

Neo-Samuelsonian methodology, normative economics, and the quantitative intentional stance

Don Ross

**School of Society, Politics, and Ethics
University College Cork**

**School of Economics and Research Unit in Behavioural Economics and
Neuroeconomics (RUBEN)
University of Cape Town**

**Center for the Economic Analysis of Risk (CEAR), Robinson College of Business
Georgia State University**

don.ross931@gmail.com

Abstract

Wade Hands (2013) critically consolidated a new and growing approach to revealed preference interpretations and methods, which he called “Contemporary Revealed Preference Theory” (CRPT). He recognised that CRPT is folded into a more comprehensive philosophy of economics due to Don Ross, which Ross dubs “Neo-Samuelsonian Philosophy of Economics” (NSEP). Hands calls this “the elephant in the room” where his criticisms of CRPT as an account of normative economics are concerned. I address Hands’s criticisms of CRPT using the full resources of NSPE. This leads to substantial reconsideration of normative economics, with respect to both assessments of general rationality in economic agents, and emphasis on welfare improvement in applied policy work. Main conclusions are that (1) economists are not properly in the business of assessing general rationality, a topic best left to a philosophical tradition descended from Aristotle; (2) the borrowing of theoretical structures from the foundations of microeconomics in the project of philosophical decision theory encourages the idea that there should be a rigorous bridge between economists’ interest in technical choice consistency and philosophers’ interest in general rationality, but NSPE implies that this approach to bridge-building is misguided; (3) economics is a policy science, but the policy domains to which it aims to be relevant are public and corporate, not personal; (4) NSPE provides clearer insight than alternative philosophies of economics as to why economists concentrate on welfare, rather than well-being, as their primary normative target.

JEL codes: A11, A12, A13, B40, D01, D60

1. Introduction

Hands (2013) was the first prominent commentator on economic methodology to explicitly signal notice of a new entrant among the competing philosophies of core economic methods, which he called “contemporary revealed preference theory” (CRPT). CRPT, on Hands’s reconstruction, differs from ‘original’ Samuelson-Houthakker revealed preference theory (RPT) in three main respects:

- (i) it applies over finite, measured, choice sets, rather than infinite, necessarily hypothetical ones; technically, it implements Afriat's (1967) Generalised Axiom of Revealed Preference (GARP) instead of Samuelson's Weak Axiom of Revealed Preference (WARP) or Houthakker's Strong Axiom of Revealed Preference (SARP);
- (ii) in interpreting utility functions (and preferences) as summaries of observed instances of an agent's choice behaviour, it denies that utility functions (or preferences) causally explain choices or other aspects of behaviour;
- (iii) in allowing that economic agency arises at scales other than that of whole individual people, it extends the reach of RPT to apply to any entities that are behaviourally sensitive to incentives; thus it extends the domain of microeconomic analysis to include non-human organisms, coordinated coalitions of organisms, and perhaps internal functional modules within nervous systems.

As Hands makes clear in his careful presentation, CRPT is not a 'paradigm shift' *against* earlier interpretations of RPT, but the product of technological evolution from them. Samuelson interpreted RPT compatibly with (ii). Numerous economists following Afriat's innovation applied (i). Followers of Gary Becker, though often only loosely committed to RPT, set precedents for (iii). What distinguishes CRPT, on Hands's story, is its fusion of (i) – (iii) so as to attempt to assemble a relatively full-bodied philosophy of economics.

Hands introduces CRPT by reference to three sources: Bernheim and Rangel (2008), Binmore (2009a, 2009b), and Gul and Pesendorfer (2008). Later he adds reference to Ross (2000, 2005, 2008, 2011) as being "the most philosophically sophisticated defense of CRPT" (Hands 2013 pp. 1104-1105). In work published subsequent to Hands's paper, particularly Ross (2014), I rounded off corners and added brushwork to this structure, and also motivated a more descriptive name for it: 'neo-Samuelsonian philosophy of economics' (NSPE). This is the version of CRPT that will be in play in what follows. Consequently, except when quoting or closely paraphrasing Hands, I will drop references to 'CRPT', a label that is bound to become increasingly awkward as the years accumulate upgrades and refinements to the view.

NSPE adds, as Hands notes, two main supporting elements to his characterisation of CRPT.

First, and most importantly, it interprets preferences and beliefs – the propositional attitudes (PAs), as Fodor (1978) nicely called them – in accordance with the now-dominant view in the philosophy of mind known as 'externalism' (Burge 1986; Dennett 1987; McClamrock 1995; Clark 1997; Bogdan 1997, 2000, 2009, 2010, 2013). Externalism comes in weak and strong forms. In the weak form due to Putnam (1975) and now endorsed by almost all philosophers, it recognises that truth conditions on beliefs depend on semantic facts of which the bearers of the beliefs in question might not be aware. As expounded in Ross (2005), what is relevant to NSPE is the more general cognitive externalism

according to which PAs do not designate theoretically constructed internal representational structures in brains, but descriptions of observed and forecast patterns in intentional behaviour, including verbal behaviour. People understand and predict one another's behaviour by adopting what Dennett (1987) calls *the intentional stance* toward one another: ascribing mutually coherent sets of beliefs and preferences that make sense of behaviour as aiming at goals that are conditional on perceptions about how the world is.

Intentional-stance taking is a skill that children are required to learn to participate in human society. This includes learning to take the intentional stance toward their *own* behaviour. This self-referential construction of networks of PAs creates distinctive individual personhood that allows for recruitment into shared projects with others. The view that PAs are constructed in shared ascription rather than discovered as expressions of representations in individual brains is sometimes mistakenly thought to assign PAs the status of mythical objects. However, as Dennett has stressed, being 'virtual' objects of social negotiation is a way of being real, not a way of being fictional. Money is similarly virtual, but its reality is self-evidently robust. Dennett (1991a) proposed a general account, using resources from computational and information theory, for distinguishing 'real patterns', including virtual ones, from descriptions of patterns that are used merely for convenient shorthand. Ross (2000) and Ladyman and Ross (2007) repaired some technical deficiencies in Dennett's account, and LR's has become the standard theory of real patterns cited for applications and criticism in the philosophy of science.

As Ross (2014) explains, this view of PAs amounts to bringing cognitive science into alignment with Afriat-style RPT, the opposite direction of reconciliation to that favoured by behavioural economists in the broad camp of Thaler (1992) and Camerer et al (2005). It also makes Hands's feature (iii) of CRPT natural: it follows from externalism that an agent need not have a nervous system to have (real) beliefs and preferences, so institutions can be fully-fledged, and not merely metaphorical, economic agents.

Second, NSPE takes a relaxed attitude to conformity of revealed preferences with axioms of either GARP or expected utility theory (EUT). If RPT is to be applied to actual rather than normatively idealised PAs, it must allow for the fact that the latter are *noisy* patterns (Dennett 1991a). Axioms provide crucial baseline specifications to be adjusted in the face of empirical choice data, but the target form of economic theory is structural and statistical; see Bourgeois-Gironde (2020) for examples of such instrumental use of axioms, and Wilcox (2008, 2021) for deeper challenges to consistent application that call for conceptual refinements. For example, as Wilcox argues, relative risk aversion must be understood as relative probabilistic risk aversion, which is crucially *not* equivalent to deterministic risk aversion with error. The sources of noise in pattern identification and specification, given externalism about PAs, arise partly from stochastic psychological mechanisms, but may equally be generated by fluctuating social dynamics, to which the very meanings of PAs are sensitive. PAs

are not, in general, *inferred from* observed behaviour through processes cognitive scientists call ‘mindreading’ (Nichols and Stich 2003). Zawidzki (2013) reviews evidence against the hypothesis that such inferences are frequently actual or indeed even typically possible in real-time social interactions.¹ Rather, Zawidzki argues, as fundamentally social constructs PAs successfully describe statistical tendencies in behaviour because they are instruments by which behaviour is *regulated* through processes of *mindshaping*. Ross and Stirling (2021) provide a formal characterisation of mindshaping in game-theoretic terms, which Ross et al (2021) extend to contexts involving risk.

