences are incoherent, thereby achieving what may be interpreted as the satisfaction of the opportunity criterion.¹³ By providing maximal opportunities for mutually beneficial transactions (including intertemporal transactions between an agent's selves), markets give each person, "rational or irrational, what she wants and is willing to pay for, when she wants it and is willing to pay for it" (McQuillin and Sugden 2012a: 630-31). In a competitive market setting with free trade, each person's opportunities are a function of other people's desires – which implies that the opportunities of all persons are potentially connected in a network of positive-sum interactions (Sugden 2010: 55-62). Sugden states that "[e]ven though we do not know what our preferences will be next year, we know that market agents will be rewarded for having anticipated those preferences and having found ways of satisfying them" (Sugden 2008a: 244). He even extends this argument to the case of preferences *created* by market participants through, e.g., advertisement (ibid.: 244-247). Hence, Sugden specifies his concept as "opportunity as mutual advantage" (Sugden 2010).¹⁴ By respecting an agent's preferences "at the point of consumption", the market provides a maximum of opportunities for mutually beneficial transactions (Sugden 2009: 100).

He describes his methodological background assumptions most lucidly in Sugden (1998b). There, he explicitly adopts a position that he attributes to Hume and Gauthier (and indirectly to Pareto), according to which an individual's actual preferences are not the proper subject of rational assessment in light of reason. Preferences should rather be understood as "passions". The underlying subjective beliefs may well be flawed, but the preferences themselves can never be "irrational".¹⁵ This implies the rejection of any kind of substantive theory of well-being (or simply "of the good"). In particular, a person's preference for A over B should not be interpreted as some kind of overall all-things-considered *reason* for choosing A over B; it does not have to be justified to anyone else – particularly not to an ethical observer.

¹³ In more abstract terms, markets can "integrate fragments of preference that are revealed in different situation and at different times", just as they can integrate fragments of dispersed knowledge (Sugden 2008a: 229, 239). As he shows in (Sugden 2004a), in a simple model of an exchange economy, the opportunity criterion is satisfied if the law of one price holds and if all markets clear. The model presented there is extended in McQuillin and Sugden (2012).

¹⁴ This reflects Hayek's – epistemologically grounded - insight that "it is because every individual knows so little ... that we trust the independent and competitive efforts of many to induce the emergence of what we shall want when we see it" (Hayek 1989: 55-56).

¹⁵ Recall Hume's dictum that "reason is the slave of the passions" (Hume 1740/1978).

From this point of view, it is easy to dismiss the welfarist approach to well-being as incoherent: Even if preferences were to turn out as stable and consistent as postulated in the textbook models, there is no ground for assuming that a person's good always equals the satisfaction of her preferences (Sugden 1998b: 41). The ethical observer's judgments about what is good for a person are based on reasons, while the person's preferences are not. Sugden aims at avoiding this inconsistency, while keeping the welfarist neutrality with respect to any particular account of human well-being. He endorses Harsanyi's principle that "we should work with each individual's *own* conception of what is good for her" (Sugden 1998b: 46).¹⁶ In order to do so, he adopts a contractarian approach.

While eschewing any substantive theory of well-being (such as happiness, or capability), Sugden nonetheless needs a "conception of each person's interests – of what each person is seeking to achieve from social cooperation" (Sugden 1998b: 60). Here, he introduces his critical premise, viz., that each person seeks to maximize the extent of her opportunities. The claim is that this offers a "reasonably realistic model of people's judgments about opportunity" (Sugden 1998b: 55). Given this assumption about people's "pervasive" preference for increases in opportunity, he then classifies this preference as a passion, which therefore does not require any justification in terms of a particular theory of well-being – it does not make sense to ask *why* someone benefits from increases in her opportunity set (Sugden 1998b: 53). Endowed with this specification of individual interest, he can hypothesize about what kind of institutional arrangements each individual, taken separately, can agree to.

Hence, it seems that the only way to engage in a critical scrutiny of Sugden's account is in probing his fundamental empirical assumption that there is indeed, among real-world individuals, a widely held "passion" for increasing opportunity sets. For this assumption backs his specification of what is in the interests of agents (as perceived by themselves) in a contractarian setting.¹⁷ Indirectly, the "judgment about opportunity" that he stipulates reflects a person's attitude toward her own future preferences. In order to make sense of that, he suggests a specific conception of personhood or identity that also serves as a normative role model: For Sugden, identity consists "not in a coherent system of preferences or in a rational plan of life, but in an *attitude towards one's past, present and future actions* which acknowledges those actions

¹⁶ Italics in the original.

¹⁷ On this, see also Sugden (2008b: 320).

as one's own and accepts responsibility for their consequences" (Sugden 2010: 54-55).¹⁸ His normative ideal is the *responsible*, rather than the rational individual. A responsible person, in Sugden's view, "treats her past actions as her own, whether or not they were what she now desires them to have been. Similarly, she treats her future actions as her own, even if she does not yet know what they will be, and whether or not she expects them to be what she now desires them to be" (Sugden 2004a: 1018). Couched in the multiple-selves language, a responsible agent is a continuing agent, i.e., a "composition of the series of time-slice agents" (Sugden 2007: 671).¹⁹ Individuals acting on incoherent and unstable preferences are, then, not treated as consisting of multiple selves with conflicting desires (of whom one ought to be privileged), but as continuing loci of responsibility. They endorse any desire their selves once had or will have – they even *identify* with their separate selves (Sugden 2007: 672, FN 5).²⁰

An individual that is responsible in this specific sense will certainly value any increase in her lifetime opportunity as unambiguously good for herself (Sugden 2004a: 1018). Importantly, this implies that the "responsible" person will reject any kind of self-regulation or self-commitment.²¹ Sugden has formalized this intuition by using the concept of "nested opportunity sets" to model different multi-period decision problems for an agent (Sugden 2006b; 2007). Conceptualizing each person's opportunity as the "set of allowable lifetime behaviours" (McQuillin and Sugden 2012a: 628), he shows that, given the inclination to identify with one's own selves, more opportunity is unambiguously preferable to less (even if the extra opportunity turns out to generate unambiguous losses in monetary terms), and tools to reduce opportunities in future periods have zero value. An essential implication is that the tools suggested by "libertarian paternalism" potentially impose normative costs – largely overlooked in the literature – by curtailing the valued range of opportunities. Specifically, Sugden wants to refute the claim, advanced by Thaler and Sunstein (2003: 1164-65, 1182), that accepting the insights of behavioral economics makes anti-paternalism "incoherent".²²

¹⁸ Italics added.

