DOI: 10.1111/ecaf.12374

DISCUSSION

WILEY ECONOMIC AFFAIRS

The community of advantage

Robert Sugden

School of Economics, University of East Anglia

Correspondence
Email: R.Sugden@uea.ac.uk

1 | INTRODUCTION

In 2018 I completed a book that I had been working on for longer than I like to recall: *The Community of Advantage: A Behavioural Economist's Defence of the Market* (Sugden, 2018). Its title is taken from a passage in John Stuart Mill's *Principles of Political Economy*, first published in 1848. Mill is contrasting what he sees as the enlightened economic views of his own time with the mercantilism of the past. Mercantilism, he says, treats economic relations between countries as if they were relations between rival tradesmen trying to undersell one another. This obsolete way of thinking “overruled for centuries all sense of the general community of advantage which all commercial countries derive from the prosperity of one another” (Mill, 1871, III. xxv. §1). Mill is telling his readers that the market is a *community of advantage*.

By ‘advantage’, Mill clearly means that a market is a network of mutually beneficial transactions. But what does he mean by ‘community’? Isn't the market simply an institutional framework in which individuals pursue their respective self-interests? Mill thinks not:

> *[C]ommerce first taught nations to see with good will the wealth and prosperity of one another. Before, the patriot ... wished all countries weak, poor, and ill-governed, but his own: now he sees in their wealth and progress a direct source of wealth and progress to his own country. (1871, II. xvii. §5)*

His idea is that, by participating in relationships of mutual benefit, people come to see one another as cooperative partners rather than as rivals in a zero–sum game. My book is an attempt to express to modern readers the idea of the market as a system of voluntary cooperation that each of us can recognise as valuable.

2 | BACKGROUND

As the subtitle of the book suggests, my stance as author is that of a behavioural economist *and* an economic liberal. I am trying to counter the increasingly widespread belief that the findings of behavioural economics justify regulatory interventions in the economy that would previously have been thought unacceptably paternalistic. This view of the normative implications of behavioural economics was expounded in two influential early papers with the significant titles 'Libertarian Paternalism' (Sunstein & Thaler, 2003) and 'Regulation for Conservatives' (Camerer, Issacharoff, Loewenstein, O'Donoghue, & Rabin, 2003). Notice the suggestions that these papers will provide justifications of paternalism and regulation to which libertarians and conservatives cannot object. Colin Camerer and his co-authors are particularly explicit about this, claiming that by "cataloging a list of common decision-making errors that even highly competent, well-functioning people make in predictable situations" (p. 1218), behavioural economics reveals new ways in which unregulated markets might fail to promote welfare. Earlier generations of economists proposed regulations to counter imperfect competition and imperfect information; in the same spirit, Camerer et al. say, behavioural economists now propose paternalistic regulations to counter the imperfect rationality of economic agents (p. 1218). Similarly, Cass Sunstein and Richard Thaler (2003, p. 1162) claim that individuals are liable to make "inferior decisions in terms of their own welfare – decisions that they would change if they had complete information, unlimited cognitive abilities, and no lack of willpower", and argue for "self-conscious efforts, by private and public institutions, to steer people's choices in directions that will improve their own welfare".

If one starts (as most economists have done) from the presupposition that normative economics is about how to organise an economy so that individuals' preferences are maximally satisfied, and if one interprets behavioural economics as showing that individuals make errors in implementing their own preferences, Camerer et al. are perhaps justified in treating such errors as just another kind of market failure. But a liberal might still be uneasy about the suggestion that it is part of the job of a social planner or regulator to determine what individuals' 'true' preferences are and to use those preferences, rather than preferences revealed in individuals' actual decisions, as the basis for normative judgements about public policy.

Some of my fellow behavioural economists think I am being contrarian in opposing paternalism while contributing to a research programme that is discovering the many ways in which individuals' decisions are inconsistent with the received theory of rational choice. But it is important to realise that, until the early 2000s, the research programme of what is now called behavioural economics was almost entirely empirical. It challenged the descriptive validity of rational choice theory, but did not develop a normative theory of its own. I worked in that research programme throughout the 1980s and 1990s, while also working as a theoretical and philosophical economist, in what I saw as the liberal tradition of David Hume, Adam Smith, John Stuart Mill, Friedrich Hayek, and James Buchanan. I never thought that these two aspects of my work were in tension. A lot of my theoretical work was concerned with identifying conceptual limitations of received rational choice theory and proposing alternative ways of understanding rationality and normative judgement. These proposals often worked by downplaying the role of preferences.

