

THE SUBJECTIFICATION OF *HOMO* *ECONOMICUS* IN BEHAVIOURAL ECONOMICS

David Primrose

In recent years, heterodox economists challenging ‘mainstream’ theories and policy prescriptions have been criticised for misrepresenting the heterogeneity of the latter (*e.g.* Fullbrook 2001; Garnett 2006). Their categorisation of ‘mainstream economics’ as synonymous with neoclassicism – grounded in ‘a strict adherence to the holy trinity [of] rationality, selfishness, and equilibrium’ (Colander *et al.* 2004: 485) – is said to hinder heterodox economists from acknowledging the nascent *pluralism* now characterising the discipline (Rodrik 2015). With the increasing prevalence of fields such as behavioural economics, new institutional economics, new information economics, evolutionary economics, various iterations of game-theory, experimental economics, social choice theory, agent-based complexity-theory and neuroeconomics, neoclassical economics is depicted to have been progressively supplanted as the dominant research programme. The ‘death of neoclassical economics’ has even been claimed (*e.g.* Colander 2000). In its place, these ‘genuinely different approaches’ (Davis 2006: 9) offer a ‘more eclectic position of purposeful behaviour, enlightened self-interest and sustainability’ (Colander *et al.* 2004: 485).

This ‘pluralist turn’ in mainstream economics constitutes the background for critically engaging with one constitutive approach: behavioural economics (‘BE’). This article considers its philosophical foundations and questions the extent to which these depart from neoclassicism. It challenges the self-representation of BE as transcending neoclassical

Primrose, D. (2017)

**‘The Subjectification of *Homo Economicus* in Behavioural Economics’
Journal of Australian Political Economy
No. 80, pp. 88-128.**

presuppositions of hyper-rational¹ decision-making in favour of more empirically realistic representations of behaviour (e.g. Rabin 1998, 2002; Camerer 1999, 2003; Altman 2015). It argues that a methodological commitment to greater empirical realism, alongside retention of neoclassical axioms of hyper-rationality, reflects behavioural economists' efforts to reconfigure, rather than reject, the *theoretical problematic of neoclassical humanism* (Madra 2016). That is, BE utilises psychological insights to understand why markets fail to function as predicted by neoclassicism. By articulating bounded-rationality as an irrational 'deviation' from *Homo Economicus*, BE positions *Homo Economicus* as the ideal subject for functioning markets and potentially realisable through corrective measures.

In presenting this case, the article is structured as follows. The next section examines the respective positions adopted by neoclassicism and BE concerning the 'rationality' of subjects. While BE criticises *Homo Economicus* as empirically unrealistic, BE's methodology retains some key neoclassical features. Drawing on this critique, section three lays the foundations for a more capacious appraisal of the performative significance of BE by introducing theoretical humanism and its relevance to neoclassicism. Emphasis is placed on delineating the two theoretical presuppositions underpinning theoretical humanism – namely, the concepts of human subject and social reconciliation – and how these structure the theoretical problematic of neoclassicism. Section four argues that, through its reworking of this problematic, BE subjectifies (Foucault 1982, 2003) *Homo Economicus* as the microfoundation for functioning markets. In pathologising 'irrational' decision-making as responsible for markets failing to operate as predicated, behaviouralism brings *Homo Economicus* back in as both a normative ideal and potentially realisable subject. Section five then addresses the policy implications arising from this pathologisation. It considers how boundedly-rational subjects are positioned as legitimate targets for corrective actions to make their decisions approximate hyper-rationality.

¹ To circumvent the neoclassical practice of contrasting hyper-rationality as perfect with actual behaviour that is imperfect, this article will refer to the neoclassical conception of rationality as the doctrine of 'hyper-rationality'. This distinguishes neoclassical conceptions of rationality from the broader philosophical principle of 'rationality' which hold that actions and opinions should be grounded in reason (Shaikh 2016: 78).

Behavioural economics: challenging *Homo Economicus*?

To begin, it is useful to be clear about what comprises BE. A distinction is often made between the ‘old and ‘new’ strands of BE (Sent 2004; Heukelom 2014). The former, associated with Simon (1955, 1956, 1959), Selten (1998) and Gigerenzer (2000, 2015), constructs an account of rationality as ‘bounded’ and rejects the neoclassical atomistic conception of individuals in favour of an evolutionary account promulgating a more holistic, ecological conception of rationality and individuality (Goldstein and Gigerenzer 2002). The ‘new’ school – arising from the work of Kahneman and Tversky (1979) – retains neoclassicism’s atomistic conception of individuals, but revises it to embed agents within an ahistorical and non-developmental social ontology (Davis 2011, 2015; Leggett 2014). This article concentrates on the newer iteration as it constitutes the mainstream of behavioural research (Heukelom 2014) and is also the most politically influential – as manifest in the growing institutionalisation of ‘nudge’ research units in governments around the world (Jones *et al.* 2013). Hence, unless otherwise-specified, all references to ‘BE’ throughout this article refer exclusively to the *new* school. The first issue to consider is how it relates to neoclassical economics.

Neoclassical Hyper-Rationality: Homo Economicus

Neoclassicism seeks to develop a positive theory promulgating the notion that all individuals approach decision-making in a universally ‘rational’ manner, thereby engendering common responses. This standard model is based on a corpus of assumptions concerning the decision-making characteristics of agents, rendered axiomatic in rational-choice theory and embodied in the avatar of *Homo Economicus* (Hollis and Nell 1975). Neoclassical rationality theory, or expected utility theory, is axiomatic in that it is constructed around a specific set of logical assumptions (completeness, transitivity, independence and continuity) formulated to produce well-defined preference orderings (Davis 2011). The ‘von Neumann-Morgenstern utility function theorem’ articulates that satisfaction of these axioms enables representation of any set of well-ordered preferences through a distinct individual utility function, as embodied in *Homo Economicus* (von Neumann and Morgenstern 2007). That is, conceiving the fundamental economic problem as relations

between a suite of scarce resources and psychological ends (utility) (Robbins 1935), actors are conceptualised as making choices which maximise individual utility. This choice, in turn, reduces to the four consistency conditions above. This formally defines hyper-rationality in that it is the violation of these conditions that results in incorrect choices and 'irrational' behaviour (Sen 1977). Hyper-rationality and individuality are, thus, deductively and logically coupled: the utility function representation of individuals is logically derived from the axiomatic assumptions governing preferences, while selection of those assumptions is driven by the imperative to extrapolate a distinct utility function securing rational choice-making, understood as choice consistency (Davis 2011, 2015).

More specifically, three qualities are attributed to the hyper-rational subject of *Homo Economicus*. First, each individual is assumed to possess consistent, well-defined and coherent preferences, grounded in self-interested forms of utility maximisation as manifest in her/his choices. Second, these preferences are taken to be rationally maximised by individuals – assuming a given set of available options, s/he evaluates the costs and benefits of countless choices and follow the appropriate strategy to maximise her/his expected utility. Third, the model presumes that individuals possess well-formed assumptions applicable to any given situation and modify these when receiving new information. Consequently, individuals are not cognitively impeded in assessing the various alternatives at hand, nor hindered by problems of self-control impairing articulation and pursuit of the optimal choice. Given the subject is guided by narrow self-interest, choice depends on individuals' subjective preferences and the constraints they face, especially income and the relative price of each alternative (Senstat and Constantine 1975; Screpanti 2000; Davis 2011). These attributes are deemed to equip individuals to pursue and realise the best possible outcomes for themselves in market interactions. The market is assumed to provide relevant information and incentives to augment choices, thereby maximising individual utility and social welfare (the sum of individual utilities) (Sen 1977).