Adjusting the focus from CRPT as Hands frames it to the more comprehensive NSPE implies removal of Gul and Pesendorfer (2008) (GP) from the corpus of representative literature. Their blunt denial that any results from psychology can be relevant to any economic model specification is implausible dogmatism on its surface, and inconsistent with their own methodological examples (e.g. Gul and Pesendorfer 2001), which incorporate the assumption that people have well defined *private* preferences over succumbing to short-term temptations. Since this assumption is a folk theoretical myth according to some externalist cognitive scientists, it is not innocent of psychology as per GP’s official methodology. A similar objection applies to recent work of Bernheim (2016), which relies on the assumption that “true” preference content can in principle be identified with outputs of (deliberative or latent) internal mental processes. In the economics literature, Infante et al (2016) and Sugden (2018) have emphasised the absence of support from cognitive science for this assumption. Therefore, in what follows, my core reference texts for methodological views consistent with NSPE will be Smith (2008), Binmore (2009a), and Ross (2014). NSPE as promoted in Binmore (2009a) and Ross (2014) explicitly incorporates preferred strategies for econometric estimations of structural utility functions. These strategies are illustrated, and accompanied by methodological reflections, in Andersen et al (2008), Harrison and Rutström (2008), Harrison and Ng (2016), and Harrison and Ross (2018).

Hands argues that CRPT avoids generic challenges raised against earlier versions of RPT. He considers Hausman’s (1992, 2000, 2008, 2012) and Sen’s (1973 and elsewhere) linked family of arguments to the effect that preferences cannot explain behaviour independently of beliefs. These arguments begin with analysis of folk-psychological explanation, and then rely on the assumption that economic consumer theory is a formalisation or regimentation of folk psychology. CRPT/NSPE is untroubled by this kind of objection, because it disavows

¹ Chater (2018) argues that *all* outputs of mindreading efforts are errors, because the view that there is anything to be read is false: all that is internal, he argues, are brain states, and these have no stable interpretations in terms of propositional content. This is consistent with Dennett’s (1991b) account of consciousness, though not, I would argue, with Dennett’s account of PAs. Chater, in my view, exaggerates the empirical instability of PAs because he focuses on psychological experiments that isolate subjects from social scaffolds, the primary force for their stabilisation.

straightforward interpretation of economic theory in terms of folk psychology. Although externalism about PAs does not eliminate preferences or beliefs from scientific ontology, in denying that these denote internal states or processes it breaks with the aspect of folk psychology on which Hausman's arguments rely. Hausman (1995) contends that unless preferences are understood as causes of choice, economists are forced to fall back on instrumentalist philosophy of their science. NSPE directly answers this objection: the core original motivation of Dennett's Real Patterns ontology that NSPE refines and incorporates is to explain how PAs can be actual, virtual, elements of the cognitive-behavioural nexus without being taken to be internal states or processes.

Hands goes on to argue, however, that the strategy by which CRPT escapes allegations of incapacity to explain behaviour succeeds by surrendering its power to normatively favour some choices over others. Both Hausman's and Sen's criticisms anticipated this problem, Hands says. He discusses two senses in which economists often proffer normative guidance: on which items in a set of policy options would make agents better off, and on which patterns of choice allow agents to be deemed practically rational. Though Hands's discussion involves various strands and sub-arguments, it can be boiled down without too much loss of accuracy as follows. Both welfare and rationality assessments as economists typically perform them inherit their plausibility from the folk-psychological assumption that agents should do what they believe is most likely to best satisfy their preferences (given constraints). Thus in breaking with folk-psychological foundations, CRPT/NSPE abandons the source of normative force, in general, in economic analysis. This implies a *reductio* against CRPT/NSPE because economists are committed to doing policy science.

In what follows I will argue that Hands's critique reveals an important and very broad fact about the normative reach of economics: that ambitions are often attributed to it that it cannot and should not aspire to achieve. Welfare is not equivalent to human well-being, and even if economists can often successfully identify welfare consequences of policies (as, I will argue, they can and do), agents should only follow their advice *to the extent that* welfare is what the agents in question care about. Fortunately for the relevance of economists' interest in policy, a dominant institutional practice in modern societies is for people to be explicitly hired by principals to implement welfare improvements for agents. As for rationality *in general*, I will argue that this is not an idea with which economists should professionally concern themselves at all. I will grant a sense in which philosophers can usefully worry about rationality in general, but the sense in question does not incorporate rational *choice* as axiomatised in the powerful tradition of Bernoulli, van Neumann, Morgenstern, and Savage.

I referred above to ambitions "attributed to" economics. Disciplines, of course, only 'have ambitions' metaphorically. Economists have ambitions literally, which vary from economist to economist. Toward the end of the essay I will suggest – though leaving the work of demonstration to future occasions – that economists' ambitions with respect to normative authority have tended systematically to be

inflated by philosophers' and other critics' interpretations, thus creating balloons ripe for popping. It is precisely one of the attractions of NSPE that it can help economists spot and resist such inflationary pressure.

Hands musters an additional criticism of CRPT/NSPE, questioning its claim to generality as a philosophy of economics. In particular, he raises doubt that a methodology for consumer choice theory should reasonably be extended across other areas of economics. In pressing this criticism, I will argue, Hands fails to fully appreciate the force of his own other arguments. These generated the conclusion that microeconomic choice theory has been taken to commit over-reach. The over-reach in question is not with respect to modeling choice itself, wherever it is found; rather, it is in extension to some *personal* policy domains, which typically implicate considerations from beyond economists' zone of focus. But *all* of microeconomics concerns choice: incentivised choice is indeed its defining subject matter. The basis for generalising NSPE across all of microeconomics emerges most clearly precisely when sweeping normative aspirations are tempered.

The paper is structured as follows. In Section 2 I argue that economists should not interpret their models of choice as aspects of the philosophical project of conceptualising general rationality. The conclusions of this discussion feed into Section 3, which addresses the limits and value of welfare analysis given NSPE. The brief Section 4 concerns the scope of NSPE in response to Hands's criticism on that front. Section 5 concludes.

2. *Rationality in general*

I will defend the following claims:

- (i) Rationality, as an everyday normative designation, has multiple potential interpretations that are disambiguated for specific application by contexts.
- (ii) The idea of rationality, and irrationality, as *general* potential characteristics of people, because it implies transcendence of contexts, is not coherent except within a carefully developed, venerable, and valuable philosophical tradition that resists presentation in other than discursive and narrative form.
- (iii) What economists model as rationality is a specific, very useful, technical idea that conforms only to occasional applications of the everyday notion, and has no very direct relationship to the general rationality of the philosophers.
- (iv) Normative philosophical decision theory is a confused enterprise that systematically confuses (ii) and (iii) above.

Culturally, 'rationality' is as thick and complex a concept as 'beauty' and 'moral goodness'. As with these other thick concepts, its normal use depends on clarity of contexts. If I can show that you believe that some proposition is true only because you wish it were true, then I can call you irrational with respect to that

belief.² If I can show that you chose some action only to increase the probability of some outcome that your action in fact makes less likely, but you persist despite evidently understanding my demonstration, I can call you irrational with respect to that combination of ends and means.

If, then, I say that you are ‘generally’ irrational, perhaps I merely mean that you are unusually statistically likely to be soundly diagnosed as irrational in a range of actual and projected specific, contextualised, instances. But this apparently innocuous extension is problematic.

Objections of Sen’s type systematically arise. In the first example above, the critical force of labeling wishful beliefs as irrational is very limited, because such beliefs often make people happier or more productive. In the second example, what sets people up for valid labeling as irrational with respect to practically self-defeating actions are narrowly instrumental and strongly partitioned motivations for the actions in question. The large proportion of human actions that do not have such characteristics, and with which actions that *do* have these characteristics are usually entangled, give examples of this kind very limited leverage as grounds for attributing general irrationality.

Confident diagnoses of local irrationality thus depend on the availability of conditions that make some choice patterns or beliefs relatively isolable from an agent’s wider cognitive and behavioural ecology. This is what we should expect, given that the intentional stance depends on *assuming* general coherence of belief and preference. Thus local departures can be identified relative to that frame, but only *as* local. If we think of the intentional-stance account of PAs as a form of irrealism about them, then, since this strategy is logically similar to the variety of RPT adopted under NSPE, then we invite Hausman’s (1995) charge of instrumentalism against CRPT. But this is where the combination of the realist interpretation of the intentional stance (again, see Ladyman and Ross 2007) and externalism about PAs bites: it is an empirical matter which entities can be constructed as agents, and the crucial contingent requirement is that describing them by reference to *generally* coherent PAs leads to successful prediction and explanation of their behaviour.

Attempting to motivate a diagnosis of general irrationality by piling up instances of local irrationality confronts the problem that there is no general counting protocol for ‘instances’. Consider the special kinds of potential local irrationalities with which economists concern themselves. The more finely we distinguish choice options, the more inconsistencies we will identify, because agents are limited in their capacities to notice and respond to informational discriminations. Empirical and analytical work is required, with respect to any given case of inconsistency, to determine whether the case is best modeled as an instance of

² I assume an externalist understanding of the belief, so this attribution does not commit me to the hypothesis that your brain performed some unsound computation. My judgment is that patterns in your behaviour, including what you say, are summarised by saying that you have this (false) belief.

subject error or unexpectedly wide indifference over prospects. And attributing error requires, in the scientific setting, specifying a *kind* of error: perceptual discrimination (Fechner) error implies different out-of-sample behaviour from errors of representational framing or errors of implementation (trembling hands). To the economist this problem calls not for exercises in interpretative hermeneutics, as apparent error requires from the qualitative intentional stance demands of the historian or literary critic, but for theoretically guided experimental design that is well integrated with good econometrics. The economist's tools allow her to operate a *quantitative intentional stance* (Alekseev et al 2018).