¹⁹ Italics omitted.

²⁰ At first sight, there is a striking resemblance of this role model with the aristocratic attitude of "never complain, never explain", usually attributed to Benjamin Disraeli – except that that attitude would presuppose a high degree of personal self-regulation, excluded by Sugden (see below).

²¹ Such a person "will not want to impose external constraints on her future choices as a way of forcing those choices to match her current conception of what is good for her" (Sugden 2004a: 1018).

²² See Sugden (2008a: 229), McQuillin and Sugden (2012). Sunstein and Thaler (2006: 237, FN 1), e.g., advance the strong claim that "the autonomy argument [meaning the "belief that people are entitled to make their own choices even if they err"] is undermined by the fact ... that sometimes preferences and choices are a function of given

Taking opportunity as the standard of value prompts the question of how to measure it (in order to determine a “fair” distribution, say).²³ Sugden opposes the approach, dominant in the literature, to do it by taking a person’s preferences as given.²⁴ Rather, “we should take a perspective in which preferences are unknown or undecided. We should ask questions of the form: *Were the individual to have* such-and-such a preference, would she be able to satisfy it?” (Sugden 2010: 50, italics in the original). In other words, the metric of opportunity should not, in his view, be based on a person’s actual preferences, but rather on her *potential* (counterfactual) preferences, i.e., those that she “might adopt”, given her “objective circumstances”, but assuming that her preferences are unformed (Sugden 2003: 791).²⁵ Without the assumption of given preferences, and eschewing the approaches, common in the literature, to rely on “reasonable” or “typical” preferences, though, it is impossible to measure a person’s opportunity directly. In Sugden’s view, opportunity rather becomes an “open-ended” concept (Sugden 2010: 48) – which it must be, in order to capture the basic intuition informing Mill’s ideas on personal liberty and self-development (Sugden 2003: 794).²⁶ In order to be nevertheless able to use the concept in normative discourse in a meaningful way, he suggests to approximate the benefits one can derive from a given set of opportunities as a function of the “resources or real income” an agent possesses (Sugden 2010: 49). Specifically, distributional analysis can only proceed by measuring *entitlements* (ibid.: 62-63). Then, the notion of opportunity requires neither the assumption of rationality nor any particular philosophical conception of the good life. One can make judgments about the relative value of alternative opportunities without knowing anything about the substance (and, a fortiori, the coherence) of one’s own future preferences.

arrangements.” From observing that preferences do not “predate social contexts” they conclude that there are no “viable alternatives to paternalism” (ibid.: 235; see also ibid.: 250).

²³ Sugden (2004b) elaborates upon a notion of “equality of opportunity” that does not rely on information which can only be known after the market game has started. This implies, inter alia, that equal efforts do not necessarily yield equal rewards.

²⁴ He also rejects a “pure quantity”-view of opportunities that measures their value independently of an agent’s preferences, actual or potential.

²⁵ Sugden defines preferences as “dispositions to make specific choices ... the product of a process of deliberation in which the agent decides what to choose” (Sugden 1998a: 323).

²⁶ As Sugden (2003: 809) explains, the value of opportunity derives from effects that are hard to predict. Recognizing this allows us to acknowledge the value of opportunity for people “to act contrary to their current preferences, contrary to the preferences that are typical for people like themselves, contrary to accepted principles of rationality and morality, and contrary to moral philosophers’ favored accounts of the human good.” His intellectual debt to Mill is outlined, e.g., in Sugden (2010: 58-61) and Sugden (2003: 785-788).

Sugden has applied his concept of opportunity to Amartya Sen's well-known "adaptation problem", as specifically described by Mill for the case of subjugated women.²⁷ Sugden endorses Mill's suggestion to simply give women – who honestly claim to be satisfied in their role as dominated housewives – the formal freedom to pursue those activities they deem worthwhile, and then to find out – in an experimental fashion – which of them suit them best. Put differently, policy should make sure that women get the formal opportunity to achieve whatever they might desire (Sugden 2006: 45). This approach neither requires substantive information about how malleable female preferences are, nor any public judgment about what constitutes human flourishing, or on which desires count as "reasonable" or "typical" in a given context.

What kind of freedom does the notion of "opportunity as mutual advantage" represent? Perhaps surprisingly, Sugden claims that he is defending a concept of positive or "effective" freedom, i.e., freedom *to do or be something* (e.g. Sugden 1998b: 59). He argues that the opportunity criterion goes beyond negative freedom, for the mere absence of restrictions is not sufficient to establish the system of free exchange that is needed in order for it to be satisfied (Sugden 2010: 58).²⁸

All in all, Sugden appears to pursue three (interrelated) objectives:

- (1) He wants to show that it is possible to "sell" the opportunity concept to people as participants of a constitutional assembly, i.e., to show every one of them separately that he benefits from adopting this concept as a standard of value.
- (2) He wants to demonstrate, *contra* Thaler and Sunstein (2008), that acceptance of key behavioral economics insights does not make anti-paternalism with respect to private and public goods "incoherent".
- (3) He wants to show how to talk meaningfully about opportunity and its fair distribution, even when this concept cannot be measured directly.

The present paper is concerned with the first point only. If we can demonstrate that Sugden's approach rests on an inadequate view of human behavior, this decreases the likelihood that the general public will accept it as informing constitutional advice. This point is crucial for his argument, for it would question the empirical relevance of points (2) and (3), although not their internal coherence.

²⁷ See Qizilbash (2011a: 40-44) on Sugden's critique of the way Sen deals with this problem.

²⁸ Qizilbash (2011a: 37), though, interprets Sugden as endorsing – like Mill – a concept of negative liberty.

We will now examine what is wrong with this account of opportunity, and suggest an alternative concept.