This formulation of *Homo Economicus* thereby shows the separation of neoclassicism from psychology (Sent 2004). While neoclassicism may informally address individuals' 'preferences' as psychological phenomena, as suggested above, the term embodies a formal ordering relation fashioned to enable equilibrium analysis. Rational 'choices'

depend on ‘well-behaved’ preferences, guaranteeing downward-sloping demand and upward-sloping supply curves with parametrically varied prices through fostering a functional representation of individuals. Reasoning is assumed to be instrumentally-oriented, while individuals are held to act as if drawing on mathematical logic: each actor is assumed to form correct beliefs about her/his environment, as well as her/his own and others’ behaviour, and select actions satisfying her/his preferences (Davis 2011). Thus, rationality and individuality are co-defined according to the logical-mathematical properties of equilibrium analysis.

The behavioural challenge: the boundedly rational individual

Taking exception to assumptions required to rationalise the utility function interpretation of preference ordering and the self-interested actor, many critical accounts have challenged the realism and logical consistency of this neoclassical formulation of hyper-rationality (e.g. Hollis and Nell 1975; Sen 1977, 2002; Hewitson 1999; Davis 2011). Nevertheless, neoclassical axioms have appeared largely impervious to falsification, partly due to increasing formalism in the tradition encouraging escalating forms of reductionism in general equilibrium theory, while hyper-rationality axioms have been taken as self-evident (Dow 2013). More recently, a different challenge has arisen from confrontations between pure theory and applied work, with hyper-rational presuppositions and the formalist structure of neoclassicism increasingly confronted by empirical counter-evidence. As Blaug (1992) demonstrates, anomalies arising from such evidence were initially rejected as random micro-level perturbations, though the mainstream has increasingly sought to theorise such deviations from the predictions of hyper-rational behaviour and the pertinence of socioeconomic institutions in shaping this behaviour and aggregate outcomes, as examined below (Wolff and Resnick 2012). This shift partially reflects the pervasive influence of BE within the mainstream of the discipline.

BE amalgamates psychological insights with the ‘tractability’ and ‘parsimony’ of neoclassical methods (Rabin 1998: 12) to construct more empirically-realistic models of behaviour. Inspired by empirical evidence countering hyper-rationality axioms, alongside articulations of the mind as constituting an information-processing device arising from cognitive psychology, BE formalises and tests psychological predictions and draws

conclusions about decision-making processes (Heidl 2016). Demonstrating that neoclassicism accords little emphasis to constructing models based on realistic assumptions, nor the institutional and socio-cultural parameters affecting decision-making, behaviouralism identifies the observable realities of decision-making as diverging from the precepts of hyper-rationality (Altman 2005, 2015). Empirical accounts – involving the study of real-world decision-making and the neural gaze of laboratory brain-scanning equipment – reveal that, while ‘irrational’ from the neoclassical perspective, such divergent behaviour is pervasive (*e.g.* Fehr *et al.* 2002; Henrich *et al.* 2005). In turn, this behaviour is theorised as depending on psychological, cultural and biological considerations affecting and constraining choice.

Table 1: A simplified model of dual process

	Automatic System	Reflective System
Processes	Fast	Slow
	Parallel	Serial
	Automatic	Controlled
	Effortless	Effortful
	Associative	Rule-governed
	Slow-learning	Flexible
	Emotional	Neutral

Source: Adapted from Kahneman (2003a: 1451).²

The entry-point for considering BE as a form of critique and alternative theory of behaviour may be summarised in the now-pervasive notion of simultaneous ‘dual processes’ operating in the brain and shaping decision-making. Although not a unified theory, this heuristic has been introduced to represent the foundations of BE research (Sloman 1996;

² See Kahneman (2003a: 1451) and Thaler and Sunstein (2008: 21) for more extensive accounts of this heuristic, which also feature additional reflections on the cognitive processes associated with dual process. An excellent critical reflection on the philosophical and its political foundations of this framework is presented in Heilmann (2014).

Kahneman 2003b). Actors' mental capacities are conceptualised as aligned with properties of two systems: one 'automatic' and one 'reflective' (Thaler and Sunstein 2008: 21). In turn, decision-making is theorised as a process in which each system concentrates on addressing differing cognitive and deliberative tasks. As depicted in Table 1 above, the *automatic* system is formulated as rapid, instinctual and emotional and, thus, capable of managing straightforward precepts and stimulation beyond contemplation. By contrast, the *reflective* system better handles concepts and deliberative behaviours considered rule-bound, deductive and logical, as it is described as controlled, effortful and neutral (Roberto and Kawachi 2015). While neoclassicism assumes decision-makers possess complete access to and utilise the latter, empirical evidence indicates that the former underpins decision-making in practice. BE thereby conceptualises the biases, heuristics and framing effects affecting decision-making, explored below, as grounded in information being simplified and distorted by the automatic system, leading actors to expedient, albeit not always prudent, choices (Heilmann 2014).

To some extent, BE models utilising this heuristic depart from the same behavioural assumptions informing neoclassicism, given the individual is assumed to be purposefully seeking to ascertain the optimal choice based on her/his preferences across the available options (Laibson and List 2015). Yet, rather than a logico-deductive theory of choice giving rise to *Homo Economicus*, individuals are described as characterised by *bounded-rationality*. Initially proposed by Simon (*e.g.* 1955, 1956), bounded-rationality proposes a foundational challenge to the axioms of hyper-rationality. Rather than decision-making reflecting consistent application of knowledge, reflection and deliberation (comprising the 'reflective system'), Simon contended that insufficient knowledge, sub-optimal cognitive practices and the imperative for regular hasty decision-making led to bounded rationality (Velupillai and Kao 2016). He utilised this conception to explain actors pursuing short-term over long-term objectives, making impulsive judgements and allowing social norms and past habits to influence decisions – that is, decision-making operates according to the 'automatic system' (Kahneman 2003b). While such behavioural 'aberrations' transcended the purview of neoclassicism, the BE tradition initiated by Simon centred them as an object of analysis (Velupillai and Kao 2016).

The 'prospect theory' of Kahneman and Tversky (*e.g.* 1973, 1974) subsequently sought to deepen this conception of real-world decision-

making by uncovering the processes engendering the bounded-rationality identified by Simon. Utilising detailed empirical studies, Kahneman and Tversky isolated various *heuristics* (or necessary shortcuts) as steering decision-making by rendering complex problems more manageable, thus inhibiting hyper-rationality. Specifically, through streamlining matters of judgement (evaluating options, including estimating probabilities) and choice (choosing between those options), heuristics constitute context-specific rules of thumb facilitating choices without thorough deliberation (Kahneman 2003b). These heuristics seem intuitively judicious to optimising individuals and enable practical strategies to manage given situations: ‘a good heuristic provides fast, close to optimal, answers when time or cognitive capabilities are limited’ (Camerer and Loewenstein 2004: 11). Yet, they simultaneously infuse systematic *biases* into decision-making and, consequently, may engender sub-optimal outcomes relative to those realised by hyper-rational individuals. BE thus presents individuals as lacking complete information and having recourse to biases and heuristics to enable swift and satisfactory decision-making, albeit leading to choices differing to those anticipated by neoclassicism.