Consider an example as applied to a single consumer. Suppose that, after becoming richer, someone becomes less sensitive to price changes in a regularly purchased item such as toilet paper. Is she rationally avoiding spending her more valuable marginal time on attending to toilet paper options, or does she believe, on the basis of no good evidence, that her habitual brand is superior to an equally accessible alternative, so is worth a premium? Or perhaps she's correlation averse with respect to intertemporal risk, so her maintenance of a brand-centered habit is minimisation of the prospect of a negative utility shock when she takes home an untested substitute. Economists know better than to necessarily believe her when she asserts a belief that "Huggies are the best," if her report is not controlled by manipulation of salient incentives. We might be able to experimentally demonstrate, to a scientific standard of evidence, that she is more rational, with respect to toilet paper, or perhaps household staples in general, than her unincentivised survey responses suggest. The results of such studies, of representative samples of consumers, are worth real money to toilet paper manufacturers and to supermarket managers deciding how to stock their shelves.

Note that although the specific data-generating process we identify matters for these purposes, it does *not* matter to them whether we stigmatise some but not other processes as 'irrational'. One might nevertheless suppose that *some* responses to price changes that we can hypothetically imagine would imply not only local but more serious and general rationality. Suppose, in the example above, that the subject revealed that she would still choose Huggies if the price were raised to \$1,000 per pack. This might license real doubts about her sanity. It is important that our methodological standard for regarding such a preference as revealed is high: behaviour in the lab that implied such a preference, if the actual experimental stakes were much smaller, would lead us to doubt our experimental design or our theoretical specification.

But doesn't this hypothetical concession suggest that contextualised consumption research can tell us something about general rationality after all? It seems that we think we know that no generally rational person is *completely* insensitive to prices. In fact I think that this is true, but it is a truth that is of no significance for policy purposes because the preference we imagine being revealed would never actually be observed. Truths of the kind at issue here matter to theory and modeling. For a more grounded example, economists learned something important about the

relationship between risk and utility when they explained why St Petersburg lotteries aren't marketed outside of economics labs. But our relevant question here is: should we interpret the resulting knowledge about theory and modeling as knowledge about general rationality? A sufficiently risk-loving person might buy a St Petersburg lottery, if there were enough such people to induce their supply. There is no evident basis for pronouncing these hypothetical gamblers 'generally irrational'. But that isn't the point of the St Petersburg thought experiment; its point is to show that a good theory of subjective utility should take account of risk.

Perhaps the following proposition could be defended: it would be a mark of general irrationality for a person to make choices with complete obliviousness to risk (i.e., to choose the lottery with the highest outcome among its prospects no matter what). Again we consider a proposition that has no practical policy significance for economists: no set of real agents economists will ever model includes any risk-oblivious ones, for the same reason they will not include any price-oblivious ones: such agents do not actually exist.³

There is an obvious objection to the claim that price-oblivious and risk-oblivious agents don't exist. It might simply be pointed out that some real *people* – my mother in her final years as a dementia victim, for example – are oblivious to risk as a special case of being oblivious to almost everything. I have argued elsewhere, however (Ross 2012), that economists working from the quantitative intentional stance cannot usefully model such people as economic agents. A natural, non-technical way of explaining why not, which is the language that many economists would actually use, is to say “they are not rational”. True to everyday usage though this is, if an economist says it and intends it strictly, then she is accidentally philosophising, and not doing so well.

According to the Dennettian account of agency that is among the foundations of NSPE, agents just are partially self-regulated systems that respond to changes in incentives. The disappearance of this responsiveness toward the end of my mother's life was thus equivalent to her disappearance from the class of agents. Economic analysis is usefully applicable to agents – to *all* agents. It is not applicable to any non-agents. But then *if* something counts as an agent, the range of technical devices available to economists for recovering real patterns in the agent's choice behaviour have no systematic relationship to the practical motivations that govern folk applications of the rationality concept (or the *useful* philosophical idea of general rationality to which I will turn presently). None of widened indifference bands, nor Fechner error, nor trembling hands are typically associated by economists with irrationality;⁴ whereas hypothesised computational or framing errors are.⁵ This, I am arguing, is an instance of economists harmfully

³ Obliviousness to risk does not, of course, mean risk neutrality, a property I expect to be observed by the financial agents I hire.

⁴ It should be noted that rhetoric here is unstable. For example, Pennesi (2021) refers to a parameter for limiting Fechner error as a “rationality measure”.

⁵ This applies to the champions of ubiquitous framing effects in real people, nudgers.

swallowing some out-of-date philosophy: the ontologies in which “computational processing error” is constructed as a concept are stretched beyond the restricted bases in which they are used by cognitive scientists.⁶ Why is a hand that trembles due to poor nervous system calibration outside cortex not regarded as evidence of irrationality, whereas a hand that trembles due to low tonic dopamine levels in midbrain is taken to impugn rationality? There is a (complicated) explanation for these semantics in the history of science, and in its relationships to folk psychology and metaphysics, but this explanation is not a justification for maintaining a scientifically arbitrary distinction.

My mother’s dreadful disease changed her gradually, not suddenly. So why not say that she lost her agency gradually, and mark this by saying that she became progressively less rational? Again, that is a natural enough thing to say for casual purposes. But as a target of economic analysis, what happened to my mother was that she steadily discriminated less and less among outcomes to which she had earlier responded as imperfect substitutes. Her indifference bands became progressively wider. But at each point in her decline she was highly behaviourally responsive to what she *did* still distinguish with respect to substitutability. She remained an agent until the point, which preceded her death by many months, at which she ceased to implement any preferences at all.⁷

My mother *did* gradually lose a lot of capacities other than agency. For most normative purposes relevant to her carers, and to the range of public policies that governed both this care and the legal rights and obligations that still applied to her (e.g., she still owed taxes), the most important thing she lost was capacity to share general assessments of the significance of her life (or of anything) with others. As she became increasingly unable to discriminate between particular people, and to understand speech, and to identify herself with any past experiences, she ceased to have opinions (in the sense of Dennett 1991c) about how to integrate what was best for her with what others regarded as good, for her and for themselves.

I think that philosophers do something useful in trying to articulate how loss of this kind warrants being regarded as loss of a kind of ‘general rationality,’ in a sense that has deeper intellectual roots than careless folk conceptions, and a far older history than scientific economics. In particular, my mother became decreasingly generally rational in the sense that Aristotle intended, which has been extensively critically refined by philosophers working loosely within a project they inherit from him.

⁶ Kahneman and Tversky took a fateful step when they moved from the purely psychological context of original prospect theory (Kahneman and Tversky 1979) to the economic theory of cumulative prospect theory (Tversky and Kahneman 1992) which was intended to allow for advice about improving rational decisions.

⁷ I do not mean to suggest that my mother was typical of Alzheimer’s patients in this respect. The majority die before their agency vanishes.

This Aristotelian tradition is by its nature discursive and entangled within a broader nexus of humanistic scholarship. It resists brief summary if anything does, but with that warning duly issued I will condense it to tweet length: it studies the deliberate personal and social management of emotions to create preconditions, dependent on good luck, for long-term personal and interpersonal peace. Some (relatively) recent philosophers who have made major contributions to it, whom I mention to help non-specialists gather what I am talking about, are Bernard Williams (1981, 2006), Martha Nussbaum (1997), Valerie Tiberius (2010), and Alan Gibbard (1990). My aim here is to promote just two points about this tradition: that it is semantically appropriate to regard its practitioners as studying general rationality, and that economists who conceive of themselves as students of rationality are interested in something largely unrelated to it.

Philosophers in the Aristotelean tradition are working witness to what many economists will regard as an audacious belief: that it is possible to make discoveries about non-trivial and non-obvious principles and commitments that characterise wise living for people trying to balance social flourishing with freedom of thought, open-ness to experience *and* particular historical-cultural identifications. There are two aspects of this project that make it audacious by reference to economists' standard assumptions about the sovereignty of strictly subjective value judgment.⁸ The first audacious idea is that it is possible to promote *general* propositions intended to be *objectively* true about wise social and personal living, notwithstanding full alertness to the heterogeneity of human personalities and proclivities, without implying any kind of moral tyranny. The second audacious idea is that it is possible for some particular people to be trained experts in such promotion, who may reasonably request deference from non-experts.

