REVIEWS

Evolution and Rationality: Decisions, Co-operation and Strategic Behaviour,
  Samir Okasha and Ken Binmore (eds.). Cambridge University Press,
  2012, x + 281 pages.

Okasha and Binmore have put together a very intriguing volume that
aims to promote the interconnections of Darwinian evolution and models
of rational choice. The volume's main virtue is also its primary flaw:
the eleven contributed chapters come from a variety of perspectives,
which simultaneously give us a fairly rich sense of the state of the art
in this intersection of fields, and also make it clear that many of these
perspectives are only minimally engaging with each other. The challenge
for the reader, made easier by the editors' thoughtful introduction, is to
find the common threads among the essays.

What I take to be the most substantive contribution of the volume
is the amount of methodological reflection seen across a number of the
papers. Interdisciplinary endeavours often have to be more sensitive
to methodological questions precisely because different fields have
different preferred methods and standards of rigour. The beginning of
this conversation takes place with Alisdair Houston's essay examining
'economic' versus 'biological' accounts of rationality, primarily through
the lens of apparent failures of transitivity. Interestingly, Houston argues
that apparent transitivity failures appear only because we model the
situation incorrectly. If we think that we are modelling individual choices,
then transitivity problems arise. But if we model selection between
strategies for making choices, then those problems disappear. Rationality
is preserved, but only by supposing that the rational choice operates on a
different level than we tend to assume.

This question about what the right model is for explaining choices
provides perhaps the most interesting theme in the volume. We see a

425

similar question posed by Brighton and Gigerenzer, who ask whether the rational actor model is appropriate in 'large' worlds of more or less unbounded choice, or only in 'small' worlds where choice sets are restricted in size. In particular, they argue that large worlds come with many sources of uncertainty that make determining their structure difficult, and this uncertainty makes it difficult to use standard optimization procedures that work so well in small worlds, where the structure of the choice environment is known a priori. This then leads them to conclude that large worlds must be treated differently than small worlds in how we model optimal choice-making.

Wolport and Jamison continue with the question of whether we are using the correct models when we model choices, by furthering a Schelling-style 'rational irrationality' model. In particular, they argue that we play 'persona games' which can help explain cooperation in games such as the Prisoner's Dilemma. Rather than making strategic choices at the level of picking strategies in a PD, we instead pick personas – ways we want to be perceived by others – and those personas do the selection at the object level strategic situation. The authors argue that this approach allows us to not only explain the high levels of cooperation seen in mixed motive games, but can also provide a generalized solution for pre-commitments. With a similar goal, but different methods, Natalie Gold argues that we can explain cooperative outcomes not in terms of personas, but in terms of team reasoning. We come to identify our goals with team goals, and so the appropriate unit of analysis shifts away from individual decision-makers and toward team decisions. Standard individualistic game theory just becomes a special case of team reasoning.

The theme of whether we have modelled our choice situation correctly is deeply important. As these papers illustrate, the consequences are philosophically serious: our normative standards, including what we count as rational, are up for grabs. But it is not only our normative standards for choice-making that are called into question here: each alternative model also changes our descriptive understanding of the modelled phenomena, through offering different explanatory pictures. Different explanations in turn affect our ability to predict behaviour, whether it be of animals (as in Houston's primary case) or of humans. The chapters themselves differ in quality, with the standout paper being the Brighton and Gigerenzer essay, which is extremely well-reasoned, but the combination of all of the chapters on this theme points out the immediate benefits we can realize from cross-disciplinary research. Choosing the basic features of a model structures so much of our reasoning about the modelled phenomenon that it is crucial that we at least occasionally reflect on those choices and investigate whether we have chosen correctly, or whether a model that has proven fruitful in one area can be extended to new ones.

A second theme of the essays of this book is much more scattershot: integrating our knowledge of evolution to investigate the Darwinian basis of our behaviour. Reinforcing the feeling that this theme is not quite so organized is the opening chapter by Peter Hammerstein. The chapter investigates how our biological understanding can be brought into our models of rationality. However, the chapter feels a bit stilted: the narrative does not flow from one section to another, even if each section is independently interesting and informative. I came into the book already convinced that evolutionary theory and rational choice theory share important commonalities, and the chapter reinforced this belief, but insofar as this chapter is meant to support an evolutionary interpretation of key concepts in game theory, it could have been done with a clearer narrative, especially to convince those who start out more sceptical.