Such insights would be relatively unproblematic for neoclassicism so long as behavioural deviations remained idiosyncratic and would, on average, cancel each other out. The problem, however, arises from BE’s universalist assumption that utility maximisation is hindered by systematic biases underpinning decision-making (Ariely 2008). As Thaler and Sunstein (2003: 176) explain:

People do not exhibit rational expectations, fail to make forecasts that are consistent with Bayes’ rule, use heuristics that lead them to make systematic blunders, exhibit preference reversals (that is, they prefer A to B and B to A) and make different choices depending on the wording of the problem.

That is, the nature of human cognitive processes inherently engenders ‘irrationality’. In turn, as identified in DellaVigna (2009), BE challenges the assumption of expected utility maximisation by positing three qualities of individuals that differ from those found in *Homo Economicus*. First, while neoclassicism assumes individual preferences concerning future plans remain identical at different points in time, BE conceptualises *non-standard preferences*. For instance, there is a tendency for decision-makers to put disproportionate weight on present over future concerns, as in the case of individuals choosing to consume income at any given moment despite committing to long-term savings

plans (O'Donoghue and Rabin 1999; Benartzi and Thaler 2007). These are said to be systemic present-biased preferences. Second, subjects are conceptualised as holding *non-standard beliefs*. This may include, for instance, systematic overconfidence, whereby subjects overestimate their own capabilities and undervalue the probability of adverse phenomena and time required to accomplish projects, such as individuals naïvely overestimating their commitment to future gym attendance (Griffin and Tversky 1992; Camerer and Lovallo 1999). Third, subjects are held to be informed by *non-standard decision-making*, whereby a decision associated with equivalent opportunity-costs and logic can be made differently depending on its *framing* in experiments. In contrast to the neoclassical assumption that behaviour is determined by individuals weighing-up the costs and benefits of alternatives, BE regards individuals as having inconsistent standpoints on risk and responding to difficulties according to their framing. This recognises that decision-making is affected by context (Tversky and Kahneman, 1986, 1991).

These features of BE present an important challenge to the usefulness of neoclassicism grounded in hyper-rationality axioms. They also offer an alternative conception of the subject to that manifest in *Homo Economicus*. The contrast is encapsulated by Thaler and Sunstein's (2008: 24) dichotomy between the 'econs' of neoclassicism and the 'humans' that inhabit the real world. The former are assumed to be hyper-rational utility maximisers replete with a given utility function, whereas the latter are shambolic in their decision-making – more akin to the Homer Simpson 'lurking somewhere in each of us' than to *Homo Economicus*. Even when strategic and purposeful, humans make repeated miscalculations in pursuing their objectives and are influenced by external factors, especially comparison against social norms. In contrast to Veblen's (1898: 398-9) sardonic depiction of the neoclassical subject as 'a lightning calculator of pleasures and pains, who oscillates like a homogeneous globule of desire of happiness under the impulse of stimuli that shift him about the area but leave him intact', BE depicts *Homo Sapiens* as failing to cognitively approximate *Homo Economicus*.

'As-if' behavioural economics

The analysis so far has depicted neoclassicism and behaviouralism as apparent antinomies. BE demonstrates the *empirical unreality* of

neoclassical axioms of hyper-rationality and, in turn, rejects *Homo Economicus* in favour of alternative accounts of decision-making. Rather than framing individual choice through a cost-benefit calculus enabling utility maximisation, BE demonstrates that individuals have recourse to myriad heuristics and biases, such that boundedly-rationality is the norm in reality. This inference has gained popularity within the economics profession and the broader social sciences, partly because it seems to advance beyond an earlier orthodoxy grounded in neoclassicism and libertarian philosophy. For instance, Conly (2012: 8) concludes that Mill ‘failed to adequately reckon with human psychology, as we now know it to be’, while ‘the existence of cognitive deficits does suggest a need for different sorts of legislation [...] coercive paternalism, for laws that force people to do what is good for them’. Similarly, Kahneman (2011: 145) asserts that ‘[d]emocracy is invariably messy, in part because the availability and affect heuristics that guide citizens’ beliefs and attitudes are invariably biased, even if they generally point in the right direction’. More broadly, the award of the 2017 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel to Richard Thaler for his contributions to behavioural economics continues popular recognition – following Herbert Simon (1978), Daniel Kahneman (2002) and Robert Shiller (2013) – for the contribution of behaviouralists to enriching mainstream economics. Such examples highlight how unrealistic hyper-rationality axioms are deemed to have been cast aside in BE.

Nevertheless, a fledgling critical literature has contested the contribution of behaviouralism to greater empirical realism.³ Specifically, it has demonstrated that BE is underpinned by an *incongruous methodological configuration*, simultaneously promulgating the need for more realistic accounts of behaviour while retaining unrealistic assumptions associated with hyper-rationality. The foundations of this critique may be best understood according to the Lakatosian yardstick of ‘verification’ (Dow 2013). Within the Lakatosian philosophy of science, a ‘theory’ constitutes a succession of gradually established marginally dissimilar

³ The foregoing examination of continuities running through the neoclassical and behavioural traditions is not intended as an exhaustive discussion of such elements, nor the critical political economy literature in which such matters are addressed. Rather, the examples considered are indicative of the focus on methodological incongruities informing such critiques of BE. These are the most important considerations for the purposes of the present analysis. For critiques of a range of the assumptions underpinning the behavioural canon, see Berg and Gigerenzer (2010); Davis (2011, 2015); and Heidl (2016).

hypotheses and experimental techniques sharing a *common hard core* (the 'research programme'). Practitioners operating within research programmes protect this core from falsification through devising a *protective belt of auxiliary hypotheses* (Caldwell 1991; Backhouse 1998). The Popperian methodological concern with the truth or falsity of a paradigm is supplemented by investigation into whether a programme is 'progressive' or 'degenerating'. The former embodies development, prediction of novel facts and precise predications; conversely, the latter is marked by minimal progress, such that its auxiliary belt does not engender novel and verifiable predictions.

From this perspective, BE is regarded as a *degenerative* programme. While its *modus operandi* has been the introduction of psychological insights to explain evidence of behaviour deviating from that assumed by neoclassicism, BE has largely incorporated such insights *within* the standard hyper-rational choice framework. Thus, behaviouralists Camerer and Loewenstein (2004: 3) highlight how BE centres on 'the conviction that increasing the realism of the psychological underpinnings of economic analysis will improve the field of economics on its own terms' through 'generating theoretical insights, making better predictions of field phenomena, and suggesting better policy'. In turn, consideration of psychological factors does not necessitate 'a wholesale rejection of the neoclassical approach to economics based on utility maximization, equilibrium, and efficiency'. Such reasoning is mirrored in the *2015 World Development Report* ('the Report'; World Bank 2015). While stating its primary objective as being 'to inspire and guide researchers and practitioners who can help advance a new set of development approaches based on a fuller consideration of psychological and social influences' (p. 2), the *Report* qualifies that 'the new tools based on this full consideration of human factors do not displace existing policy approaches based on affecting self-interested personal incentives; rather, they complement and enhance them' (p. 3). Such examples do not point to a new Lakatosian research programme, so much as a marginal modification of the protective belt of auxiliary hypotheses: the neoclassical hard core of hyper-rationality axioms is retained, with constraints on hyper-rational choice elucidated through more complex representations of 'rationality' or otherwise as 'irrationality' (Dow 2013).