These propositions are audacious for a combination of two reasons: first, they seem *prima facie* implausible; second, if they are true then well-trained Aristotelean philosophers assume remarkable responsibility. There is no general analytic argument to be made for the audacious propositions. The only possible proof is in demonstration: read the best exemplars, such as the philosophers I mentioned above. It is possible to be a reasonably well-informed skeptic about the propositions: I think that Karl Popper was one (though not his general intellectual friend Hayek). It is no part of my aim to convince such skeptics. I do, however, seek to persuade economists that *casual* skepticism here is rash.

First, it is mulish to insist that what the Aristotelian philosophers study is not really rationality. The entire point of the project is to gain insights into how a thoughtful person should control and temper her impulses and tendencies to lazy thinking. Hands (2013) points out that CRPT/NSEP takes to its logical conclusion economists' identification of 'rationality' (in their special sense) with choice

⁸ These assumptions do not divide economists from philosophers *in general*. Hume was among the forgers of economists' working professional value system, and there are today at least as many Humean philosophers as broadly Aristotelian ones.

consistency. The Aristotelian philosophers are also interested in consistency, but only *secondarily*, and to the extent that some manifestations of inconsistency undermine principles of wise judgment about general values. Inconsistency with respect to responses to price changes in toilet paper is not among such manifestations. Inconsistency with respect to the value of trying to prevent biodiversity collapse, or of seeking to thwart the ambitions of cruel people, or to looking for reasons to judge people generously, are among such manifestations. I conceded earlier that a hypothetical person who was oblivious to prices or risks would signal general irrationality. But these strange insensitivities would be indicators that they were too cognitively *and* emotionally peculiar to likely be generally rational; they would not in themselves be the *substance* of the general irrationality they suggested. And I don't think that philosophers should be any more *interested* in these outlandish thought experiments than economists should be.

The basic point is that economists have no proper business, unless they mean to develop a sideline in promoting aggressive reductionism about value in denying any of the following claims: there are generally better and worse general ways of forming and implementing opinions on important social and moral questions; “generally better and worse” doesn't grant trumping authority to consistency, or normally refer to any *purely* subjective states of individual opinion holders; and ‘rationality’ in Aristotle's sense is something that people who are emotionally attracted to peacefulness should want to encourage.

Second, economists invite confusion about their very important project, of trying to comparatively quantify opportunity costs of public and private choices, with vigilance about unintended consequences factored into the estimations, if they adopt rhetoric that seems to reduce Aristotelian general rationality to the necessary conditions for identifying and estimating choice functions as utility functions. Few economists intend such reduction, because few economists are aggressively philosophical. However, it is common for them to defend ‘rational’ economic man’ as a normative idealisation. Most see that ‘man’ is a remnant of old blind spots, but fewer see that ‘rational’ is too. It is the semantic trace of a bruised conviction, which was particularly widespread in general mid-20th century culture, that all important normative problems can in principle be transformed into technical ones (Amadae 2003, 2016). The important concept of economic agency is indeed technical (Ross 2012). But insofar as its attribution to a person is a normative complement, it celebrates a form of cognitive and professional achievement, not general wisdom of character.

There is an admirable tradition in the culture of economics that emphasises the modesty of economists' normative ambitions. Keynes (1936) urged economists to refrain from promoting sweeping normative programmes and to adopt the posture of “humble, competent people, on a level with dentists” (p. 373). Keynes is explicit that his intended force of “humble” is to exclude promotion – *or* blockage – of sweeping programmes for reforming human values. This attitude followed the example of one of Keynes's principal mentors, Marshall, who emphasised that

an economist's normative opinions *qua* economist concern “the ordinary business of life”, as contrasted with the deep moral and historicist quandaries assessed by Aristotelian philosophers (Backhouse 2002, pp. 180-181). More recently, Colander and Su (2018) argue for precisely the Keynesian vision of humbleness, stressing that even with respect to considerations of practical efficiency economists should not derive advice for actual people *directly from* theory that applies to economic agents.⁹ Economic theory should not be seen as designed to formally characterise whole people (Ross 2005). That is why, when it is applied to the problems of whole people, the understanding it facilitates must be blended with knowledge of history, political psychology, demographics, analytical sociology, and artistic assessment. Keynes said roughly the same thing.

There have been cultural moments in which economists' rhetoric suggesting that technical utility analysis is a superior substitute for philosophical assessment of Aristotelian rationality has been based in earnest on technocratic hubris. However, my sense is that contemporary expressions of the sentiment much more frequently reflect reaction by economists, after having been lectured by their critics and their own heroes to manifest humbleness,¹⁰ against the audacity of philosophers' claims to expertise about general normative wisdom. Again, we can distinguish between a view of philosophers' ambitions as being grandiose, and a view of their claim to special expertise as being presumptuous. Either view, and obviously both held together, can motivate impatience with or indeed hostility to the Aristotelian project with respect to normative rationality.

It would take us far beyond the scope of this essay to engage in meta-philosophical *evaluation* of the extent of grandiosity and presumptuousness among philosophers.¹¹ My personal opinion, for what it is worth, is that grand questions will be posed regardless of whether they can be thought to have timeless best answers, that we should therefore prefer them to be addressed by careful thinkers who rigorously immerse themselves in deep scholarly traditions,

⁹ Duflo (2017) superficially seems to echo Keynes when she sketches a role for development economists as “plumbers”. By this she refers to the value of attention to local-scale features that accommodate or resist *installation* of policy ideas, as opposed to just intended inputs and desired outputs. However, Duflo emphatically fails to follow Keynes in urging humbleness upon economists. She thinks that while some economists engage in plumbing, *other* economists (“engineers”) should be entrusted with formal mechanism design and still others (“scientists”) should develop and test relevant generalisations. This amounts to assigning economists *all* of the primary roles in shaping and promoting public policy!

¹⁰ I am not claiming that economists who give policy advice generally *are* humble. Many are, but many others are obviously not. My claim is rather that a norm of humbleness now prevails, such that un-humble economists are expected to at least pay lip service to that attitude. Such signaling was a trope in the rhetoric even of Samuelson, during the era of maximal technocratic confidence (Backhouse 2017).

¹¹ Setting out to evaluate such presumed presumptuousness would perhaps be more presumptuous than the work to be evaluated! On the other hand, see Eklund (2017).

and that skilled practitioners can be objectively identified. My limited aim to this point has been to defend the claim that when economists model people or other natural choosers as economic agents, they are not engaged in a project that competes with, or even has any particularly close relationship with, philosophical reflection on rationality as an over-arching criterion for assessing whole human lives.

As a critical target, this is far from a straw person. A major scholarly industry, philosophical decision theory, has devoted itself to the project of building a general account of practical rationality on the basis of models of utility maximisation. I think that this enterprise rests on confusion, in the sense that its aim is Quixotic: pursuit of an unachievable objective that would not be worth having if we could get it.

I will demonstrate the grounds for this fundamental skepticism about a whole province of analytic philosophy by considering a recent example that meets a high standard of technical accuracy – that, indeed, is sound economic methodology if, contrary to its author’s stated purpose, it is read as such. Lara Buchak (2013) frames her work as inquiry into general practical rationality. She asks whether, to earn designation as generally rational, an agent’s choices should be captured by a utility function such that the choices optimise their subjective expected utility as per Savage (1954). The direct critical target can be regarded as Savage himself, who famously responded to Allais’s putative ‘paradox’ by revising his original preference across the two Allais choice contexts so as to conform to Subjective Expected Utility (SEU) theory and persevere what he characterised as his commitment to his own rationality. Buchak argues that Savage’s revision, rather than his first response, was his site of error: the rational agent, she argues, will not be generally constrained by Savage’s tradeoff consistency axioms (the sure-thing principle [STP] and the independence axiom), or the reduction-of-compound lotteries axiom (ROCL). Such agents thus shouldn’t be judged to be irrational if they violate SEU in these specific ways. However, their choices must, Buchak contends, be representable as optimising some member of the family of rank-dependent utility (RDU) functions (Quiggin 1982, 1993). Buchak develops axioms for what she calls a “Risk-Weighted Expected Utility” (REU) theory of the rational agent. The economic substance of REU is identical to that of generalised RDU; the difference is simply that RDU is descriptive whereas REU is explicitly normative. SEU is formally nested in RDU, so Buchak’s rational agent will conform their choice behaviour to SEU in some contexts but will depart from it in others. Choices among insurance policies, she argues, provide the clearest test cases for these different contexts.

