Journal of Economics (2023) 140:275–278
https://doi.org/10.1007/s00712-023-00836-9

BOOK REVIEW

Dhami, Sanjit and Sunstein, Cass R., *bounded rationality. Heuristics, Judgment and Public Policy*, MIT Press, 2022, 552 pp., $ 60.00 (Paperback)

Ronnie Schöb[1]

Published online: 10 August 2023

This is an exciting book to read. In a well-balanced combination of formal strictness, well-founded intuition and many examples, it offers the reader a comprehensive treatment on recent research on bounded rationality and its potential implications for public policy. As there is no precise definition of rationality, the authors start from the bayesian rationality approach (BRA), for which they offer a precise definition that serves as a benchmark for what constitutes rational behavior. Rational behavior, according to this view, requires that people have unlimited computational and cognitive abilities and make choices according to some well-established axioms of rationality in the absence of any emotions. Their decisions are based on unbiased memories as well as correct forecasts.

Chapter 2 describes the BRA in a very accessible and self-contained way so that it constitutes a neat foundation for the following chapters. Rationality in terms of the BRA approach is then defined with respect to situations made under certainty, under risk and uncertainty as well as with respect to decisions concerning future outcomes and strategic interaction. Having defined rationality with respect to these different types of choices, the authors continue to define what rationality is not (Chap. 2.8). They are right to point out that rationality neither assumes purely selfish preferences nor does it require the absence of any emotion. While the former is well-known, the latter may be debatable. Dhami and Sunstein take the stand that an individual's preferences may depend on emotional states such as guilt and shame and that preference relations may differ depending on the emotional state one is in when making decisions. It is necessary to be precise about this. Other-regarding preferences and emotions play an important role in Behavioral Economics, but this does not imply that they should be considered as non-rational behavior. The second part of Chap. 2 then provides overwhelming evidence that people make decisions not in line with the

✉ Ronnie Schöb
Ronnie.Schoeb@fu-berlin.de

1 Freie Universität Berlin, Berlin, Germany

BRA. Evidence includes phenomena such as preference reversals, limited attention, overconfidence, nonlinear probability weighting, reference dependence and hyperbolic discounting. The appendix offers a concise formal treatment of decision theory, which turns out to be very useful for teaching purposes.

The third chapter then compares the BRA with alternative approaches. It first discusses the implications of dynamic optimization and then compares the optimization approach with the heuristics one. Although the example of microfinance is discussed in more detail than required to make their point, the authors nevertheless provide a very elaborate discussion of the debate, culminating in a review of the debate between Harstad and Selten and Crawdford and Rabin that appeared in the Journal of Economic Literature. Of particular importance here is the discussion about the endogeneity of preferences and beliefs. For the BRA it is crucial to assume stable preferences while only allowing to update beliefs (according to Bayes' law). But preferences may not be fixed, they may change with experiences and may be context-dependent (or may even be formed ad hoc).

The next chapter tackles a wide variety of behavioral models of heuristic-based choices and lays the ground for confronting the heuristics and biases research program initiated by the pioneering work by Amos Tversky and Daniel Kahneman (Chap. 5) with its critics, in particular by Gerd Gigerenzer and his coauthors (presented in Chap. 6). The overall conclusion drawn from this discussion is that the heuristics and biases research program "was able to demonstrate that the BRA in economics is not tenable, not even in an as-if-sense." (p. 265). Having said this, the question of policy implications arises.

The second part, starting with Chap. 7, consequently turns to the policy implications of bounded rationality. Under the heading "Philosophical Foundations", the authors lay down their normative framework of behavioral welfare economics. They argue that despite the many empirical behavioral findings that people do not behave according to the BRA, policymakers should adopt a working presumption in favor of respect for people's self-regarding choices. However, they add one crucial modification: Politicians should only respect people's individual choices if those choices are based on adequate information and furthermore are sufficiently free from behavioral biases. But do politicians or experts really know whether people are adequately informed or sufficiently free from behavioral biases or identify areas in which this is not the case? This is anything but clear. As long as we do not fully understand apparently inconsistent behavior, we cannot make coherent statements about welfare. Dhami and Sunstein implicitly assume that we have gained a clear understanding of potential biases. It is true, on the one-hand side, that many decisions are inconsistent (e.g. if we consider preference reversals) but on the other-hand side, it is not clear in which direction the bias goes, in particular if choices are made on ad hoc constructed preferences or a choice-set dependent in cases where we have only limited information about the whole choice set. And even if we would be able to determine what the "true" preferences are compared to the assumed preferences on which decisions are based, this may not justify policy intervention. Individuals' well-being may not only depend on outcomes but also on the way in which theses outcomes are achieved, i.e., welfare may also crucially depend on procedural utility. If this is the case, people might feel humiliated and disrespected if an authority restricts their choice set solely to prevent them from not acting in their apparent own best interest. If there is one weakness about the book, it is the fact that the authors have devoted too little room to the normative challenges arising from the positive analysis of bounded rationality. As they call for significant departures from standard economic approaches to welfare analysis, one should be more explicit to what extent such a departure not only rests on the insights from the positive analysis of bounded rationality but also on the implicit modification of the normative foundation. For instance, in the BRA, aspects such as self-determination, autonomy or the role of procedural utility do not matter, because if we are interested in rational people's own best interest, we have no reason to disrespect self-determination and autonomy and do not have to worry about procedural utility. If people face bounded rationality, however, we have to deal with conflicts between different norms, and this requires us to make these conflicts transparent and to take a clear normative stance. To some extent, the authors do so, but some of the normative foundations remain implicit.

Dahmi and Sunstein justify paternalism with respect to indirect judgments and call for humility with respect to direct judgments (ultimate goals). But even with respect to direct judgments they make an argument for policy intervention: "But what if the analyst has an actual data, suggesting that people's direct judgment produces large welfare losses …? What if the analyst has information about what people are likely to like, and what if that information suggests that people's ex ante predictions are incorrect? What if the analyst has an account of what makes for a good or fulfilling life, and not simply an "opinion" …?" (p. 351). They argue in favor of paternalistic interventions, but what about the alternative of giving individuals information about their potentially bounded rationality and leaving it to them to make sense of it?

Given the normative foundations outlined in Chap. 7, the last three chapters deal with different forms of paternalism. Optimal taxation as a form of hard paternalism is discussed in Chap. 8 by using the model of multiple selves to show the similarity of sin taxes to Pigouvian taxes, and the concept of limited attention to address the impact of salience on both the deadweight loss of taxation and on the tax incidence. To what extent do these approaches justify significant departures from standard economic approaches? While the classical public economics approach derives the justification for policy intervention only from the harm principle (and redistributive concerns), the behavioral public economics approach, if it wants to promote such policies, needs to apply a very particular concept of welfare, assuming that people are not responsible for their future selves and democratic systems need not commit to political transparency and accountability when they tax their citizens to increase welfare. Chapter 9 and 10 turn to the soft paternalism in form of the libertarian paternalism and show how it works in practice. Here, indeed, the normative challenges are much smaller.

The book covers a wide range of aspects concerning bounded rationality and thus offers a great compendium about recent developments in behavioral economics and psychology. It thus paves the way for a more intense discussion of how public economics can build on insights about the bounded rationality of individuals. The normative implications, however, are discussed only in passing, but, given the authors' expertise, I hope that they will give them more space in a next edition, which this book definitely deserves.

Funding Open Access funding enabled and organized by Projekt DEAL.