The result is an incongruous methodology. On the one hand, BE holds that the empirical lessons of psychological experiments will advance economic analysis when filtered through models allowing for

behavioural phenomena diverging from hyper-rationality. On the other hand, such modelling often does not pursue psychological realism, depending on Friedman's (1953) instrumentalist 'as-if' defence to justify *increasingly unrealistic* formulations of psychological choice as resolving more elaborate constrained optimisation problems than the simpler neoclassical models they intended to transcend (Berg and Gigerenzer 2010). That is, BE has retreated from exploring and theorising *actual* decision-making processes. Instead, it regularly conforms to Friedman's defence of unrealistic models endowed with additional psychological parameters, assuming that individuals behave *as if* they are resolving intricate constrained optimisation problems and modelling cognitive limitations in the objective function or constraint set. Indeed, Jolls *et al.* (1998: 1471) mirror Friedman in asserting that 'economics should not be judged on whether the assumptions are realistic or valid, but rather on the quality of its predictions'. The result of this seemingly inconsistent methodological posture may be labelled 'as-if behavioural economics' (Berg and Gigerenzer 2010).

To clarify, much behavioural modelling generalises otherwise-familiar neoclassical accounts, incorporating novel parameters in the objective function or constraint set to characterise psychological phenomena (Pesendorfer 2006). Reflecting Kahneman's (2003a: 1469) concession that it largely 'retain[s] the basic architecture of the rational model, adding assumptions about cognitive limitations designed to account for specific anomalies', BE *expands* the scope of the neoclassical utility function to represent heuristics through incorporating psychology. It consequently aims to accord increased theoretical exactitude to notions of bounded-rationality and their consequences for decision-making outcomes. Thus, as Camerer (1999, n.p.) notes, 'psychology provide[s] a way to model bounded rationality which is more like standard economics than the more radical departure Simon had in mind...[and]...incorporate this kind of psychology into [neoclassical] economics.'

The result is that potentially valuable behavioural insights are lost by maintaining that actors optimise more-or-less complex utilities. For example, in conceptualising inter-temporal choice, models of time inconsistency fuse discounting parameters with non-exponential weighting schemes while preserving the assumption that individuals maximise a time-separable utility function (*e.g.* Laibson 1997). Similarly, social preferences theorists affix parameters weighting individuals' concern for gaining more or less than others to an otherwise-neoclassical

utility function (*e.g.* Fehr and Schmidt 1999). Such models retain the neoclassical supposition that all behaviour is oriented around constrained optimisation – decision-makers systematically explore every potential consumption sequence, calculate the weighted sum of utility terms for each and choose that with the highest weighted utility. Incorporating psychological elements thereby simply produces more complex optimisation problems to solve (Gigerenzer and Selten 2001). Given the lack of empirical evidence that individuals actually solve such problems or more substantive empirical accounts of decision-making processes, BE thus relies on arguments approximating Friedman’s ‘as-if’ doctrine (Gigerenzer and Berg 2010).

Such accounts of the incongruous methodological configuration of BE offer a pertinent critique of its claims to greater empirical realism. The progression of behaviouralism beyond neoclassicism remains limited by the former’s acceptance of the latter’s rational-choice axioms as their hard core, while only modifying its auxiliary hypotheses through amending its models to take account of restricted cognitive limitations and asymmetrical preferences.

The theoretical problematic of neoclassical humanism

Because BE has largely sought to *modify* rather than *reject* neoclassical models of hyper-rationality based on constrained optimisation, its capacity to utilise insights from cognitive psychology and social experiments is restricted. This tempers its claim to be constructing more realistic descriptions of individuals’ cognitive processes. It would, therefore, be tempting to dismiss behaviouralism on Lakatosian grounds as yet another string in the epistemological bow of neoclassicism and, instead, turn to alternative schools adopting different stances on the nature and meaning of theory and evidence. Heterodox economics approaches such as Marxism, old institutionalism and post-Keynesianism offer that option because they employ different, more holistic methodological and conceptual frameworks and, thus, constitute competing Lakatosian research programmes.

While not disregarding this need for alternatives, the political economic significance of behaviouralism’s retention of hyper-rationality extends beyond methodological incongruity. Rather than focussing on the apparent fissure between pursuing greater realism and preserving

neoclassical axioms as a methodological *contradiction per se*, the remainder of this article argues this is actually *consistent* with the *theoretical problematic of neoclassical humanism* (‘theoretical problematic’) that structures BE. That is, behaviouralism remains within a problematic in which the interests of hyper-rational and autonomous human subjects are to be reconciled through markets.

From this perspective, beyond the ‘as-if’ critique outlined above, retention of hyper-rationality alongside more realistic descriptions of psychological processes has critical *performative* implications. Specifically, as demonstrated below, BE engages in the subjectification of *Homo Economicus* by *reworking*, rather than rejecting, the theoretical problematic. Explaining the failure of markets to perform as predicted by neoclassicism as stemming from the bounded-rationality characterising decision-making in reality, BE *pathologises* the latter as an irrational divergence from *Homo Economicus*. It thereby fosters a functionalist explanation of non-market institutional arrangements to buttress markets through correcting such deviations. The remainder of this article is, thus, not concerned with the empirical truth or logical consistency of behaviouralism *per se* – that is, its failure to represent the *truth* of the subject. Rather, it examines how BE discursively produces and accords authority to a particular hyper-rational subjectivity as the normative ‘microfoundation’ for harmonious economic and social order.

To appreciate this position, it is first necessary to comprehend the key philosophical tenets of theoretical humanism (‘TH’) and its significance in structuring neoclassicism. Initially arising with the Renaissance and developing alongside the transition from Feudalism to Capitalism in Europe, TH constitutes a post-Enlightenment philosophical orientation informing numerous social theories (Rabil 1988a, b; Wolff and Resnick 2012). It reflects a focus on individual human subjects as the fundamental driving force for and essence of society and thought. In contrast to extant religious thought, human subjects thereby supersede the central ontological position (*logos*) previously occupied by God as the *sine qua non* for deliberation and practice (Copson 2015). From this foundation, TH is underpinned by two theoretical conceptions: (i) the human subject and (ii) social reconciliation (Madra 2016). With regard to the former, as Ruccio and Amariglio (2003: 48) explain:

Placing humans at the center [*sic*] of schemas of progress and history and meaning is what distinguishes theoretical humanism, as the human subject is thus the beginning and ending point of all movement

from the growth of knowledge [...] understood as undertaken by, for, and through human subjectivity) to the transformation of the natural world (through science and technology oriented to human desires and ends, such as happiness).