3. TWO SHORTCOMINGS OF THE OPPORTUNITY CONCEPT

In what follows, we will challenge Sugden's claim (let's call it "claim I") that the standard of value which he proposes represents a "reasonably realistic model of people's judgments about opportunity" (Sugden 1998b: 55) and, hence, builds upon "credible moral psychology" (Sugden 2006b: 218). His claim II, concerning the possibility to formulate the standard of value (viz., the opportunity criterion) and its implied anti-paternalism in an internally coherent way has been convincingly established – its practical relevance, however, depends on the validity of claim I.

Let us be clear at the outset that we join Sugden in adopting a contractarian perspective that refrains from addressing some allegedly neutral ethical observer; we also concur with him in rejecting the welfarist attitude that tries to safeguard objective preference coherence by discounting, from some external viewpoint, any deviations from rational choice axioms as "errors" or "mistakes" in need of correction.

We however argue, first, that Sugden's characterization of his normative role model as a "responsible" person reflects a counterintuitive view of what responsibility implies. Closely related to this, Sugden neglects the way agents come to acquire the preferences they subsequently wish to have the opportunity to satisfy. In short, there is no theory of learning in his account. Our second claim is that Sugden's approach is inadequate in the way it specifies the attitude that his "continuing agent" is likely to assume toward the fact that her own preferences will change in future periods. We suggest to take a *motivation-theoretic* perspective to clarify this point. It is unlikely that most agents will adopt the kind of attitude that Sugden attributes to his normative role model, namely, the unconditional endorsement of (and even *identification* with) any preference she may have in the future, which involves the rejection of any kind of self-regulation. We submit instead that most people have an inherent tendency to strive for subjective coherence, and that in order to achieve this, they require techniques of self-regulation that the market may or may not provide.

3.1 “Responsible” behavior toward one’s future preferences

As we have seen, Sugden’s normative behavioral benchmark is the “responsible” person. He defines as “responsible” someone whose continuing agency is expressed by her unconditional endorsement of any choice or preference she might have in future time periods. Such a person is prepared to make her own mistakes and “live with the consequences” (Sugden 2007: 680).²⁹ She endorses, as a *passion*, the desire to have the chance to cater to all the preferences she might have. This person would, then, see opportunities to satisfy potential preferences as inherently valuable. By implication, this means that a person would be labeled “irresponsible” by virtue of him valuing the option to engage in self-commitment or self-regulation, thereby endorsing the preferences of his multiple selves in a selective manner. Knowing about the habit-forming and generally path-dependent nature of many preferences, he would presumably do this with the aim to avoid harming himself through future choices that may turn out to be detrimental to his health, say. This kind of concern, though, hardly qualifies as “irresponsible”. Rather, it seems to be the lack of it that typically makes children act “irresponsibly”.

Moreover, granting that a passion for maximum opportunity does not need to be justified, we can nonetheless question its empirical plausibility by asking what could possibly motivate such an attitude under real-world conditions. It seems that to assume “responsibility” for some attitude A (in the sense of treating the corresponding choice as her own), an agent B would at least need to discriminate among the various processes by which A may come about, and then to see to what degree these processes are under her control. In other words, some degree of personal autonomy would be required in order to qualify A as sufficiently *authentic* to make its full endorsement plausible. Autonomy, however, implies “self-regulation”. In Mill’s account, it crucially also involves dynamic “self-development”, which presupposes an individual’s opportunity to try out new experiences, make mistakes, learn from them, and develop personal capacities for judgment and decision-making.³⁰ While Sugden attributes to the “responsible” person the willingness to make her own mistakes and live with the consequences (see above), he is silent about whether to “live with the consequences” involves any kind of *learning*. Tellingly, in the model of opportunity sets he offers in Sugden (2006b; 2007), his agent’s (risk) preferences

²⁹ See also Sugden (2010: 55).

³⁰ See Christman (1991) for a similar approach that makes a person’s autonomy a function of her learning history. On Sugden’s intellectual debt to Mill, see the references given in footnote 25, above.

are assumed to be mood-dependent (Sugden 2007: 667), i.e., they switch from one state to the next in a systematic, though non-cumulative way.³¹ In other words, they merely reflect the instability aspect of preference change. We argue that this only captures a part of what real-world people experience when faced with changing preferences. In order to be both in line with Mill's account – i.e. to make the opportunity concept express the idea of self-development –and to offer a credible notion of what “responsible” behavior implies, Sugden must implicitly assume that his agents are actively engaged in learning in the sense of irreversible preference change (Elster 1982). But neither he nor his “responsible” person are concerned about the (possibly dysfunctional) way in which the learning of new preferences proceeds.

3.2 Striving for subjective coherence

Our second claim is that real-world individuals generally strive for a minimum degree of subjectively perceived coherence between their personal beliefs and preferences in a synchronic and diachronic sense. In order to achieve this, they require tools of self-regulation and self-commitment. This lends further substance to the claim that unconditional endorsement cannot qualify as a plausible attitude toward one's own future preferences.

Notice that “subjectively perceived coherence” differs from what is usually referred to as “internal consistency” in the rational choice paradigm. What individuals perceive as coherent may be, from an observer's viewpoint, totally “irrational” and unstable. We argue that the personal striving for subjective coherence represents a fundamental human drive that is expressed in behavioral dispositions to reduce cognitive dissonance (Akerlof and Dickens 1982)³², to preserve one's preferences over time (Sally 2000), to consolidate one's preferences once they have been constructed (Hoeffler and Ariely 1999), to maintain a positive self-image, and to please others. Typically “willpower” has to be mustered to achieve this (Loewenstein 1999). While in the literature, the demand for self-commitment devices is usually seen as being driven by the desire to maintain objective “internal” consistency, i.e., rationality in the

³¹ For evidence, he refers to Isen (1999). Note also the emphasis Sugden (1998b: 42-43) puts on the *stochastic* nature of preferences.

³² The psychological theory of cognitive dissonance describes the human disposition to adapt one's *beliefs* to one's changing preferences. Since Sugden assumes that people are “responsible” for their own preferences, he must postulate that they are also able to control their preferences. This may be preferable than adjusting one's beliefs when, e.g., there is a strong desire for certain beliefs (e.g. Caplan 2001).

neoclassical sense,³³ we argue that it should rather be understood as expressing the desire to maintain a minimum degree of coherence, *as perceived by the individual herself*.