Mouden, Burton-Chellew, Gardner and West follow in this theme by arguing that our preferences are shaped by evolution such that they maximize inclusive fitness. This is the first of several chapters that try and wrestle with the relationship between fitness and preferences. It is also the least successful. The authors offer fairly unconvincing adaptationist arguments that in the end feel like epicycles on top of epicycles. More convincing is the chapter by Bernignhaus, Guth and Kliemt, which looks how to incorporate intentional and evolutionary motivations into a unified theory, through an examination of the Trust Game. They explore how the 'push' of evolution and the 'pull' of intentional action can be combined to gain a better explanation of our behaviour. It is a clear account that is integrative rather than reductive. Another thoughtful paper in this line of reasoning is by Jack Vromen. Through a very interesting discussion of the relationship between strong reciprocity and reciprocal altruism, his chapter offers a careful differentiation between psychological and evolutionary motivations, and uses that distinction to consider proximate versus ultimate explanations in strategic behaviour. This is a sophisticated look at the evolution of preferences, which many would profit from reading.

The final two essays in this theme are also the last two essays in the book. They are also the most systematic attempts to think about the relationship between evolution and rationality, and the methodological consequences of doing so. The first, by Herbert Gintis, argues for a unified behavioural science, integrating evolutionary biology, decision science, economics, and related fields, with game theory and correlated equilibria at its core. This is engaging, and offers a sweeping vision for how we should conduct the work that is the topic of this volume. It is also more or less a *précis* of his 2009 book, *The Bounds of Reason*. Though readers who wish to see Gintis' full picture might simply want to read that book, this volume ought to remain of some interest, since the last essay in the book, by Kim Sterelny, engages with Gintis' work

directly. The fact of engagement is nice to see, as outside of this final essay, and Gintis' discussion of some of Gigerenzer's earlier work, there is not much of a conversation between the authors. Sterelny looks at three different modelling approaches for human rationality, devoting much of his time to Gintis' approach, and argues that we are forced towards model pluralism, in part because humans faced distinct sets of conditions during different eras. Holocene humans and Pleistocene humans require different models of agency. Sterelny recognizes that this approach is perhaps messier, but argues that we ought to be sensitive to how our agency changed in response to different ecological situations. Sterelny's essay argues first that we need to be much more sensitive to this in our modelling practice, making a thoughtful contribution to the methodological discussion present throughout this volume, but then it shifts to argue that as behavioural ecology shifted, so did the connection between biological fitness and economic utility. This is a much more sophisticated account than what we found earlier in the volume. Sterelny's essay is a good one to finish the volume, as it ably demonstrates how careful considerations of methodology can lead to substantive conclusions about the nature of the relationship between evolution and economics.

Hutteger and Zollman offer an essay that doesn't quite fit in with either theme that I have identified, but present a powerful methodological argument that is worth highlighting. They argue against what they call 'ESS Methodology' – the tendency to only examine Evolutionary Stable Strategies (ESS) of an evolutionary game to gain an understanding of the game in question. The authors connect the concept of an ESS to the broader literature on refinements to the Nash Equilibrium in economics, and then argue that while ESS has several virtues conceptually, ESS Methodology, which only looks at an evolutionary game's ESSs for an understanding of the game, offers a very limited view of the population's behaviour. In particular, it can only tell us about equilibrium analysis – what happens if we are at or near a game equilibrium. We do not learn anything about the dynamics of the population through time, nor how likely we are to end up at one ESS versus another, or some other non-ESS state. Furthermore, ESSs are not the only attractors in the dynamics of a game, even if we limit ourselves to models that employ the replicator dynamics. Hutteger and Zollman use previous work they have done on the Sir Philip Sidney (SPS) game, which is a signalling game that allows for conflicts of interest between the sender and receiver. They demonstrate that there are other evolutionarily significant equilibria, notably a pooling equilibrium and a polymorphic equilibrium, neither of which are ESS, but both are Liapunov stable, implying that they do serve as attractors in the overall population dynamics of the game when they are present. These are examples of the more general phenomenon of Neutrally Stable Strategies

(NSS) in games. NSSs are less strict than ESSs, but they are Liapunov stable, and can thus play a major role in the game's dynamics. Hutteger and Zollman point out that we must move toward a more dynamics-focused methodology to learn about evolutionary games. This is clearly right.