This *subject* is an essentialist conception of a *rational, centred and autonomous self-consciousness* possessing, or capable of possessing, complete knowledge of its 'true' preferences formed largely in isolation from external influence, and largely unbounded computational capacities to purposefully utilise whatever measures are required to realise her/his individual ends (Screpanti 2000; Madra 2016). Thus, the TH subject is both an epistemological and ontological entity: enjoying the cognitive capacity to *recognise* its interests and *understand* how to satisfy them, and benefiting from an intentional 'agency' to purposefully attain these ends through all means necessary (Madra 2016).

The corollary of this subjective essence is the utopian vision of an accommodating *harmonious and contradiction-free social order* reconciling the aggregation of subjects' interests (Garnett 1994; Madra 2016). Regardless of whether this order is realised, the concept remains a reference-point to juxtapose actually-existing states as imperfect approximations of the ideal. Within neoclassical theory, for example, this notion of *social reconciliation* is most discernible in notions of equilibrium, which conceptualise an ideal social order grounded in the unimpeded competitive interaction of rational and autonomous subjects (*i.e.* individuals who possess consistent preferences reflecting their welfare and choose according to these preferences). If such subjects only interact through markets, social relations are reduced to exchange relations, such that competitive equilibrium is underpinned by *Homo Economicus*. Reconciliation of actors' various interests through the market thereby ensures that hyper-rational individuals, in pursuing their own objectives, enable all others to pursue theirs (Garnett 1994; Screpanti 2000). Thus, while not necessarily representing reality, reconciliation is normatively founded on an essentialist subject. This, in turn, legitimises free markets as a theoretically efficient apparatus to secure Pareto-optimal states of equilibrium (Shaikh 2016).

These two presuppositions – the essentialist hyper-rational subject and harmonious reconciliation of these subjects' interests – inform the constitutive problematic of neoclassicism. In contrast to the structuralism informing Keynesianism and some variants of Marxism (see Wolff and Resnick 2012), TH structures this problematic within neoclassicism as

studying the conditions of existence for reconciling individual and collective rationality (Madra 2016). That is, neoclassicism is concerned with conceptualising how reconciliation (for instance, equilibrium) of the diverse demands of hyper-rational individual subjects may be realised.

This problematic has taken heterogeneous forms within neoclassicism. For instance, while the invisible-hand constitutes one formulation – assuming that equilibrium maximising social welfare is secured through subjects' activities being channelled by competitive markets and private ownership – its form varies across the tradition (Aydinonat 2009; Wolff and Resnick 2012). This divergence is evident amongst early neoclassical practitioners where, for example, Walrasians constructed general equilibrium models bottom-up from individual agents, whereas the British tradition of utility calculus articulated a partial equilibrium approach utilising representative agents and market-level analysis (Hennings and Samuels 1990; Ingraio and Israel 1990). In a static equilibrium and price-adjustment formulation, the former conceptualises competitive markets as an auction, such that the invisible-hand equates to a hypothetical auctioneer. Conversely, the Marshallian School favoured an evolutionary interpretation of the problematic, whereby competition is theorised as a quasi-Darwinian selection process clearing out inefficient agents, with the 'invisible hand' reflected in this mechanism (Koppl 1992; Roncaglia 2005). Thus, while formulating the problematic differently, each neoclassical school remains structured around investigating reconciliation of individual and collective rationality (see Madra 2016 for a survey).

The subjectification of *Homo Economicus*

BE reworks the theoretical problematic by continuing to posit human rationality as the *logos* of economic processes, while examining how and why societies fail to secure reconciliation through markets. It does not reject *Homo Economicus*, but *reinterprets* it as the normative ideal against which reality is compared. The behavioural criticism of neoclassicism for failing to realistically describe decision-making is concerned not with the anthropological *truth* of *Homo Economicus*: rather it introduces psychological insights to explain why subjects in reality deviate from this ideal. Behavioural discourse thus contributes to the discursive *subjectification* of *Homo Economicus* as the normative

microfoundation for functioning markets. Bounded-rationality, arising from systemic biases and heuristics, is *pathologised* as ‘anomalous’ relative to this ideal and articulated as explaining why markets fail to perform as anticipated by neoclassical models. This section explains each of these features in BE and the pitfalls therein.

Subjectification and economic discourse

In challenging the notion of the subject as ontologically given, Foucault (1982, 2003) conceptualises ‘subjectification’ as entailing the process whereby a discursive formation (in this case, BE theory) produces and accords authority to a particular subjectivity as the essence of economy and society (Butler 1995; Callari 2011). That is, while the TH discourse previously examined *presupposes* hyper-rational subjects as a site of epistemological certitude, Foucault’s deconstructs the category of ‘subject’ as an entry-point by examining how it is produced in a particular form through a structured set of discursive processes and practices (Callari 2011). By questioning the naturalised ontological status of *Homo Economicus* as the *logos* for economic processes, Foucault’s insights enable evaluation of how this subject is articulated as an authoritative figure within BE.

To understand the logic of subjectification through economic discourse, it is necessary to briefly consider Foucault’s approach to the discursive construction of semantic fields. Interrogating the conditions engendering certain objects of analysis, Foucault (*e.g.* 1972, 1980) held that, as part of the signification system producing and reproducing meaning and identity, discourse exerts power over knowledge by serving as a locus of truth. In turn, it legitimises authority and social control. Rather than being given, analytical objects arise in a contextually-specific form through a nexus of interrelated institutional and discursive practices, such as behavioural patterns, norms and forms of classification (Foucault 1991). Concepts (*e.g.* ‘student’) are established not simply through their pronouncement as such. Rather, the articulation of mutually-reinforcing practices enables the referenced concept to materialise as a prospective *given*, thereby buttressing specific formulations (for instance, labelling a particular contribution in class as ‘disruptive’) and points of authority (*e.g.* ‘teacher’) (Mehta 2013). These givens structure relationships accordingly: operating beyond actors’ conscious deliberation (*e.g.* that of

the teacher, student or University attesting ‘disruptiveness’) and, instead, functioning as *givens* (outside questioning) and points of *authority* (authorising the practices and relationships) (Callari 2011).

Concomitantly, economics constitutes a modernist discourse abstracting the human experience and demarcating it into distinct fields of inquiry: ‘history’, ‘society’ and myriad subdivisions (*e.g.* ‘the economy’) and combinations (*e.g.* ‘economic history’) (Screpanti 2000). These fields are constitutive of a Foucauldean process of subjectification because they engender *subject-effects*: insofar as they are conceptualised and come to be known as products of a subject (or collectivities thereof), these fields and the discursive protocols through which they are constituted make the subject *visible* and imbue that vision with epistemic *authority* (Callari 2011). This combination of visibility and authority discursively constructs the subject as the essence of society and economy.

Reworking the theoretical problematic of neoclassical humanism

Examined through this Foucauldean lens, BE subjectifies *Homo Economicus* as a *normative ideal* for cognition and behaviour, and a *potentially realisable* subject to be procured through policy measures. To develop this case, the remainder of this article examines the conceptual and discursive logic and terms distinguishing boundedly-rational subjects from their hyper-rational counterparts. This shows the values that are embodied in giving normative priority to the latter. Specifically, the articulation of *Homo Economicus* as the authoritative subject in BE is underpinned by the identification and pathologisation of ‘anomalies’. Boundedly-rational decision-making is derided as both ‘irrational’ and requiring correction to become congruent with the hyper-rational ideal. That is, BE discourse determines and hierarchically organises categories of ‘(hyper-)rationality’ and ‘bounded-rationality’ based on assumptions about what constitutes ‘normal’ or ‘desirable’ decision-making characteristics (Mehta 2013).