Through her first three chapters, Buchak explicitly stalks the usual target of normative decision theory: the agent who deserves to be congratulated for general rationality thanks to making choices that respect appropriate consistency axioms. The standard seems to be an a priori ideal: we begin knowing what rationality conceptually demands, and we stress-test some axioms to see if they capture this conception.

However, to steal a nice joke from Ken Binmore (1998), in Chapter Four Buchak effectively de-Kants the operation when she narrows her focus to what she says is the basis for the *importance* of practical rationality. This is that we care whether the agent (1) tends to realize consequences that are aligned in expectation with her goals and (2) manifests choice patterns that allow error to be distinguished from bad luck. This is just what an economist practicing NSPE aims to do when she applies the intentional stance to behaviour that is sensitive to incentives defined by reference to consequences. Every hypothetical test case of a choice that she considers in her extended defence of REU can be reinterpreted directly as an instance of standard welfare comparison. Thus the economist who might feel like a visitor to a foreign land in the early chapters can settle in comfortably for the later ones. If REU were supposed to replace SEU as an account of welfare optimisation across the board (rather than of rationality, as Buchak intends), then most economists would want to hold out for a richer range of evidence and argument. But the economist reader need not accept REU to assess Buchak's analyses of cases as welfare assessment. Indeed, Buchak argues in her chapters 6 and 7 that when we consider decisions in diachronic contexts – precisely the contexts that the neo-Samuelsonian requires in order to characterise stochastic choice patterns as revealed preference – it's best to have *both* SEU and REU in the tool-box for her normative analysis. This is no casual concession: the second half of Buchak's book is devoted to chasing down the answer to what she regards as a live question, whether SEU *ever* does *normative* work that REU by itself does not. Ultimately she concludes that it does: the REU agent, she shows, will be ambivalent about some insurance offers where the SEU agent should rationally decline them. Insurance issues, Buchak maintains, provide deep test cases for normative assessments, and she is explicit that one should not try to settle them in general by dogmatically imposing REU as a *general* prior. In the end, REU wins as the general account of rationality because even in the cases where SEU delivers more specific insight, REU as a general policy minimises risk from fine errors of risk calculation.

Some existing economic literature echoes Buchak's conclusion as an insight about welfare evaluation rather than rationality. Harrison and Ng (2016) analyse an experimental insurance product market in terms closely aligned with Buchak's intuitions, though using structural estimation of real choice data rather than hypothetical toy problems that require no econometrics. Harrison and Ng also reach the kind of conclusion Buchak expects: optimal individual choices of insurance are sensitive to whether subjects are descriptively better classified as expected utility maximisers or rank-dependent utility maximisers, but this distinction does *not* distinguish the subjects whose choices imply welfare losses from those whose choices realise available consumer surplus. In another experimental setting, Harrison and Ross (2018) argue against any general programme, such as that constructed by Bleichrodt et al (2001), for paternalistically over-riding risk attitudes estimated as revealed preferences under RDU in favour of what agents would choose if they conformed to expected utility

theory (EUT).¹² We nevertheless offered policy advice in the specific real-world problem context we analysed (choices of household investment portfolios from a particular fixed menu, with specified information provision about the options) based on optimisation under EUT. As humble economists, in Keynes's sense, we based this on our empirical findings, along with knowledge we had about our clients' exogenous value context (wanting households' expected retirement resources to be maximised), not on any general theory of rationality. The key point is that Buchak does not advocate appealing, on a priori grounds, to REU (or SEU) in this kind of application either. Economists may therefore doubt that if their tool-box includes RDU, adding REU to it can bring any additional value with respect to their purposes.

But what other purpose is there? There is *only* Buchak's original, stated purpose of deciding who deserves to be called 'rational'. By virtue precisely of the sensible arguments about (hypothetical) cases to which she turns after her third chapter, and her solid understanding of the relevant economics, she reaches what at least some economists consider sound theoretical conclusions – thereby demonstrating that the philosophical framing is redundant. Philosophers who conceive of general rationality in Aristotle's sense are apt to be no more convinced of the project's importance.

The defender of philosophical decision theory might, however, here try to turn the tables by arguing that defenders of NSPE could read Buchak as throwing a life preserver to save us from the other challenge arising in Hands's critique, that of saying what differing descriptive models of choice have to do with welfare. That is, my reconstruction of her overall argument could be read as showing that if the NSPE advocate can somehow salvage the claim that rationalisability under the intentional stance delivers a theory of normative rationality, then analyses like Buchak's show us that we can derive intuitive welfare consequences from the kinds of empirical methods we favour. However, the price of this strategy is too high. Use of the intentional stance is indeed normative, but in itself it holds the bar of rationality very low. Merely applying the intentional stance can't tell us that welfare is what we should focus on for policy analysis. Additional grounds must be sought.

3. *Welfare*

¹² I refer here to EUT rather than SEU to emphasise that we have descended from Buchak's philosophical heights to the ground of descriptive economic identification, paralleling reference to RDU instead of REU. A further complication here is that Bleichrodt et al say they compare analysis of observed choices conditional on EUT with analysis of the same choices conditional on Cumulative Prospect Theory (CPT). But as they consider data in which subjects encounter no loss frames, they cannot empirically pin down CPT's special λ parameter for loss aversion, and should be properly interpreted as talking about an RDU model.

Economic policy advice is for humans and human institutions, not non-existent ‘inner rational agents’ (Infante et al 2016, Sugden 2018), and also not, according to my argument above, to humans who aspire to be called ‘rational’ by making their discrete choices consistent. The goal of the policy maker, one might then think, is to try to create conditions in which as many people as possible, given decent luck, can enjoy lives that combine psychological peace with earned pride in judgments well made and mistakes well managed and learned from. Thus it might be supposed from what I have said that the policy maker should seek advice on how to help people be generally rational in Aristotle’s sense. And I have just argued that general rationality is not connected in any direct sense with maximising economic utility.¹³

But in fact economists, in offering policy recommendations, do not try to promote general rationality. Economists whose methodology is accurately characterised by NSPE give advice to agents about how to make future choices, and current choices with future consequences, consistent with more efficiently achieving goals consistent with those revealed by the agents’ own past patterns of choice. They refer to this as promoting ‘welfare’. Hands’s critical demand is for arguments, consistent with CRPT/NSPE, as to why economists expect that people should want to receive and follow such advice. More directly, given economists’ main practical role, why should people be expected to want their civil servants, political representatives, and corporate managers to follow such advice on their behalf?

It follows from the disassociation I have urged between utility maximisation and general rationality that people should want *personal* advice from economists only when their well-being is significantly put at risk by choices that are inconsistent in the hard-to-monitor or hard-to-forecast ways that economists specialise in identifying. There are familiar such contexts in Marshall’s “ordinary business of life”. Most people’s well-being will be seriously compromised if they fear being drastically consumption-constrained in the future relative to the present, and will be still more seriously undermined if such fears turn out to be realised. In modern societies, avoiding such outcomes requires accumulating and maintaining at least minimum levels of wealth, and hard-to-identify inconsistent investments of existing wealth strongly tend to impede such accumulation and maintenance. Wise people look for such expert financial advice as they can find and afford, and such advice is economic advice regardless of whether those who provide it are economists strictly speaking. The best such advice, in general, is what would be given to a modeled economic agent who shared the advisee’s future consumption aspirations and dynamic budget constraints.

This is the severely limited domain of economics as Aristotle understood it (Meikle 1995). In modern times economists have mainly delegated it to financial

¹³ As conceded above, general rationality plausibly excludes economic *insanity* – being oblivious to relative prices and relative risks – but economists don’t give advice against this because it is never an actual problem.

services professionals, though it is implied second-order advice when economists advise retail investment providers, as in Harrison and Ross (2018). An economist who ventures beyond it in the realm of personal advice – for example, telling a person that she should align her revealed preferences over health risk with her revealed preferences over education choice risk, even where she can't buy insurance policies that make these mutually fungible – should not be surprised to be asked by the advisee why this should matter to her. In the actual world, economists normally do not give unsolicited advice about personal choice patterns that cannot be financially mediated and thus transformed into advice about alternative investments. Perhaps this is just because economists don't try to supply non-existent demand, but perhaps the non-existence of 'personal economists' also reflects recognition that welfare is not the same thing as personal well-being, and economists are experts on the former. This point responds directly to an example used by Hands (2013, p. 1100) when he complains that economics based on CRPT doesn't help him improve his stock-picking.