While a desire to establish and maintain the conditions of “internal consistency”, as described in the economics textbook, can hardly be given plausible motivational underpinnings (for instance, it is questionable whether it yields higher payoffs than “irrational” play in strategic interactions; see Berg and Gigerenzer 2007),³⁴ the human drive to establish and maintain subjectively perceived coherence is confirmed by psychological evidence. For instance, Ryan (1995: 398) argues that “innate tendencies toward assimilation and integration play a critical role in social development”. In most domains of life, humans exhibit an inherent disposition to “actively grow, learn, and integrate” (ibid.). Paradigmatic examples include intrinsically motivated behavior and the internalization or rather “active assimilation” of behavioral regulations that are “originally alien or external to the self” (ibid.: 405-408). An important subtype of internalization, relevant in the context of the present discussion, is “emotional integration”: “[P]eople have to learn to manage the expression of drives and affects as well as conflicts between them, and do so in more or less integrated ways” (ibid.: 409). At first sight, viewing the human personality as a “self-unifying system” (ibid.) seems to clash with the evidence of inconsistency in human behavior. But it may be reconciled with this evidence by arguing that a person’s various “selves” are subject to comparative evaluation in the sense that they are experienced as more or less “authentic” from the point of view of the “true” self (a notion that itself can, of course, fluctuate over time).

As Ryan explains, integrative processes are dynamic and ongoing and therefore “highly dependent upon contextual supports for basic psychological needs” (ibid.: 399). While the insights of the “self-determination theory” proposed by Ryan and others support the “opportunity” view of human well-being in stressing the importance of “opportunities for choice” and the “absence of salient external controls” (Ryan 1995: 404), they also tend to emphasize that the basic needs for autonomy, competence and relatedness need to be supported by appropriate institutional conditions that go beyond the simple removal of constraints (“ambient supports” in the words of Ryan et al. 2008). Put differently, autonomy is not just

³³ The literature is huge, see, e.g., Schelling (1984), O’Donoghue and Rabin (1999).

³⁴ See, however, Yaari (1977) for the possibility that an agent’s endogenously changing preferences may be exploited to the detriment of the agent, even when she is fully informed.

“given”; it is a capacity that “must be developed through various processes involving educational, social, and personal resources” (Christman 2005: 87, italics omitted).

A universal desire for “self-command” in order to maintain subjectively perceived coherence has also been postulated by Adam Smith in his *Theory of Moral Sentiments* (Smith 1976). In order to receive pleasure from the approval by others, people use the “impartial spectator” (Smith’s “man within the breast”) to align their preferences with certain social norms that prevail in their community. Personal goal formation is shaped by our sense of what is worthwhile, and this in turn is influenced by subjective perceptions of social approval. Sugden, in his own interpretation of Smith, speaks of “sentiments” or “sentimental repertoires” (Sugden 2002: 77). In Smith’s account of fellow-feeling, the impartial spectator manages one’s preferences, not with the aim to make them cohere with rational choice axioms,³⁵ but rather to gain benefits through the approval of one’s fellows. In turn, this approval may induce a perception of relatedness, which furthers the satisfaction of fundamental needs necessary for autonomous action (Ryan 1995).

Notice that a substantive account of human well-being is introduced here. In order to formulate a viable critique of Sugden’s approach “from within” his own paradigm, we cannot muster such an account to scrutinize the *reasons* people may or may not have for holding a preference for increases in opportunity. In Sugden’s view, if people have such a preference – understood as a pre-rational passion – it just reflects their values and cannot be subject to reason. The account given here can, however, be used to challenge the *empirical* claim that (most) people (apart from Sugden himself) do hold such a passion.

Thus, the “responsible” agent proposed by Sugden does not represent “credible moral psychology”. Hence, it lacks descriptive adequacy in the sense of Sumner (1996), and cannot serve as a convincing normative role model that guides contractarian advice.

4. TOWARD A CONCEPT OF “OPPORTUNITY TO LEARN”

In light of these objections, we suggest to replace Sugden’s concept of “opportunity” by a more dynamic concept of “opportunity to learn”. We follow Sugden in adopting a contractarian perspective and in rejecting the welfarist attitude that tries to safeguard preference consistency

³⁵ That would be the viewpoint of, e.g., Khalil (2010).

by discounting, from some external viewpoint, deviations from rationality standards as “errors” or “mistakes”.

We however depart from his concept in arguing that the notion of a “responsible” person is empirically implausible. It is psychologically more realistic to assume that real-world individuals care at least for the way their future preferences build upon each other, than to simply attribute to them the desire to be able to satisfy their preferences, independently of their substance or the way they have been formed.

The concept of “opportunity to learn” builds upon the hypothesis that the space of possible paths of preference development that can be realized is not infinitely open. Rather, it is structured, due to the fact, e.g., that certain preferences are habit-forming and path-dependent. The concept proposes that the proper criterion to judge institutional arrangements is their capacity to allow people to engage in “effective preference learning”, i.e., to acquire new preferences, and to develop a specialized and refined system of preferences. In order to achieve this, they must be able to regulate their exposure to learning opportunities (that are potentially overwhelming, see Schwartz, 2004) in a self-determined way. This implies, first, that preference change is understood as a perfectly natural concomitant of the real evolving world in which we live, rather than as a potential source of irrationalities. The second implication is that it is not the satisfaction of given or even “potential” preferences *per se* that constitutes well-being, but rather the ability to use acquired preferences as stepping stones toward their further development. The underlying normative intuition owes much to Frank Knight (1923: 14-15, italics in the original), who argued

“the chief thing which the common-sense individual actually wants is not satisfactions for the wants which he has, but more, and *better* wants ... Wants and the activity which they motivate constantly look forward to new and “higher”, more evolved and more enlightened wants and these function as ends and motives of action beyond the objective to which desire is momentarily directed ... Life is not fundamentally a striving for ends, for satisfactions, but rather for bases for further striving ... the true achievement is the refinement and elevation of the plane of desires, the cultivation of taste”.