For example, this was a time when many economists and philosophers were struggling with the problem of how to represent the value of individual liberty within the formal framework of social choice theory. The moral requirement that each individual should have a 'protected sphere' of personal choice was construed as implying some restriction on the relationship between individual preferences and 'social preferences', but attempts to formalise that restriction ran into difficulties if individuals had 'meddlesome' preferences about one another's private affairs. I showed that the problem dissolved if one forgot about preferences and defined each individual's protected sphere in terms of his opportunities for choice. How each individual chose to use these opportunities was up to him, and irrelevant for the definition of liberty (Sugden, 1985). In another contribution to social choice theory, Peter Jones and I considered how the extent of an individual's 'effective freedom' might be measured. We argued that this measure should not depend on the individual's *actual* preferences over the options in her opportunity set, but instead on how well that set catered to the range of preferences that would be reasonable or normal for someone in her position (Jones & Sugden, 1982). In a third strand of work, I developed a theory of *team reasoning* in which non-selfish behaviour does not directly reveal altruistic or 'social' preferences; instead, each individual is motivated to play his part in mutually beneficial practices, conditional on others doing the same (Sugden, 1993).

When, in 2000 or thereabouts, I set about trying to find a form of normative economics that was compatible with behavioural findings, I built on this earlier work. Since the fundamental problem was that individuals often seemed to lack coherent, context-independent preferences, my strategy was to see what could be done without using the concept of preference at all. In place of the preference-satisfaction criterion of traditional welfare economics, I substituted a criterion of opportunity. By defining people's opportunities independently of their preferences, I hoped to cut through the problems caused by the inadequacy of the standard assumptions of rational choice theory. By attributing normative value to individuals' opportunities without taking any view about how those opportunities should be used, I saw myself as upholding the non-paternalistic traditions of liberal economics. The first finished product of this programme of work was a paper that showed how some of the classic theoretical results of welfare economics could be reformulated in terms of opportunity (Sugden, 2004). The underlying idea was that competitive markets are desirable, not as an efficient means for satisfying people's *given* preferences, but as an institutional structure that provides opportunities for whatever voluntary transactions people might want to make. In the words of Buchanan (1964, p. 219), who has been a major influence on my work:

> *The 'market' or market organization is not a means toward the accomplishment of anything. It is, instead, the institutional embodiment of the voluntary exchange processes that are entered into by individuals in their several capacities. That is all there is to it.*

As my approach to normative analysis is very different from familiar forms of welfare economics, I have not been surprised that it has remained a minority taste. But I was, and still am, surprised that behavioural economists have been so ready to interpret contraventions of rational choice theory as evidence of decision-making error. In the pioneering era of the 1980s and 1990s, this was exactly the interpretation of anomalies that mainstream economists typically favoured, and that we behavioural economists disputed. (Look, for example, at the exchange between Charles Plott, 1996 and Daniel Kahneman, 1996. Plott is suggesting that decision-making behaviour that contravenes rational choice theory is the result of errors which, with sufficient experience, individuals learn to correct. Kahneman [1996, pp. 251–2] argues that, viewed in the perspective of empirical psychology, Plott's distinction between "latent" rationality and error is "deeply problematic".) As some of us used to say, it is as if decision-makers are held to be at fault for failing to behave as the received theory predicts, rather than that theory being faulted for failing to make correct predictions. It seems to me that, in their desire to make normative pronouncements, many of my fellow behavioural economists are not taking the psychological evidence sufficiently seriously.

3 | THE ARGUMENT OF THE BOOK

My book weaves together four main strands of argument.

3.1 | The addressee

The first strand concerns a question that economists rarely ask explicitly: To whom should normative economics be addressed? Neoclassical and behavioural economists usually write as if addressing an imagined 'social planner', implicitly conceptualised as a benevolent autocrat. Philosophers who write about normative economics sometimes imagine instead that they are engaging in 'public reasoning', addressing an assembly of moral agents who are trying to decide what, all things considered, is good for people. Both approaches treat normative analysis as an attempt to find a 'view from nowhere' – an impartial view of what is good for people that is not any actual person's view of what is good for him or her.

Following Buchanan, I propose a different approach – that of *contractarianism*. Instead of taking a view from nowhere, a contractarian imagines herself addressing citizens as individuals, advising them how to reach mutually beneficial agreements. The logic of this approach disallows paternalism. If you are addressing a social planner or an assembly of public reasoners, you can coherently recommend a policy on the grounds that it promotes the welfare of certain individuals, even though those individuals do not want it to be adopted. But the contractarian has to show that her recommendations are in the separate interests of each individual, *as that individual perceives those interests*. For a contractarian, what is wrong with paternalism is not that it has morally objectionable properties, as viewed from nowhere. It is that a paternalistic proposal is not properly addressed. If you want to prevent someone from making a choice that you think would be bad for her, the person you need to address is *her*, not someone who has assumed the role of a self-appointed guardian.