Where economic advice and welfare maximization meet, a domain Aristotle didn't imagine, is in policy for people whose choices constrain the options of other, generally anonymous, people. In standard cases, free of corruption, they don't give advice that would maximise the personal utility of the advisees, who are treated as agents, but advice that addresses the utility of the third parties with legitimate stakes in the outcomes of the choices, whom they thus effectively treat as principals. Since these principals usually have heterogeneous, and very often conflicting, utility functions over the potential outcomes, economists shape their advice by looking for aggregate-scale efficiencies in welfare. These are often very difficult to uncontroversially identify, but economists are the most reliable experts in identifying them, and agree about them relatively often. For a familiar example, they regularly remind governments that raising special cost barriers against imported products, if they are safe and not socially or environmentally harmful in themselves, is almost always welfare inefficient. This analysis is not always decisive; the argument against trade barriers is defeasible if institutional arrangements allow a special subset of the principals to hog almost all the welfare gains for themselves and another subset suffers losses.

This is of course the standard picture of applied economics. As raised at the outset, the critical question about it posed by Hands, following Hausman, is: what is the normative justification for advocating policies that target welfare improvement if, as per CRPT/NSPE, welfare is disassociated from subjective satisfaction in the folk psychological sense – and, as I have now argued, also from well-being on its best considered philosophical interpretation? The economist's instinctive first answer might be: we have an efficiency measure for welfare but not for folk-psychological satisfaction or Aristotelian well-being. The standard retort to this occurs repeatedly in the critical literature about standard economic practice: that it is analogous to searching for dropped keys under the streetlight. A procedure, that is, is defended on the grounds that we know how to use it, rather than on the basis of an argument for its actual normative importance.

I think, however, that the economist's instinctive response is correct. NSPE, I contend, helps us see its justificatory logic better than older accounts of economic agency according to which economics idealises practical folk psychology or utilitarian theories of well-being.

A first gain in clarity from NSPE here is that it naturally separates criticisms based on concern for folk-psychological satisfaction from criticisms based on normative prioritisation of well-being. The externalist / Dennettian account of intentionality denies that the traditional, internalist model of psychological satisfaction captures anything that satisfies the criteria for acknowledgment as a real pattern. The closest real phenomena to the objects of the internalist account are relatively stable patterns of response that people give when they are asked to choose actions that reflect comparative evaluations of possible prospects. That is to say: they are the patterns we identify when we probe for revealed preferences using our most reliable methods, particularly incentivised experimentation.

That this reconceptualisation is not mere semantic footwork is shown if we ask whether these revealed preferences are 'subjective'. On the internalist account, the immediate affirmative answer derives from the assumption that they are privately experienced and reported on the basis of introspection or phenomenology, a person's reading of her own mind. The externalist, by contrast, grants them subjectivity only in the sense they reflect an individual's idiosyncratic social history, on which the individual in question has the greatest *pre-scientific* expertise because she has the densest record of observations. This expertise is limited in scope. Experiments might show, for example, that the person's choices over risky monetary prospects are best characterised by a concave utility function combined with an inverse-S shaped rank-dependent probability weighting function with particular parameter values in the flexible Prelec (1998) specification, and this is highly unlikely to be something the subject antecedently believed about herself on the basis of everyday observation (Harrison and Ross 2018). Furthermore, the revealed preferences are not primitive, fixed properties of the individual in the sense of being either innate or wholly self-constructed: they are products of social mindshaping, and they are labile.

The social origins of externally attributed preferences – along with the beliefs and other PAs that, as Hausman (2012) rightly stresses, are essential for predicting behaviour – partly explains their central importance in policy selection. People engage continuously and pervasively in mindshaping because most of their social behaviour involves cooperation and normatively or literally regulated competition. Mindshaping creates shared reference points that furnish enough structure for people to select strategies, divide labour in sensible ways, and stabilise descriptive and normative expectations (Ross 2005, Bicchieri 2006, 2017, Ross et al 2021). Thus their revealed preferences literally *constitute* the normative elements around which social policies are conceptualised and bargained over (implicitly or institutionally; see Binmore 1994, 1998, 2005). Since, according to NSPE, welfare is nothing over and above the measurement system for comparatively assessing payoffs of social interactions, it comes close to being a tautology that

welfare is the target of public policy (not just economic policy), and particularly of policies that are institutionally situated as ‘official’.

*Close to a tautology, but not quite.*¹⁴ In the previous section I argued that general rationality applies, or fails to apply, to whole-life narratives, and that these are conceptually distinct from the kinds of preferences over which policies are coordinated. One crucial dimension of difference is that, as all of the leading philosophers in the Aristotelian tradition stress, general rationality revolves around integrating what Gibbard (1990) calls “wise choices” with “apt feelings”. The latter are richly particular and sensitive to distinctive personal history profiles, which is why general rationality is usually assessed through criticism of narratives. Administrative systems typically abstract away from aptness of feelings, precisely because these are so thick and particular. For one thing, this causes them to resist aggregation and quantitative representation and measurement. For another thing, public institutions often incorporate explicit mandates to handle similar claims to public resources similarly, which requires the construction of relatively wide equivalence classes of outcomes. If I favour a policy as a sad necessity while you favour the same policy as a splendid and progressive triumph, from the perspective of welfare analysis we both count as winners if the policy is adopted. But our intimates and our biographers will register our attitudes very differently when incorporating our respective emotions about the policy into their assessments of our general rationality.

On the modernist conception of the state that has dominated wealthy societies for about two centuries – the conception for which Hegel struggled to forge a descriptively adequate language (Herrmann-Pillath and Boldyrev 2014), and which was first fully articulated in empiricist-friendly terms by Max Weber (1922) – professional civil servants regard themselves as agents who are tasked by a public mandate to design and implement policies and regulations on behalf of an aggregated set of principals (roughly, citizens) by reference to the principals’ welfare (Besley 2006; Dowding and Taylor 2019). Economists choose and analyse problems as if they are technical advisors to these agents. As Sugden (2018) rightly complains, they often generate such advice without any explicit commission, and with no attention to the limited power of *actual* such agents. That is, as Sugden puts it, they direct advice to hypothetical benevolent dictators. However, often enough economists receive genuine commissions to assess or design policies, and combine broad practical awareness of solution constraints with Keynes’s recommended normative humbleness (Ross and Townshend 2021).

My brief here is to defend the normative importance of welfare as understood in NSPE, not to argue that all is well in the world. *Exclusive* focus on welfare in public policy is open to criticism for ignoring dimensions of value that resist quantitative measurement (Alexandrova 2017). Another, particularly serious and consequential, problem is that economists are often not very loud about reminding

¹⁴ There would not be a basis for objection here if it were entirely a tautology. Theories that specify identification restrictions are not empirical hypotheses.

governing elites that to be satisfied with mere Kaldor-Hicks efficiency, regardless of whether gains from policies are in fact used to compensate losers (Blinder 2018; Boix 2019), can generate the opposite of welfare improvement. Arguably, this recurrent failing has been an important contributing cause of the widely perceived crisis of trust in government and expertise, and associated surges of populism that threaten the legitimacy of elections, the rule of law, and public willingness to receive and respect technically informed advice (Nichols 2017; Tanzi 2017).

It is far from clear that, even in the absence of crisis-level strains in welfare-focused governance, most members of democratic societies genuinely assent to the role of principals who hire officials as agents to promote welfare. Expressive voting in electorates can be construed as reflecting an idea that political leaders should promote general well-being by exemplifying ideals of character, and citizens under the sway of this idea may display near-indifference to policy choices framed by reference to welfare (Caplan 2007; Achen and Bartels 2016). Demand for politics of this kind is energetically amplified by supply from political entrepreneurs.

The result tends to be disastrous for general well-being as philosophers understand it. In pluralistic societies, a politician whose central campaign theme is “I am a representative of your generally rational ideal” can only give concrete policy content to this fantastic pretence by favouring some interest groups over others, with no obligation to address welfare efficiencies at all. Worse, if a sub-population’s interests are promoted on grounds that they are more generally rational than those of rival coalitions, then the tight entanglement of general rationality with deep moral commitments leads quickly to the moralisation of all disagreements, to elections that are existential crises for all sides, and to willingness to abandon constitutional rules and scruples because winning at all costs seems justified by higher morality. The situation just described is hardly hypothetical in current circumstances in many countries, especially larger ones, and its baleful effects on *both* general well-being and welfare can hardly be overstated. The situation is especially dire when extremely urgent and difficult threats, such as climate change, require widespread pragmatic bargaining over trade-offs on multiple intersecting margins. That task demands policies that target carefully estimated welfare effects.