Comprehending this subjectification initially requires grasping how BE remains oriented to the theoretical problematic. BE constitutes one of a plethora of contemporary mainstream traditions which has, since the

1970s, responded to the crisis in Walrasian general equilibrium analysis⁴ – formalised in Arrow-Debreu (A-D) general equilibrium models – through rehabilitating and reworking the two presuppositions underpinning the problematic explored above.⁵ As Madra (2016) demonstrates, such efforts have taken one of two forms: i) *isolating* specific axioms in A-D and constructing models derived from these weakened assumptions without emasculating the neoclassical humanist presuppositions; or ii) *reformulating* the problematic, through either integrating *non-Walrasian* neoclassical schools or incorporating concepts and methods from *other disciplines*. These two steps have enabled the production of concepts explaining why the requisite conditions for the operation of the ‘invisible hand’ cannot be secured (markets are insufficient) *or* why its realisation necessitates measures instituting favourable conditions (insufficient markets exist).

Numerous examples are evident in contemporary mainstream economics. Concepts associated with *information failures* (moral hazard and adverse selection), for instance, relax assumptions of perfect information to rationalise phenomena such as unemployment (*e.g.* Shapiro and Stiglitz 1982; Bowles and Gintis 1990). Similarly, notions of *transaction costs* loosen assumptions of the smooth formulation and enforcement of contracts enabling perfect competition (*e.g.* Coase, 1937; 1960). These conceptual innovations, in turn, inform functionalist formulations of ‘institutions’ as arising to resolve the inability of the price-mechanism to produce unique Pareto-efficient equilibrium outcomes. While taking many forms in such accounts, by theorising the conditions unsettling or inhibiting markets from conforming to the predictions of perfect competition and how (extant or potential) institutions may resolve such tensions, the theoretical problematic is thus not abandoned but *reproduced* in novel forms.

BE distinctively reformulates this problematic through the notion of *bounded-rationality* discussed above. In contrast to the representation of the subject in the A-D model, BE relaxes the assumption that actors

⁴ For consideration of this crisis in general equilibrium theory since the 1970s and some of the ensuing responses, see *e.g.* Bell and Kristol (1981); Hahn (1984); Arrow (1987); Katzner (1999, 2010); Bowles and Gintis (2000); Ackerman (2004); and Madra (2016).

⁵ For surveys of such mainstream approaches, see *e.g.* Colander *et al.* (2004); Colander (2006); Milonakis and Fine (2008); Fine and Milonakis (2009); Chernomas and Hudson (2016); Madra (2016); Crespo (2017); and Mirowski and Nik-Khah (2017).

exercise hyper-rationality in decision-making. Concomitantly, this failure to conform to *Homo Economicus* is used to rationalise why individuals and, in turn, markets fail to function as predicted by neoclassicism. Analogous to concepts such as information failures and transaction costs, BE posits bounded-rationality as *another* reason individual and aggregate rationality fail to reconcile through markets, engendering functionalist explanations of complementary institutional arrangements to buttress markets through enabling individuals to make choices *as if* they were *Homo Economicus*, as explored further below.

Specifically, BE reworks the theoretical problematic by investigating conditions for social harmony premised upon hyper-rational subjects, *given that subjects are assumed to be characterised by bounded-rationality in reality*. Recall from above that the problematic is underpinned by a notion of the subject as an autonomous, self-transparent and rational self-consciousness cognisant of her/his true, welfare-enhancing preferences and capable of translating these into choices within markets. BE departs from this conception to the extent that it acknowledges actual decision-makers fail to approximate *Homo Economicus*. However, BE continues to accord this ideal central ontological status as *logos* determining the economy as a whole in that deviations from it are held responsible for markets failing to function effectively in practice. Thus, Akerlof and Shiller (2015: 164) summarise their objective as being to analyse ‘the role of markets when people have weaknesses, so markets are not efficient’. Similarly, Camerer and Fehr (2006: 47) contend that while behaviour deviating from hyper-rationality may occlude welfare-maximising outcomes, the presence of sufficient subjects approximating *Homo Economicus* ‘may cause aggregate outcomes to be close to the predictions of a [neoclassical] model that assumes that everyone is rational and self-regarding.’

Illustrative examples abound in the behavioural literature. Consider, for example, Thaler and Sunstein’s (2008) study of high rates of household indebtedness and financial instability in the Global North. They contend that these have arisen partly due to individuals’ self-control problems (favouring immediate gratification over the long-term costs of impulsiveness) and difficult-to-comprehend information, resulting in injudicious consumption and investment decisions.⁶ Similarly, Akerloff

⁶ More explicitly, in referring to mortgage markets, Thaler and Sunstein (2008: 134) argue that:

and Shiller (2009) appropriate Keynes' 'beauty contest model' of expectations formation to attribute persistent global macroeconomic problems to financial volatility stemming from boundedly-rational self-fulfilling beliefs. In both cases, while *Homo Economicus* is positioned as the normative microfoundation for properly functioning markets, divergence from this subjective ideal due to bounded-rationality is held to compromise such order.

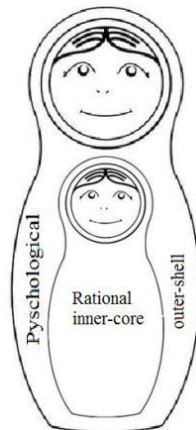
In this respect, behaviouralism exists symbiotically with neoclassicism, as it is primarily organised around explaining failures of the latter. Rather than laying the foundation for a more complex theory of subjectivity *per se*, BE introduces psychological factors as *means* to explain why individual behaviour does not correspond to *Homo Economicus* and, concomitantly, how this prevents markets from functioning effectively. For example, although appreciating Berg and Gigerenzer's (2010) critical assessment, discussed above, that BE models assume that agents solve complex optimisation problems *as if* they were *Homo Economicus*, Rabin (2013) defends modelling behaviour as stemming from such optimisation problems while recognising contextual influences on preferences. For Rabin, it is problematic to assume that all deviations from neoclassical models stem from agents' incapacity to solve complex optimisation problems. In some situations, agents may be *theoretically* capable of solving such problems and try to do so, yet commit systemic errors in doing so due to their underlying biases and heuristics. As Rabin (2002: 611) elsewhere articulates, the objective of BE thereby becomes utilising empirical methods and experimental evidence to identify which departures from neoclassicism are necessary through categorising such departures in a manner which enables theorists to distinguish 'as precisely as possible where and how [neo]classical economic assumptions go awry' (see also Camerer and Loewenstein 2004).