I close my discussion with the Hutteger and Zollman essay in part because it offers an opportunity to comment on the limits of the methods discussed in this volume. Hutteger and Zollman point out that equilibrium analysis is not good enough if we want to understand the dynamics of evolutionary games. Brighton and Gigerenzer point out that small-world optimality modelling is not appropriate for larger worlds where we have more sources of uncertainty. More generally, the first theme of the book examines whether we model choice situations with the right frameworks, and the second theme investigates how we should integrate biological reasoning into our economic modelling. These are deep and important methodological insights – modellers would be well-served by paying attention to them. But it is worth noting that methodological critiques can go considerably further, particularly when we look at evolutionary modelling. For example, none of these approaches can say anything about basic features of real biological evolution or economic choice-making such as innovation or novelty. Game theory requires a fixed game: a set of strategies and their corresponding payoffs that does not change over time. We can, and have, learned many interesting things about the world by limiting ourselves to this framework. But because this limitation exists in the modelling framework that so many of us employ, we ought to be sensitive to how this matters methodologically.

Most basically, this limitation means that we cannot rigorously describe novelty. When we have mutations in an evolutionary game, they are always just changes to a different, pre-existing, strategy. But part of what makes the real evolutionary process so amazing is the emergence of new forms. It is why Darwin finished the *Origin of Species* by saying '... from so simple a beginning endless forms most beautiful and most wonderful have been, and are being evolved'. Evolutionary game theory cannot offer us any insight into this. Nor can our standard models of rationality offer anything about how we come to see new choices as being possible. This is not merely a lament that our tools lack the power to allow us to examine extremely complex phenomena like novelty. Fixed strategy sets can be complicated enough – dynamic strategy sets would be enormously challenging. Rather, it should also make us less certain that we are modelling these simpler cases as well as we think we are. Think of stability analysis. An ESS is only an ESS with respect to the other strategies in the defined game. If we have incorrectly identified the strategy set, and simply add a new strategy to the game, what was previously an ESS may not be one anymore. Previous invasion analysis is

useless. This is precisely why invasive species are problematic – they upset previously established equilibria by introducing a new strategy into the mix. This is not just a problem for biological models – economic models have the same basic methodological weakness. Evolutionary game theory and evolutionary reasoning more generally are often offered as providing more secure foundations for our models of decision making and strategic behaviour. But it is worth considering that while evolution itself might provide such a secure foundation, our models of it are not there yet.

A volume like *Evolution and Rationality* is ideal for getting up to speed on the issues at the overlap of economic rationality and evolutionary theory. As with any edited volume, there is a variation in quality across articles, and at least this reader would have liked to see a bit more direct engagement amongst the authors, particularly where they were writing on more or less the same topic. The different perspectives both enrich the picture offered by the editors, and contribute to a feeling that the book lacks a central vision about how we ought to understand the overlap of evolution and rationality. Combined with the variance in quality between the essays, the differences in methodology and topic across chapters make for a very uneven reading experience. I would have organized the chapters a bit differently, and introduced sections with short introductions by the editors to increase cross-chapter coherence. As it stands, the book feels like it is a particularly high-powered special issue of a journal. By its nature, the book is not looking to stake out much new ground, but rather give a curated sample of what is already there.

Ryan Muldoon
University of Pennsylvania, USA

The Genesis and Ethos of the Market, Luigino Bruni. Palgrave Macmillan, 2012, 240 pages.

Both capitalism and modern political economy rest on the separation of economics from ethics, which in turn can be traced to a number of shifts within philosophy and theology – notably the move away from practices of reciprocity and the common good towards the sole pursuit of individual freedom and self-interest. In his latest book, Luigino Bruni provides a compelling critique of capitalist markets and an alternative vision that fuses Aristotelian–Thomist virtue ethics with the Renaissance and Neapolitan Enlightenment tradition of 'civil economy'.

The book develops three broad yet closely intertwined theses. First, that Greco-Roman Antiquity and the Christian Middle Ages invented models of civil life that transcended tribalism and political absolutism but produced sacral communities wherein the power and privilege of