The normative intuition expressed here bears, of course, resemblance to Mill’s account of personal self-development in the sense of character formation withstanding the “tyranny of

public opinion” (McPherson 1982).³⁶ Thus, we argue that when faced with the prospect of uncertainty about their own future preferences, individuals will value the “opportunity to learn”, i.e., the ability to engage in effective preference learning in a self-determined way, rather than the ability to satisfy whatever preference they may end up with in future periods.

The qualifier “effective” presupposes that we can make a distinction between learning processes that are conducive to what Knight refers to as “cultivation of taste”, and those that are not. It would be tempting to base such a distinction on hedonic grounds by simply classifying those kinds of learning as effective that produce pleasurable outcomes. Given Sugden’s strong reservations against normative hedonism, though, such an approach is a non-sequitur. For similar reasons, it is hopeless to formulate the distinction in terms of the evaluative space proposed by the capability approach or the one suggested by the “behavioral welfare economics”.³⁷ It seems that the only option we have is to show that there are conditions under which, perhaps paradoxically, the ongoing learning of new preferences (i.e., the use of opportunities to learn) can result in preference dynamics that reduce well-being in terms of available “opportunities”. Freedom to choose may then, under certain conditions, not further a person’s self-development! Put differently, within Sugden’s non-welfarist paradigm, it is plausible to assume that a “responsible” agent is prepared to accept consequences of her own choices that are harmful in terms of utility, be it understood as happiness or the satisfaction of preferences. But it would be incoherent to assume her to accept and “live with the consequences” of harm in terms of lost opportunities for her future selves – for that would run counter to her postulated contractarian interest.

One may indeed argue that the “wrong” kind of learning can be ultimately self-defeating in undermining effective freedom in the sense understood here, i.e., as the ability to engage in further learning. In order to clarify this possibility, let us introduce the notion of “perceived opportunity”, meaning, in slight variation of Sugden’s definition of the term “opportunity”,³⁸ something that the agent *perceives to have* – or: is aware of having – the power to bring about, if he so chooses. In most cases, real (or objective) and perceived (or effective) opportunity will be roughly identical. Consider, however, being addicted to some kind of behavior (such as drug

³⁶ See Mill (1962: 186): “It is the privilege and proper condition of a human being, arrived at the maturity of his faculties, to use and interpret experience in his own way”.

³⁷ All these candidate notions of well-being are discussed in Sugden (2008b) and McQuillin and Sugden (2012b).

³⁸ See Sugden (2010: 49).

consumption or gambling). Given the peculiar satiation characteristics of addictive consumption activities (Witt 2001, 2010), agents may be “trapped” in behavioral dynamics that systematically frustrate their attempts to satisfy their preferences in a more than transient way. What is more important in our context, though, is the possibility that the agents “turn blind” toward products and activities that are unrelated to their addiction. While they may still be, from an objective point of view, free to pursue these alternative paths of preference acquisition, their perceived freedom to do so can be blocked. Analogous reasoning may be applied to the cases of overborrowing (Anand and Gray 2009) or “status races” (Frank 2008).³⁹ All these are cases where a loss in perceived opportunities effectively reduces the agent’s opportunities to learn. Apart from that, they also indicate a loss of well-being in terms of self-determined integration (Ryan 1995).⁴⁰

The shift in emphasis from “opportunity” to “opportunity to learn” opens up interesting new perspectives. As we have seen, Sugden rejects the orthodox way to model deviations from rational choice axioms as conflicts between different selves – in contrast, the continuing person, understood as a locus of responsibility, endorses all her selves’ preferences. As an alternative, we could model the continuing person as a “locus of learning”. She would be concerned, at any given period of time, with the way her future selves’ preferences build upon each other. We submit that that the continuous pursuit of acquiring new preferences, subject to the maintenance of a minimum degree of subjectively perceived coherence, constitutes what real-world individuals consider worthwhile. In other words, autonomy as the actual source of personhood would then be of a procedural nature – namely, the individual experience, shaped by the dynamic interplay with the social environment, of learning new preferences in an autonomous way. Seen in this perspective, a “responsible” agent is concerned with his own ongoing personal capacity to acquire ever new preferences, rather than with his opportunity to satisfy whatever preference he happens to have (as a result of this capacity) in the future.

Methodologically, this refocusing involves a shift toward a motivation-theoretic perspective that looks at *how sequences of preferences are interconnected through time* – rather

³⁹ Cordes and Schubert (2010) present a formal model of these self-defeating behavioral dynamics.

⁴⁰ The implications of this kind of self-defeating learning dynamics may also be illustrated in the following way: Under normal circumstances, the fact that my preferences are shaped by arbitrary contextual factors does not make it implausible for me to value opportunity. For I may wish to retain the chance to consciously choose *which of those factors I let shape my preferences*. If I end up stuck in one of the self-defeating learning dynamics described above, though, I lose the ability to do this “meta”-choosing.

than the decision-theoretic focus on isolated choice situations. We submit that this perspective is particularly adequate for a continuing person, understood as a composition of all her past, present and future “time-slice selves”, as Sugden puts it. In that role, the continuing person will typically be engaged in managing the future development of her own preferences. One of her aims in that process can be assumed to be to find out – in an effortful process of trial and error – what she “really” wants (where that of course will itself change over time).

Hence, the appropriate question to ask is not whether individuals, when faced with the prospect of incoherent preferences, will *endorse* (Sugden’s view) or *disavow* (the received view) these preferences. Rather, the question is whether under these conditions “ambient supports” are in place ensuring that the individuals’ ability to engage in the ongoing learning of new preferences is not jeopardized.

Some clarifications are in order at this point: First, the criterion that we suggest does not aim at establishing some external standpoint from which the actual or potential preferences of individuals ought to be judged and possibly overridden. It does not rely on a “reason to value” formula (as does Sen’s approach) that is supposedly substantiated in some democratic process, reflecting a particular theory of the good life. The “opportunity to learn” criterion simply points to the need to establish the “ambient supports” that are required to maintain people’s ability to engage in the ongoing learning of new preferences. It does not depend on any *substantive* account of human well-being.