3.2 | Behavioural welfare economics and paternalism

The second strand is a critique of 'behavioural welfare economics' – the approach to normative analysis that was proposed by Camerer et al. (2003) and Sunstein and Thaler (2003), and which is now favoured by most behavioural economists. This critique originated in a paper I wrote jointly with Gerardo Infante and Guilhem Lecouteux (Infante, Lecouteux, & Sugden, 2016). Behavioural welfare economics implicitly assumes that people have context-independent 'true' or 'latent' preferences which, because of psychologically induced errors, are not always revealed in actual choices. It aims to reconstruct latent preferences by identifying and removing the effects of error on decisions, and to design policies to satisfy those preferences. Its implicit model of human agency is of an 'inner rational agent' that interacts with the world through an imperfect psychological 'shell'. I argue that there is no satisfactory evidence to support this model, and no credible psychological foundation for it. If the concept of true preference has no empirical content, the idea that such preferences can be reconstructed is a mirage.

3.3 | Normative economics without preferences

The third strand of the book proposes an alternative form of normative economics that does not use the concept of preference. I begin by looking for a normative criterion that citizens might reasonably want economists to use when making recommendations about public policy. As a first step, I propose the Individual Opportunity Criterion – the principle that, as viewed by each citizen separately, more opportunity for that person is better than less. When an individual makes choices over time, 'more opportunity' includes opportunities for her to change her mind from one period to another. I argue that each citizen can endorse this criterion without needing to believe that her own preferences are rationally consistent.

However, it is intrinsic to the idea of a market that the set of opportunities available to any one person depends on the choices that other people make from *their* opportunity sets. Thus, what each person can expect to gain from participating in the market cannot be described by any particular opportunity set, specified independently of other people's choices. What people collectively can expect from the market is a rich array of opportunities to trade with each other on terms that they might find *mutually* acceptable. Building on Sugden (2004) and on work I have done with Ben McQuillin (McQuillin & Sugden, 2012), I propose an Interactive Opportunity Criterion which (roughly speaking) says that individuals collectively should have as much opportunity as possible to carry out whatever voluntary transactions they might want to make. ('Might' here ranges over the entire range of preferences that a person might plausibly have: there is no reference to *actual* preferences.) I show that this criterion is satisfied in the equilibrium state of a competitive market, also defined without reference to preferences. This result is encapsulated in one of my favourite slogans. Let us say that a person is willing to pay for a good if he is willing to give up what would induce others (or, if he is liable to change his mind, himself at another point in time) to supply it. Then to say that the Interactive Opportunity Criterion is satisfied is to say that *each person is able to get whatever he wants and is willing to pay for, when he wants it and is willing to pay for it*.

Recognising that this is a result about a highly idealised economy, I consider a range of conditions that are usually considered as 'market failures' to be corrected by governmental regulation. I discuss these conditions, and possible responses to them, from a contractarian viewpoint. I argue that neoclassical arguments for regulations against cartels and against the exploitation of monopoly power can be endorsed on contractarian grounds, as can certain kinds of regulations against spurious complexity in pricing. But I raise doubts about the significance of behavioural arguments for regulation that assume choice overload or preferences for self-constraint.

Finally, I ask what properties a market economy must have if it is to be *psychologically stable* – that is, if it is to reproduce a general belief that its governing principles are fair. I argue that, because of the division of knowledge and because the opportunities open to each person depend on how other people choose to use *their* opportunities, markets cannot guarantee that individuals are rewarded according to the efforts they make, or according to the talents they put to use. Here I follow Hayek (1976, p. 94), who recognises that "in the cosmos of the market we all constantly receive benefits which we have not deserved in any moral sense". Psychological stability has to rest on continuing expectations of mutual benefit, defined relative to a baseline that evolves over time and that cannot be justified in terms of abstract principles of fairness. Nevertheless, if the market is to be recommended to each individual separately, each individual must be able to expect to share in the benefits that markets create. Maintaining such expectations typically requires some form of social insurance, but if markets are to work, there cannot be full insurance against underserved losses.

3.4 | The Principle of Mutual Benefit

The final strand of the book considers the moral status of markets. Virtue ethicists such as Elizabeth Anderson (1993) and Michael Sandel (2012) have argued that there are 'moral limits' to the proper domain of the market: market relations (it is said) are based on self-interested and instrumental motivations, and are thereby morally impoverished. A related idea in economics is that socially valuable practices (particularly those of trust and reciprocity) can depend on pro-social and intrinsic motivations which the market tends to 'crowd out'. An important strand of behavioural economics is concerned with modelling intrinsic motivation, 'social preferences', and preferences for conforming to social norms. Building on work I have done with Luigino Bruni (Bruni & Sugden, 2013) and Andrea Isoni (Isoni & Sugden, 2019), I argue that market relations can be viewed by those who participate in them as *cooperative*: each party can act with the intention that the parties together achieve mutual benefit. Such an intention is neither self-interested nor altruistic.