Some critics of standard economic policy assessment, for example Sen (1999) and Nussbaum (2000), argue that economic development policy should directly target general well-being rather than welfare. Answering these criticisms of ‘welfarism’ extends beyond the response to Hands’s and Hausman’s concern, which requires only defending the claim that welfare as conceptualised according to NSPE is normatively significant. Neither Sen nor Nussbaum assert that welfare is normatively irrelevant. However, a few words should be said about their views in light of my use of the contrast between welfare and well-being in explaining why most applied economists focus on welfare.

Sen and Nussbaum seem to offer (slightly different) ways around the worry that political emphasis on general rationality leads to demagoguery. That is, they identify what they claim are measurable proxies for general well-being and argue that public policy should target these proxies. Claims I have defended here indicate lines of defence for the welfarist. Nussbaum agrees with my claim that the articulation of both general well-being and general rationality is essentially discursive and thus doesn't lead toward a measurable common currency for the many dimensions of general well-being. Consequently, her approach is open to Dasgupta's (2009) criticism that it offers no guidance about trade-offs given scarce resources. (The scarcest resource in question is likely to be political rather than material capital.) Similar comments apply to Sen's approach, along with the additional problem that he defends so many proxies for good development policy that almost any official agent will always be able to claim success along some measure his account promotes. Given that accountability is arguably the single greatest real barrier to sustained development promotion, I think this is a decisive objection. Sen also does not rigorously address the question of how his proxy measures dynamically cross-predict one another empirically. If, as seems probable, household consumption expenditure predicts the others better than they predict themselves (Ravallion 1992), then we could preserve accountability by demanding that development policies improve that. And household consumption expenditure seems closer to what economists aim to capture in welfare theory than to what philosophers aim to characterise when inquiring into general well-being.

4. *The reach of NSPE*

Hands (2013, p. 1103) concludes his criticism of CRPT/NSPE by suggesting that empirical evidence for its success is restricted to a particular corner of economics, applied demand analysis, but that as yet we have no grounds to believe "that it would be possible, and better, to do all of choice-theoretic economics – in macroeconomics, finance, industrial organization, law and economics, cost-benefit analysis, and all of the other things that modern economists do – by merely projecting patterns gleaned from GARP-consistent choice data onto new sets of parameters." Thus Hands doubts that CRPT can yet be sold as the basis for a general philosophy of economics, as NSPE promises. He is clear, however, that the main basis for regarding the general extension of CRPT as far-fetched is his reading of it as a form of instrumentalism about choice. He diagnoses "a notorious problem with instrumentalism – it doesn't travel well from domain to domain and must prove its robustness in each new application."

This complaint exposes the point that the crucial foundation of NSPE is its identification of a version of RPT consistent with stochastic choice with the realistic interpretation of the intentional stance wobbily defended by Dennett (1991a) and formulated more categorically by Ross (2000) and Ladyman and Ross (2007). The intentional stance is a theory of choice as responsiveness to incentives, because it is to begin with a theory of what it is to be an agent, according to which the extensions of 'agent' and 'incentive-influenced entity' are

identical. NSPE's central claim is that CRPT is simply the theory of the quantitative application of the intentional stance. Because the intentional stance is a kind of realism (dubbed 'rainforest realism' by Ladyman and Ross), not instrumentalism, Hands's general philosophical basis for doubting its reach is rejected.

Of course it must be conceded that NSPE will need to prove itself in empirical applications. That applies to every philosophical theory of anything. A survey of applications from across the empire of economics would necessarily be a monograph-length enterprise. Here I will indicate just a couple of signposts, aligned with Hands's list of candidates above. Ross (2014) cites Frydman and Goldberg (2007, 2011) as a programme for macroeconomics that is naturally interpreted as applied NSPE (though it needs to be dissociated from spurious connection by its authors to prospect theory, when the established high frequency of rank-dependent preferences suffices as an assumption for their modeling). As for industrial organisation, for decades theory in that area has been almost exclusively developed using non-cooperative game theory, thus depending on assignment of utility functions to individual agents that are typically firms. NSPE, in dissociating preferences from internal psychological states, licenses first-order ascription of real utility functions to firms, as opposed to taking them to be second-order approximations based on constructed representative agents.

All of microeconomics concerns choices of agents. NSPE, in incorporating a theory of choice and agency, thus applies to microeconomics without restriction. The same goes for macroeconomics *if* it is thought to depend on microfoundations.¹⁵

5. Conclusion

Hands is correct that CRPT is primarily motivated by interest in connecting utility theory with described empirical behaviour. NSPE, the philosophical generalisation of this methodology, is still more explicit in aiming to ground a *realist* interpretation of descriptive economic models. This furnishes its defender with straightforward resources for responding to those elements of Hands's criticism that depend on reading CRPT as instrumentalism. However, Hands puts useful pressure on an incomplete aspect of NSPE: the scope it provides for justifying normative economics. To the extent that theories of rationality and welfare rest on folk psychological conceptions of human reasoning and well-being, NSPE's abandonment of folk psychology presents a challenge.

I have argued that the challenge should be met by conceding that economists are not in the business of assessing general rationality. Casual references to rationality

¹⁵ In fact, I am skeptical of that widely held view, for reasons given in Hoover (2009) and Ross (2014). Macroeconomists aim to identify real patterns, but these might abstract away from reference to agents' choices. I thus think that details of relationships between macroeconomic theory and NSPE remain unresolved. This is hardly surprising, given the current absence of consensus on the foundations of macroeconomic theory in general.

that litter meta-textual commentary across the discipline obscure this. More directly, the wholesale borrowing of theoretical structures from the foundations of microeconomics in the project of philosophical decision theory encourages the idea that there should be a rigorous bridge between economists' interest in technical choice consistency and philosophers' interest in general rationality. NSPE implies that this approach to bridge-building is misguided. Dynamic rationality in a whole person is essentially a distinct normative issue from descriptive consistency across discrete choices by economic agents. This would be a serious problem if economists, like decision theorists, were in the business of giving general advice to individual people. But they are not. Economics is a policy science, but the policy domains to which it aims to be relevant are public and corporate. Once this is recognised, NSPE actually provides clearer insight than alternative philosophies of economics as to why economists concentrate on welfare, rather than well-being, as their primary normative target. As welfare is not the only valuable dimension of social life, let alone personal life, economists offering contributions to social guidance should read their Keynes and be humble about their role.

References

- Achen, C., & Bartels, L. (2016). *Democracy for Realists*. Princeton University Press.
- Afriat, S. (1967). The construction of utility functions from expenditure data. *International Economic Review* 8: 67-77.
- Alexandrova, A. (2017). *A Philosophy for the Science of Well Being*. Oxford University Press.
- Alekseev, A., Harrison, G., Lau, M., & Ross, D. (2018). Deciphering the noise: the welfare costs of noisy behavior. Center for the Economic Analysis of Risk Working Paper 2018-01. <https://cear.gsu.edu/wp-2018-01-deciphering-the-noise-the-welfare-costs-of-noisy-behavior/> .
- Amadae, S. (2003). *Rationalizing Capitalist Democracy*. University of Chicago Press.
- Amadae, S. (2016). *Prisoners of Reason*. Cambridge University Press.
- Andersen, S., Harrison, G., Lau, M., & Rutström, E. (2008). Eliciting risk and time preferences. *Econometrica*, 76: 583–618.
- Backhouse, R. (2002). *The Ordinary Business of Life*. Princeton University Press.
- Backhouse, R. (2017). *Founder of Modern Economics: Paul A. Samuelson* (Volume 1). Oxford University Press.
- Bernheim, B. D. (2016). The good, the bad, and the ugly: A unified approach to behavioral welfare economics. *Journal of Benefit Cost Analysis* 7: 12-68.

- Bernheim, B.D., & Rangel, A. (2008). Choice-theoretic foundations for behavioral welfare economics. In A. Caplin & A. Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*, pp. 155-192. Oxford University Press.
- Besley, T. (2006). *Principled Agents: The Political Economy of Good Government*. Oxford University Press.
- Bicchieri, C. (2006). *The Grammar of Society*. Cambridge University Press.
- Bicchieri, C. (2017). *Norms in the Wild*. Oxford University Press.
- Binmore, K. (1994). *Game Theory and the Social Contract Volume 1: Playing Fair*. MIT Press.
- Binmore, K. (1998). *Game Theory and the Social Contract, Volume 2: Just Playing*. MIT Press.
- Binmore, K. (2005). *Natural Justice*. Oxford University Press.
- Binmore, K. (2009a). *Rational Decisions*. Princeton University Press.
- Binmore, K. (2009b). Interpersonal comparisons of utility. In H. Kincaid & D. Ross, eds., *The Oxford Handbook of Philosophy of Economics*, pp. 540-559. Oxford University Press.
- Bleichrodt, H., Pinto, J., & Wakker, P. (2001). Using descriptive findings of prospect theory to improve the prescriptive use of expected utility. *Management Science* 47: 1498-1514.
- Blinder, A. (2018). *Advice and Dissent*. Basic Books.
- Bogdan, R. (1997). *Interpreting Minds*. MIT Press.
- Bogdan, R. (2000). *Minding Minds*. MIT Press.
- Bogdan, R. (2009). *Predicative Minds*. MIT Press.
- Bogdan, R. (2010). *Our Own Minds*. MIT Press.
- Bogdan, R. (2013). *Mindvaults*. MIT Press.
- Boix, C. (2019). *Democratic Capitalism at the Crossroads*. Princeton University Press.
- Bourgeois-Gironde, S. (2020). *The Mind Under the Axioms*. Academic Press.
- Buchak, L. (2013). *Risk and Rationality*. Oxford University Press.
- Burge, T. (1986). Individualism and psychology. *Philosophical Review* 95: 3-45.

- Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature* 43: 9-64.
- Caplan, B. (2007). *The Myth of the Rational Voter*. Princeton University Press.
- Chater, N. (2018). *The Mind is Flat*. Allen Lane.
- Clark, A. (1997). *Being There*. MIT Press.
- Colander, D., & Su, H.-C. (2018). *How Economics Should Be Done*. Edward Elgar.
- Dasgupta, P. (2009). Facts and values in modern economics. In H. Kincaid & D. Ross, eds., *The Oxford Handbook of Philosophy of Economics*, pp. 580-640. Oxford University Press.
- Dennett, D. (1987). *The Intentional Stance*. MIT Press.
- Dennett, D. (1991a). Real patterns. *Journal of Philosophy* 88: 27-51.
- Dennett, D. (1991b). *Consciousness Explained*. Little, Brown.
- Dennett, D. (1991c). Two contrasts: folk craft versus folk science, and belief versus opinion. In J. Greenwood, ed., *The Future of Folk Psychology*, pp. 135-175. Cambridge University Press.
- Dowding, K., & Taylor, B. (2019). *Economic Perspectives on Government*. Palgrave Macmillan.
- Duflo, E. (2017). The economist as plumber. *American Economic Review*, 107: 1-26.
- Eklund, M. (2017). *Choosing Normative Concepts*. Oxford University Press.
- Fodor, J. (1978). Propositional attitudes. *The Monist* 61: 501-524.
- Frydman, R., & Goldberg, M. (2007). *Imperfect Knowledge Economics: Exchange Rates and Risk*. Princeton University Press.
- Frydman, R., & Goldberg, M. (2011). *Beyond Mechanical Markets*. Princeton University Press.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Harvard University Press.
- Gul, F., & Pesendorfer, W (2001). Temptation and self-control. *Econometrica* 69: 1403-1435.
- Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. In A. Caplin & A. Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*, pp. 3-39. Oxford University Press.

- Hands, W. (2013). Foundations of contemporary revealed preference theory. *Erkenntnis* 78: 1081-1108.
- Harrison, G., & Ng, J.-M. (2016). Evaluating the expected welfare gain from insurance. *Journal of Risk and Insurance* 83: 91-120.
- Harrison, G., & Ross, D. (2018). Varieties of paternalism and the heterogeneity of utility structures. *Journal of Economic Methodology* 25: 42-67.
- Harrison, G., & Rutström, E. (2008). Risk aversion in the laboratory. In J. Cox & G. Harrison, eds., *Risk Aversion in Experiments*, pp. 41-196. Emerald.
- Hausman, D. (1992). *The Inexact and Separate Science of Economics*. Cambridge University Press.
- Hausman, D. (1995). Rational choice and social theory: A comment. *Journal of Philosophy* 92: 96-102.
- Hausman, D. (2000). Revealed preference, belief, and game theory. *Economics and Philosophy* 16: 99-115.
- Hausman, D. (2008). Mindless or mindful economics: A methodological evaluation. In A. Caplin & A. Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*, pp. 125-151. Oxford University Press.
- Hausman, D. (2012). *Preferences, Value, Choice, and Welfare*. Cambridge University Press.
- Herrmann-Pillath, C., & Boldyrev, I. (2014). *Hegel, Institutions, and Economics*. Routledge.
- Hoover, K. (2009). Microfoundations and the ontology of macroeconomics. In H. Kincaid & D. Ross, eds, *The Oxford Handbook of Philosophy of Economics*, pp. 386-409. Oxford University 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.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263-292.
- Keynes, J.M. (1936). Economic possibilities for our grand-children. Re-printed in Keynes, *Essays in Persuasion*, pp. 358-373. Norton (1963).
- Ladyman, J., & Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalised*. Oxford University Press.
- McClamrock, R. (1995). *Existential Cognition*. University of Chicago Press.
- Meikle, S. (1995). *Aristotle's Economic Thought*. Oxford University Press.

- Nichols, S., & Stich, S. (2003). *Mindreading*. Oxford University Press.
- Nichols, T. (2017). *The Death of Expertise*. Oxford University Press.
- Nussbaum, M. (1997). *Cultivating Humanity*. Cambridge University Press.
- Nussbaum, M. (2000). The costs of tragedy: Some moral limits of cost-benefit analysis. *Journal of Legal Studies* 29: 1005-1036.
- Pennesi, D. (2021). Intertemporal discrete choice. *Journal of Economic Behavior and Organization* 186: 690-706.
- Prelec, D. (1998). The probability weighting function. *Econometrica* 66: 95-113.
- Putnam, H. (1975). *Mind, Language, and Reality*. Cambridge University Press.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3: 323-343.
- Quiggin, J. (1993). *Generalized Expected Utility Theory. The Rank-Dependent Model*. Boston: Kluwer.
- Ravallion, M. (1992). Poverty comparisons: a guide to concepts and methods. World Bank Living Standards Measurement Study Working Paper 88.
- Ross, D. (2000). Rainforest realism: A Dennettian theory of existence. In D. Ross, A. Brook, & D. Thompson, eds., *Dennett's Philosophy: A Comprehensive Assessment*, pp. 147-168. MIT Press.
- Ross, D. (2005). *Economic Theory and Cognitive Science: Microexplanation*. MIT Press.
- Ross, D. (2008). Ontic structural realism and economics. *Philosophy of Science* 75: 732-743.
- Ross, D. (2011). Estranged parents and a schizophrenic child: Choice in economics, psychology, and neuroscience. *Journal of Economic Methodology* 18: 217-231.
- Ross, D. (2012). The economic agent: not human, but important. In U. Mäki, ed., *Handbook of the Philosophy of Science, v. 13: Economics*, pp. 691-735. London: Elsevier.
- Ross, D. (2014). *Philosophy of Economics*. Palgrave Macmillan.
- Ross, D., & Stirling, W. (2021). Economics, social neuroscience, and mindshaping. In J. Herbecke & C. Herrmann-Pillath, eds., *Social Neuroeconomics*, pp. 174-201. Routledge.
- Ross, D., Stirling, W., & Tummolini, L. (2021). Strategic theory of norms for empirical applications in political science and political economy. In H. Kincaid &

- J. Van Bouwel, eds., *The Oxford Handbook of Philosophy of Empirical Political Science*. Oxford University Press.
- Ross, D., & Townshend, M. (2021). Everyday economics. In H. Kincaid & D. Ross, eds., *Modern Guide to Philosophy of Economics*. Edward Elgar.
- Savage, L. (1954). *The Foundations of Statistics*. Wiley.
- Sen, A. (1973). Behaviour and the concept of preference. *Economica* 40: 241-259.
- Sen, A. (1999). *Development as Freedom*. Oxford University Press.
- Smith, V. (2008). *Rationality in Economics*. Cambridge: Cambridge University Press.
- Sugden, R. (2018). *The Community of Advantage*. Oxford University Press.
- Tanzi, V. (2017). *Termites of the State*. Cambridge University Press.
- Thaler, R. (1992). *The Winner's Curse*. New York: Free Press.
- Tiberius, V. (2010). *The Reflective Life*. Oxford University Press.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297–323.
- Weber, M. (1922). *Wirtschaft und Gesellschaft*. Mohr Siebeck.
- Wilcox, N. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. Cox & G. Harrison, eds., *Risk Aversion in Experiments*, pp. 197-292. Emerald.
- Wilcox, N. (2021). Utility measurement: Some contemporary concerns. In H. Kincaid & D. Ross, eds., *Modern Guide to Philosophy of Economics*. Edward Elgar.
- Williams, B. (1981). *Moral Luck*. Cambridge University Press.
- Williams, B. (2006). *Philosophy as a Humanistic Discipline*. Princeton University Press.
- Zawidzki, T. (2013). *Mindshaping*. MIT Press.