Second, the “locus of learning” concept implies that the present self is indeed privileged relative to the future selves. The present self however does not wish to restrict the future selves’ opportunities to indulge in their own preferences. She does not want to constrain them in order to align their choices with her own current – and probably transient – conception of what is good for her. So far, we follow Sugden (2004a: 1018). The difference is, though, in the assumption that, being concerned with those selves’ abilities to continue the process of preference learning, the present self guards them from following those paths of learning that ultimately risk undermining this ability. Again, this concern goes beyond any particular conception of the good.

We now face a dilemma: On the one hand, the “guarding” seems to require restrictions on the freedom to choose of the future selves in some cases; on the other hand, maintaining the ongoing opportunity to learn calls against any kind of restrictions – particularly in light of the epistemic constraints under which all individuals and policy advisors operate. We wish to heed

the general liberal claim that “being able to choose how to live one’s life is an aspect of individual well-being in its own right” (Sugden 2003: 785). Specifically, we want to respect Mill’s famous dictum that options also contribute to well-being if they are in fact not chosen and even if they are, from some moral point of view, “worthless”, for it is through the *act* of choosing that an agent cultivates the (intrinsically valuable) faculties of observation, reason, judgment and self-control.

It seems that the policy tool of “nudges” (Thaler and Sunstein 2008) can overcome this dilemma, provided its well-known “freedom requirement” is taken seriously. Put differently, nudges allow it to be the case that specific individual liberties are promoted without abandoning the protection of other liberties.⁴¹ That means that individuals can benefit from effective safeguards against slipping into self-defeating learning dynamics in the sense specified above, without foregoing their freedom of choice. This freedom of choice would not be justified as giving *homines oeconomici* the chance to realize their rational life plans, but rather as allowing cognitively constrained learning agents to try out new “experiments in living”, as Mill would put it (see also Schubert 2012: 256).

The general question one were to ask in thinking about policy implications would be: “Would the individual be able, given her material circumstances, to acquire such-and-such preference?”⁴² To illustrate, consider the way Sugden deals with the “adaptation problem” brought up by Sen (1988: 45-46). Sugden focuses on the “hopeless beggar’s” formal freedom to “walk away”, once he has “changed his mind” (read: his preferences).⁴³ This however hardly captures the normative intuition that made Sen worry about the case. We submit that it would be far more plausible to diagnose the beggar as having lost the “opportunity to learn”, i.e., the ability to engage in effective preference learning, where this presupposes the ability to *perceive* the opportunities available. Analogously, instead of asking “What if the slave would cease to be contented...” (Sugden 2006: 46), we suggest to ask “What if the slave *cannot, on his own*, cease to be contented....” Despite their (assumed) demonstrated happiness, the beggar and the slave have adapted to their situation by *resignation*.

⁴¹ Qizilbash (2011a: 37-39) for the distinction between promoting and protecting liberties.

⁴² Compare this to Sugden’s guiding question: “*Were the individual to have* such-and-such a preference, would she be able to satisfy it?” (Sugden 2010: 50, italics in the original).

⁴³ See section 2, above, and Sugden (2006a).

5. SOME POLICY IMPLICATIONS

Sugden's concept of opportunity is motivated, first and foremost, by the wish to refute the claim, advanced by some proponents of "behavioral welfare economics", that the empirical evidence on the irrationality of individuals' preferences (i) justifies paternalistic interventions, and (ii) even makes opposition against them "incoherent". Such conclusions can be avoided by establishing a normative standard that attributes value to the unfettered exercise of "choice" independently of its consequences in terms of well-being. It is instructive to compare Sugden's strategy to do so with that of Buchanan (1991) in his classic defense of normative individualism. Both argue from within the contractarian approach to normative reasoning.

As Buchanan argues, the liberal case against paternalism is only vulnerable to the argument from behavioral economics if it is itself based on the assumption that the individual holds an "epistemic privilege" with respect to "what is best for her" – in other words, if it is presumed that the individual's choices exist independently of some construct named "utility" that is supposed to be maximized. Buchanan rejects this presumption in favor of a subjectivist viewpoint that states that "[a]ll there is are individual choices" (ibid.: 225) and that "[t]hat which is best' is objectively meaningful only at the moment of choice" (ibid.: 229). Freedom of choice, then, is valuable independent of whether people actually care about the process of choosing. Buchanan's strategy, while relying on a strong normative notion of personal responsibility, has the advantage of allowing the possibility that individuals may, through their voluntary choices, attribute legitimacy also to "choice-restricting" institutions. While Sugden repeatedly concedes that genuine⁴⁴ restrictions on one's freedom to choose may be appropriate in "exceptional cases" (Sugden 2006b: 218), and that opportunity does not "trump other moral imperatives",⁴⁵ he does not integrate these cases in his theoretical concept in a systematic way.

Sugden's alternative way to prove the possibility of coherent anti-paternalism via an assumption about people's attitudes toward their future potential preferences avoids postulating a normative principle that is completely detached from individuals' own values. Nevertheless, it is supposed to support the claim that the individual "benefit" is constituted by "voluntary choice" only (McQuillin and Sugden 2012a: 618). As we have seen, though, the specific assumption

⁴⁴ As opposed to merely apparent restrictions that in fact enlarge the freedom of the individual (at least in her role as consumer): Mill's well-known example of the prohibition to sell oneself into slavery and mandatory cooling-off periods are cases in point (Sugden 2008b: 320). See also Qizilbash (2011b: 31-32).

⁴⁵ See Sugden (2008a: 247; 2009: 95-96; 2010: 61), McQuillin and Sugden (2012a: 630).

about agents' attitudes it is based upon lacks empirical plausibility. As a genuinely psychological hypothesis it is vulnerable to the objection that it is descriptively inadequate. Hypotheses about such attitudes can hardly be defended without considering the way those potential preferences may (one day) be actualized, i.e., effectively acquired. That in turn invites a focus on the way preferences are interconnected over time, and on the individual's attitude toward that interconnection. What Sugden's concept *can* offer is to reveal some "hidden" normative costs of paternalistic interventions. We suggest, however, (i) that our criterion of "opportunity to learn" can achieve this even better, besides (ii) providing a coherent account of showing which of the tools proposed by libertarian paternalism tools are legitimate under which conditions.