To explain this kind of intention, I develop a version of the theory of team reasoning that uses opportunity-based rather than preference-based concepts and makes minimal assumptions about people's rationality. The essential idea is encapsulated in a moral rule or norm, the *Principle of Mutual Benefit*. This requires that when you are participating with others in an interaction that each of you has chosen to enter, and for as long as the other participants follow the principle too, you should behave in such a way that the other participants are able to satisfy normal expectations about the consequences of the interaction for them. In a population of people who act on the principle, those actions tend to reproduce social practices (for example, practices of honest trading) that provide opportunities for mutual benefit. In contrast to the implications of many social-preference theories in which reciprocity is treated as returning kindness for kindness, the hypothesis that people act on the Principle of Mutual Benefit does not lead to the 'Paradox of Trust' – the paradoxical conclusion that if everyone were completely trustworthy, trust would not reveal pro-social intentions, and so could not prompt trustworthiness.

In the final chapter of the book, I argue (in what I take to be the spirit of John Rawls's 1971 account of justice) that a contractarian can recommend to individuals that, consulting only their own interests, they agree that some principle of behaviour is to *count as* moral – that each will use her best endeavours to conform to it, to support others in conforming to it, and to censure others if they fail to conform. I end with a contractarian recommendation to my fellow-citizens: that it is their common interest to accept the Principle of Mutual Benefit.

ACKNOWLEDGEMENT

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement No. 670103, and from the Economic and Social Research Council of the UK (award no. ES/P008976/1).

REFERENCES

Anderson, E. (1993). *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.

Bruni, L., & Sugden, R. (2013). Reclaiming virtue ethics for economics. *Journal of Economic Perspectives*, *27*, 141–164.

Buchanan, J. (1964). What should economists do? *Southern Economic Journal*, *30*, 213–222.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for 'asymmetric paternalism'. *University of Pennsylvania Law Review*, *151*(3), 1211–1254.

Hayek, F. (1976). *Law, Legislation and Liberty. Vol 2: The Mirage of Social Justice*. Chicago, IL: University of Chicago Press.

Infante, G., Lecouteux, G., & Sugden, R. (2016). Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics. *Journal of Economic Methodology*, *23*, 1–25.

Isoni, A., & Sugden, R. (2019). Reciprocity and the Paradox of Trust in psychological game theory. *Journal of Economic Behavior & Organization* (forthcoming). https://doi.org/10.1016/j.jebo.2018.04.015 (accessed 10 September 2019).

Jones, P., & Sugden, R. (1982). Evaluating choice. *International Review of Law and Economics*, *2*, 47–65.

Kahneman, D. (1996). Comment [on Plott (1996)]. In K. Arrow, E. Colombatto, M. Perlman, & C. Schmidt (Eds.), *The Rational Foundations of Economic Behaviour* (pp. 251–254). Basingstoke: Macmillan & International Economic Association.

McQuillin, B., & Sugden, R. (2012). How the market responds to dynamically inconsistent preferences. *Social Choice and Welfare*, *38*, 617–634.

Mill, J. S. (1871). *Principles of Political Economy with Some of their Applications to Social Philosophy* (7th ed.). London: Longmans. First published in 1848.

Plott, C. (1996). Rational individual behaviour in markets and social choice processes: The discovered preference hypothesis. In K. Arrow, E. Colombatto, M. Perlman, & C. Schmidt (Eds.), *The Rational Foundations of Economic Behaviour* (pp. 225–250). Basingstoke: Macmillan and International Economic Association.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Sandel, M. J. (2012). *What Money Can't Buy: The Moral Limits of Markets*. New York: Farrar, Straus & Giroux.

Sugden, R. (1985). Liberty, preference and choice. *Economics and Philosophy*, *1*, 213–229.

Sugden, R. (1993). Thinking as a team: Toward an explanation of nonselfish behavior. *Social Philosophy and Policy*, *10*, 69–89.

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

Sugden, R. (2018). *The Community of Advantage: A Behavioural Economist's Defence of the Market*. Oxford: Oxford University Press.

Sunstein, C., & Thaler, R. (2003). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review*, *70*, 1159–1202.

How to cite this article: Sugden R. The community of advantage. *Economic Affairs*. 2019;39:417–423. https://doi.org/10.1111/ecaf.12374