Such reasoning, whereby observed deviations from hyper-rationality are conceptualised as resulting from mistakes, effectively presupposes the neoclassical optimising model of decision-making as correct at one,

When markets get more complicated, unsophisticated and uneducated shoppers will be especially disadvantaged by the complexity. The unsophisticated shoppers are also more likely to be given bad or self-interested advice by people serving in roles that appear to be helpful and purely advisory. In this market, mortgage brokers who cater to rich clients probably have a greater incentive to establish a reputation for fair dealing. By contrast, mortgage brokers who cater to the poor are often more interested in making a quick buck.

normative level. As explored below, BE thus seeks to demonstrate, before offering a psychological explanation of, the unreality of neoclassical hyper-rational presuppositions in practice through pathologising cognitive biases and heuristics as perceptual and inferential mechanisms linked to subjects (Gilovich *et al.* 2002; Pesendorfer 2006). Rather than being employed to conceptualise an emotionally-whole actor, BE utilises psychology to comprehend the factors hindering actors' capacity to approximate *Homo Economicus*. Psychology is thereby reduced to an arrangement of dynamics *affecting* behaviour by *interfering* with hyper-rationality and articulating bounded-rationality as an individual failure to process knowledge (Pykett 2012; Wilkins 2013). The result is an epistemology grounded in an implicit dualistic model of the hyper-rational subject, as the normative microfoundation for functioning markets, effectively constrained inside an obstructive psychological shell (Sugden 2015; Infante *et al* 2016) – depicted in Figure 1 as a *behavioural matryoshka*:

Figure 1: The behavioural matryoshka



Source: Author's design, adapted from Infante *et al* (2016).

Coterminous with the TH subject, the 'inner' actor is delineated as the essential locus of rational human identity and foundation of normative authority concerning its interests and objectives. Assuming its hyper-

rationality as given, BE does not seek to investigate the psychology of this subject; instead continuing to assert that '[t]here is basically only one way to be rational' (Laibson 2002: 22). Conversely, psychological mechanisms inducing deviations from hyper-rational decision-making are relegated as constitutive of the 'outer-shell' inhibiting the agent in the rational 'inner-core' from acting according to its true preferences, 'as if the individual's psychology were an external force subverting the will of the true self' (Sugden, 2015: 584). Thaler and Sunstein (2008), for example, articulate psychological limits – including time, cognitive capacity, information availability and unconscious emotional drives – as hindering hyper-rational decision-making. Boundedly-rational decision-making is, thus, conceptualised as reflecting interactions between the self-governing reasoning of the rational inner-agent and constraining properties of the psychological outer-shell. However, the inner-agent remains the normative authority with regard to matters of judgement and preference, such that BE theory and policy seeks to comprehend and reconstruct the preferences of this fictional agent, as examined below.

Pathologising the 'deviant' subject

This reworking of the theoretical problematic through investigating the psychological factors hindering the realisation of *Homo Economicus* and, thus, the effective functioning of markets is underpinned by the discursive *pathologisation* of boundedly-rationality. BE positions *Homo Economicus* as the normative microfoundation for functioning markets and introduces psychology to explain their failure to perform as predicted by neoclassicism. In turn, it *subjectifies* the former as the *norm* from which bounded-rationality deviates. That is, while recognising that this hyper-rational subject does not *describe* human rationality in reality, through revising the theoretical problematic, behaviouralism discursively prioritises it over explanations of actual behaviour as an *ideal*. Moreover, by treating it as a *potentially-realizable* state, it provides a rationale for policies seeking to change individuals' behaviour to more 'rational', market-conforming forms.

The normative priority accorded to *Homo Economicus* is informed by a methodological delineation between the *descriptive* and *normative* orientations of BE. Within neoclassicism, departures from hyper-rationality are deemed anomalous and usually self-correcting, such that it

is unnecessary to abandon it as a description of decision-making or the assumption that markets are generally efficient (Shaikh 2016). Conversely, BE introduces the novel – from the perspective of neoclassicism – differentiation between ‘descriptive’ and ‘normative’ theories of choice (Heukelom 2011). Questions of ‘normativity’ have been identified in neoclassicism at least since Keynes’ (1973) *Scope and Method of Political Economy*: the ‘normative’ addressed ethical questions, while the ‘positive’ embodied the empirical basis of a value-free discipline (Davis 1998). In contrast, BE defines ‘normative’ as the rubric defining how one *ought* to behave to be ‘rational’ (Heukelom 2011, 2014). Used to ‘characterize rational choice’ (the universal rules of rational decision-making accorded by actors’ reasoning), the term ‘normative’ is thereby distinguished from ‘descriptive’ accounts of ‘actual choices’ (empirically-grounded theories of actors’ decision-making) (Thaler 2000: 138).

From this distinction, BE articulates the characteristics of ‘rational’ behaviour in accordance with *Homo Economicus*, while recognising that individuals must be described as generally acting otherwise (Davis 2011). BE theorises decision-making by utilising an individual value function expressed according to two scales. The first assigns weights to diverse probabilities, whereby individuals overreact to minor probability events and underreact to medium and large ones; and the second ascribes subjective value to prospects relative to choice reference points, whereby individuals are risk-averse about gains and risk-seeking regarding losses. Neoclassical utility functions are then reclassified as *normative* explanations of choice, while the BE function is classified as a *descriptive* representation of actual choices (Davis 2011; Heukelom 2011). As outlined below, this implies that properly-designed choice architectures (Thaler and Sunstein 2008), a policy-driven construction of the choice-making environment, should facilitate decision-making *as if* individuals had complete information about the consequences of their choices and could compensate for biases and heuristics.

From Kahneman and Tversky (1974) onward, BE has thereby proceeded from presuppositions of hyper-rationality and analysed deviations from this ‘rational’ yardstick as ‘irrational’ (Berg and Gigerenzer 2010). Thus, as Kahneman (2003a: 1449) suggests, ‘[t]he rational-agent model was our starting point and the main source of our null hypotheses.’ In making this case, bounded-rationality is articulated as *departing* from hyper-rationality, thereby reinforcing the latter as the essential reference-point

in decision-making theories. While BE acknowledges that *Homo Economicus* constitutes an empirically unrealistic representation of subjects, it remains the *norm* against which reality is compared. For instance, although Kahneman and Tversky (1979) contend that neoclassical theories of constrained optimisation fail to realistically describe human behaviour and propose their own prospect theory as an alternative descriptive account, they do not abandon the former as the normatively favourable account of rationality. Consistent with the reasoning of Rabin (2013) above, they posit that ‘departures from expected utility theory must lead to normatively unacceptable consequences, such as inconsistencies, intransitivities, and violations of dominance’ (Kahneman and Tversky 1979: 277).

The categories utilised to classify and distinguish boundedly-rational individuals from their ‘rational’ counterpart subsequently pathologise the former. While characterising these subjects as ‘deviants’ requiring rehabilitation and boundedly-rational behaviour as ‘anomalous’, *Homo Economicus* is presented as a complete, healthy individual simultaneously advancing their own interests and that of society through behaviour enabling functioning markets. For instance, Thaler (1987: 198) elucidates an ‘anomaly’ as an empirical ‘result inconsistent with the present economics paradigm’, one ‘difficult to “rationalize” therein, or requiring ‘implausible assumptions [...] to explain it within the [dominant] paradigm’. Analogously, Medin and Bazerman (1999: 543), frame BE as concerned with delineating the ‘systematic ways in which people deviate from optimality or rationality’, while Diamond (2008: 1859) contends that BE aims to identify ‘circumstances where people are making “mistakes” arising from their deviations from hyper-rationality. Koszegi and Rabin (2008) and Beshears *et al.* (2008) then formulate techniques to distinguish such mistakes – understood as revealed preferences – as deviations from ‘normative’ preferences conforming to neoclassical presuppositions. Critically, this pathologisation does *not* stem from empirical investigation correlating deviations to negative outcomes, with psychological enquiry restricted to gauging whether behaviour conforms to hyper-rationality (Infanti *et al.* 2016).⁷ Such use

⁷ Bruni and Sugden (2007: 148) identify the analogous methodological defence rationalising the shared normative elucidations informing neoclassicism and BE:

The essential idea behind the discovered preference hypothesis is that rational-choice theory is descriptive of the behaviour of economic agents who, through

of the terms ‘anomaly’ and ‘deviation’ here articulate a hierarchical correlation between the categories of (hyper-)rationality and bounded-rationality, whereby the latter accedes to and is refracted through the former (Mehta 2013). Through positioning these representations of decision-making in *binary opposition*, each procures meaning when juxtaposed, such that bounded-rationality is articulated as contrary to that deemed normal or desirable.