As to point (i), it seems obvious that many of the applications proposed by "libertarian paternalists" risk constraining the opportunities of individuals to try out and acquire new preferences on their own. To illustrate, consider the effect that nudges promoting the choice of healthy food, say, are likely to have on the evolution of corresponding social norms (Schnellenbach 2012). Lacking a well-defined normative foundation, they will probably stabilize social norms that ostracize the choice of unhealthy food, thereby promoting conformism and increasing the psychological costs of trying out any preference for this kind of products. By hindering "opportunities to learn", they qualify as genuine normative costs within the dimension of well-being proposed here (and beyond those mentioned by Sugden).⁴⁶

Concerning point (ii), the criterion suggested here would license the use of nudges only in those cases, where the acquisition of a preference risks leading the agent into learning dynamics that are self-defeating in the sense of constraining the set of individually perceived opportunities. Addictive goods and activities, as well as "status races" are cases in point. Specifically, we submit that if there exist privately construed nudges in the marketplace that steer consumers into acquiring preferences of this kind, it is legitimate for the state to respond by introducing robust "counter-nudges". The negative effect on individual's learning opportunities is then potentially offset by a positive effect within the same dimension of benefit.

⁴⁶ It seems that this category of costs is overlooked by Sugden, despite the prominence that the risk of a "tyranny of public opinion" assumes in the work of Mill.

6. CONCLUDING REMARKS

The present paper has argued that the “opportunity criterion” proposed by Robert Sugden suffers from the fact that its microfoundations lack empirical plausibility. Hence, his specification of contractarian interest is inadequate. We have suggested an alternative criterion that can be understood as a dynamic variant of Sugden’s concept, in that it stresses people’s opportunity to learn new preferences.

In light of our criterion, the market is valued from the perspective of individuals who do not yet have well-formed preferences (or possibly even no preferences at all) and *who have to acquire them*. Agents value markets to the extent that they provide the chance to do so and to experience self-development. In order to achieve this, self-regulation tools may be required that the market may not provide. Importantly, the mere fact that preferences are incoherent does not justify paternalistic intervention. We thus reject the claim made by some libertarian paternalists that the – undisputable – fact that social context shapes preferences “undermines” the belief that agents should be free to choose and to make their own mistakes. Paternalistic interventions are, however, legitimate in those (probably rare) cases where the acquisition of new preferences engenders learning dynamics that are likely to result in a loss of perceived opportunities, for that amounts to a loss of opportunities to learn.⁴⁷ Contrary to Sugden’s view, real-world people are unlikely to unconditionally endorse any preferences they may acquire in the future. What truly matters, in our view, is not the chance to satisfy whatever preference I may have in future periods, but, first and foremost, the chance to develop those preferences.

The concept of “opportunity to learn” may allow measurement of the scope of an agent’s opportunity set, something that was foreclosed by Sugden’s insistence on purging the concept from any reference to actual preferences. It could do so by imposing some (limited) conceptual structure on the space of a person’s options; given what we know about the universal mechanisms driving human preference learning, the range of potential preference paths is not infinite. Exploring these issues must obviously be left to future research.

⁴⁷ Neglecting this issue may even have problematic distributive implications: As Beaulier and Caplan (2007) point out, the poor tend to be more susceptible to this kind of dysfunctional learning dynamics than the rich.

References

- Akerlof, G.A., Dickens, W.T. 1982. The Economic Consequences of Cognitive Dissonance. *American Economic Review* 72: 307-319.
- Anand, P., Gray, A. 2009. Obesity as Market Failure: Could a 'Deliberative Economy' Overcome the Problem of Paternalism? *Kyklos* 62: 182-190.
- Arrow, K. 1995. A Note on Freedom and Flexibility. In *Choice, Welfare and Development – A Festschrift in Honour of Amartya K. Sen*, ed. K. Basu, P. Pattanaik, K. Suzumura, 7-16. Oxford: Oxford University Press.
- Beaulier, S., Caplan, B. 2007. Behavioral Economics and Perverse Effects of the Welfare State. *Kyklos* 60: 485-507.
- Berg, N., Gigerenzer, G. 2007. Psychology implies paternalism? Bounded rationality may reduce the rationale to regulate risk-taking. *Social Choice & Welfare* 28: 337-359.
- Bernheim, D.B. and A. Rangel 2009. Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics* 124: 51-104.
- Buchanan, J.M. 1991. The foundations for normative individualism. In *The Economics and the Ethics of Constitutional Choice*, ed. J.M. Buchanan, 221-229. Ann Arbor: University of Michigan Press.
- Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T. and Rabin, M. 2003. Regulation for Conservatives: behavioral economics and the case for 'asymmetric paternalism'. *University of Pennsylvania Law Review* 151: 1211-1254.
- Caplan, B. 2001. Rational irrationality and the Microfoundations of Political Failure. *Public Choice* 107: 311-31.
- Christman, J., 1991. Autonomy and Personal History. *Canadian Journal of Philosophy* 21: 1-24.
- Christman, J., 2005. Saving Positive Freedom. *Political Theory* 33: 79-88.
- Cordes, C., Schubert, C. 2010. Role Models that make you unhappy: Light Paternalism, Social Learning and Welfare. *Papers on Economics & Evolution* #1022.
- Deci, E.L., Ryan, R.M. 2000. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry* 11: 227-268.
- Dworkin, G. 1972. Paternalism. *The Monist* 56: 64-84.
- Elster, J. 1982. Sour Grapes - Utilitarianism and the Genesis of Wants. In *Utilitarianism and Beyond*, ed. A.K. Sen, B.A. Williams, 219-38. Cambridge MA: Cambridge University Press.
- Frank, R.H. 2008. Should Public Policy Respond to Positional Externalities? *Journal of Public Economics* 92: 1777-1786.
- Hayek, F.A. 1989. Order – with or without design? London: Centre for Research into Communist Economies.
- Hoeffler, S., Ariely, D. 1999. Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology* 8: 113-139.
- Hume, D. 1740/1978. *Treatise of Human Nature*. Oxford: Clarendon Press.
- Isen, A.M. 1999. Positive affect. In *Handbook of cognition and emotion*, ed. T. Dalgleish and M.J. Power, 521-539. Wiley, London.
- Khalil, E. 2010. Adam Smith's concept of self-command as a solution to dynamic inconsistency and the commitment problem. *Economic Inquiry* 48: 177-191.
- Knight, F.H. 1923. *The Ethics of Competition and other Essays*. New Brunswick: Transaction Publishers.

- Loewenstein, G. 1999. Willpower: A Decision-Theorist's Perspective. *Law and Philosophy* 19: 51-76.
- McPherson, M.S. 1982. Mill's Moral Theory and the Problem of Preference Change. *Ethics* 92: 252-273.
- McQuillin, B. and R. Sugden 2012a. How the market responds to dynamically inconsistent preferences. *Social Choice & Welfare* 38: 617-634.
- McQuillin, B. and R. Sugden 2012b. Reconciling normative and behavioural economics: the problems to be solved. *Social Choice & Welfare* 38: 553-567.
- Mill, J.S. 1962. *Utilitarianism and On Liberty*. London: Fontana Library.
- O'Donoghue, T., Rabin, M. 1999. Doing it now or later. *American Economic Review* 89: 103-124.
- Qizilbash, M. 2011a. Sugden's Critique of the Capability Approach. *Utilitas* 23: 25-51.
- Qizilbash, M. 2011b. Sugden's Critique of Sen's capability approach and the dangers of libertarian paternalism. *International Review of Economics* 58: 21-42.
- Qizilbash, M. 2012. Informed desire and the ambitions of libertarian paternalism. *Social Choice & Welfare* 38: 647-658.
- Roemer, J.E. 1998. *Equality of Opportunity*. Cambridge MA: Harvard University Press.
- Rothenberg, J. 1962. Consumers' Sovereignty Revisited and the Hospitability of Freedom of Choice. *American Economic Review – Papers & Proceedings* 52: 269-283.
- Ryan, R.M. 1995. Psychological Needs and the Facilitation of Integrative Processes. *Journal of Personality* 63: 397-427.
- Ryan, R.M., Huta, V., Deci, E.L. 2008. Living well: a self-determination theory perspective on eudaimonia. *Journal of Happiness Studies* 9: 139-170.
- Sally, D. 2000. I, too, sail past: Odysseus and the Logic of Self-Control. *Kyklos* 53: 173-200.
- Schelling, T. 1984. Self-Command in Practice, in Policy, and in a Theory of Rational Choice. *American Economic Review* 74: 1-11.
- Schnellenbach, J. 2012. Nudges and Norms: On the political economy of soft paternalism. *European Journal of Political Economy* 28, 266-277.
- Schubert, C. 2012. Pursuing Happiness. *Kyklos* 65: 245-261.
- Schumpeter, J.A. 1949. English Economists and the State-Managed Economy. *Journal of Political Economy* 57: 371-382.
- Schwartz, B. 2004. *The paradox of choice*. New York: Harper Collins.
- Sen, A.K. 1988. *On Ethics and Economics*. Oxford: Basil Blackwell.
- Sen, A.K. 1992. *Inequality Reexamined*. Cambridge MA: Harvard University Press.
- Sugden, R. 1989. Maximizing Social Welfare: Is it the government's business? In *The good polity*, ed. A. Hamlin, P. Pettit. Oxford: Basil Blackwell, 69-86.
- Sugden, R. 1998a. The metric of opportunity. *Economics & Philosophy* 14: 307-337.
- Sugden, R. 1998b. Measuring Opportunity: Toward a contractarian measure of individual interest. *Social Philosophy & Policy* 15: 34-60.
- Sugden, R. 2002. Beyond Sympathy and Empathy: Adam Smith's Concept of Fellow-Feeling. *Economics & Philosophy* 18: 63-87.
- Sugden, R. 2003. Opportunity as a Space for Individuality: Its Value and the Impossibility of Measuring it. *Ethics* 113: 783-809.
- Sugden, R. 2004a. The Opportunity Criterion: Consumer Sovereignty without the Assumption of Coherent Preferences. *American Economic Review* 94: 1014-1033.
- Sugden, R. 2004b. Living with unfairness: The limits of equal opportunity in a market economy, *Social Choice & Welfare* 22: 211-326.

- Sugden, R. 2006a. 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. 2006b. Taking unconsidered preferences seriously. In *Preferences and Well-Being*, ed. S. Olsaretti. Cambridge: Cambridge University Press, 209-232.
- Sugden, R. 2007. The value of opportunities over time when preferences are unstable. *Social Choice & Welfare* 29: 665–682.
- Sugden, R. 2008a. Why incoherent preferences do not justify paternalism. *Constitutional Political Economy* 19: 226-248.
- Sugden, R. 2008b. Capability, Happiness and Opportunity. In *Capabilities and Happiness*, ed. L. Bruni, F. Comim, M. Pugno, 299-322. New York: Oxford University Press.
- Sugden, R. 2009. Market simulation and the provision of public goods: A non-paternalistic response to anomalies in environmental valuation. *Journal of Environmental Economics and Management* 57: 87-103.
- Sugden, R. 2010. Opportunity as Mutual Advantage. *Economics & Philosophy* 26: 47-68.
- Sugden, R., 2011. The behavioural economist and the social planner: To whom should behavioural welfare economics be addressed? *Papers on Economics & Evolution* #1121, ed. Max Planck Institute of Economics, Jena, Germany.
- Sumner, L.W. 1996. *Welfare, Happiness and Ethics*. New York: Oxford University Press.
- Sunstein, C.R., Thaler, R.H. 2003. Libertarian Paternalism is not an Oxymoron. *University of Chicago Law Review* 70: 1159-1202.
- Thaler, R.H. and C.R. Sunstein 2008. *Nudge: improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.
- Witt, U. 2001. Learning to consume – A Theory of Wants and the Growth of Demand. *Journal of Evolutionary Economics* 11: 23-36.
- Witt, U. 2010. Economic Behavior – Evolutionary vs Behavioral Perspectives. *Papers on Economics & Evolution* # 1017, ed. Max Planck Institute of Economics, Jena, Germany.
- Yaari, M.E. 1977. Endogenous Changes in Tastes: A Philosophical Discussion. *Erkenntnis* 11: 157-196.