This hierarchical ordering is especially evident in the growing literature of behavioural development economics (‘BDE’). While claiming that such research should not be conflated with ‘pejorative attempts to label the poor as “irrational” [...] blame the poor for their poverty [or] argue that the poor have specific irrationalities’ (Mullainathan 2005: 47), it proceeds to consider how impoverished individuals’ failure to conform to the hyper-rational presuppositions of *Homo Economicus* engenders and entrenches poverty in the Global South by limiting engagement in more ‘rational’, welfare-enhancing economic activities (e.g. Banerjee and Duflo 2011; Datta and Mullainathan 2014). This is routinely conceptualised through identification of ‘anomalies’ restricting potentially profitable, albeit risky, entrepreneurial endeavours (Berndt 2015). For instance, the principle of *loss aversion* is held to explain individuals’ apprehension about the prospect of loss rather than risk *per se* (Mullainathan and Thaler 2001; Gächter *et al.* 2009). This leads to vulnerable farmers being reluctant to modify their behaviour and thereby rationalises their unwillingness ‘to put assets at risk by buying agricultural inputs they are not guaranteed to recoup’ (Fafchamps 2009: 16) and entrenching extant practices (Fowler and Brand 2011). Thus, Banerjee (2004) and Liu and Huang (2013) interpret the failure of impoverished peasants to adopt the ‘optimal’ (largely commercial) technology with which to cultivate their land as arising from biases against loss aversion, such as exhausting a considerable portion of their savings or the potential health effects of utilising pesticides. In these cases, the impoverished are held to be sufficiently averse to loss that they inadvertently propagate the poverty-cycle by declining otherwise-judicious investment opportunities.

experience and deliberation, have learned to act in accordance with their underlying preferences; deviations from that theory are interpreted as short-lived errors.

Coterminous with the TH centring of the subject as *logos*, such discourse articulates a social ontology according to the accepted behavioural norms from which boundedly-rational individuals are deemed to have diverged. That is, the notions of ‘deviation’ and ‘anomaly’ encode the BE concern with investigating how an aggregation of hyper-rational actors enables markets to operate effectively. As outlined above, neoclassicism holds the essence of properly-functioning markets to be decision-making by the hyper-rational individual subject, while all subjects will derive collective benefits from the aggregation of such behaviour. Conversely, individuals exhibiting limited capacity to operate in this manner impose social costs by undermining functioning markets. Importantly, however, boundedly-rational behaviour is not *inherently* problematic to others: it only appears so in the context of neoclassical presuppositions concerning market operations and social welfare as depending on the reconciliation of individual and aggregate rationality in these markets (Mehta 2013). BE discourse thus legitimises interventions to ‘correct’ boundedly-rational behaviour, as examined below.

To substantiate the implications of this pathologisation of bounded-rationality, it is apposite to consider the behavioural assumption that real-world decision-making may be conceptualised through the ‘dual process’ heuristic, outlined above. Recall that this depicts the human brain as comprised of two systems of thought operating simultaneously: the emotional and impulsive *automatic system* and the more deliberate, rational *reflective system*. Reliance on the former augments decision-makers’ vulnerability, as the ensuing biases and heuristics lead to faulty perceptions about the effects of her/his actions, preferences detrimental in the long-term, or individuals choosing damaging behaviours even when preferring otherwise. In challenging the universalist presumptions of hyper-rationality, this articulation of a more complex account of decision-making appears to complement long-standing calls for such an approach within political economy (*e.g.* Simon 1959). In particular, through embracing a form of economic analysis previously denigrated as ‘feminine’ and ‘soft’ (Nelson 1995: 134; Hewitson 1999), BE seemingly answers the case made by critical feminist political economists for a theory of cognition and behaviour transcending the modernist reason-emotion dualism embodied in ‘Descartes’ Error’ (Damasio 1994). As illustrated in Table 2, rather than privileging characteristics of ‘rationality’ (left column) in the reason-emotion dualism, BE recognises

behavioural diversity and explicitly models emotion, intuition and social cognition (right column).

Table 2: Dichotomy of the ‘rational’ and ‘emotional’

Rational	Emotional
Deliberative	Intuitive
Objective	Subjective
Thoughts	Feelings
Conscious	Unconscious
Mind	Body
Cerebral	Visceral
Dynamic	Predetermined
Reason	Deception

Source: Adapted from Wilkins (2013: 7).

Yet, as depicted in Figure 1, rather than examining emotion and intuition primarily to *understand* human behaviour, they are articulated as hindering hyper-rational decision-making and preventing markets from functioning effectively. BE thereby continues to implicitly separate the rational and emotional in explanations of decision-making. This depends on neurobiological interpretations of the latter as cognition: determining immanent thought processes identified within a particular (anterior insula) region of the brain (Gladwell 2005), rather than a differentiated, socially contextualised feeling (Pykett 2012; Wilkins 2013). By subsuming emotion into cognition, it remains grounded in ‘a system of binary opposites that privileges the masculine over the feminine’ and ‘co-opts emotional experience’ (Mumby and Putnam 1992: 469) into a limited rationalist frame (*cf.* Thaler and Sunstein 2008: 21). Put differently, BE perpetuates a binary construction subordinating the automatic, emotional, ‘feminised’ brain to the norm of its deliberative, rational, ‘masculinised’ counterpart: relegating feelings and ambiguity to the physical and feminine sphere of irrational, corporeal reaction and demarcating it from non-rational judgement and instinct (Pykett 2012).

Individuals are thus held as predictably erring due to their brain chemistry or inability to manage it and, consequently are unable to consistently formulate utility-maximising choices.

This appropriation of psychology to pathologise bounded-rationality formulates the limitations of actors' cognitive capacities as inherent, biomedical in origin or indicative of mental abnormality (Mehta 2013). Discursive allusions to irregularity are exacerbated by evocations of mental disorder and bodily affliction, such as the *Report* (World Bank 2015) identifying 'willpower deficits' (p. 115), 'low bandwidth' (p. 87) and diminished capacity to 'control temptation and delay gratification' (p. 123) as hindering the capacity of the poor to escape impoverishment. Similarly, in examining the prevalence of HIV in specific sub-Saharan African countries, (de Walque *et al.* 2012: 6) assert that individuals engage in risky sexual behaviour despite the pervasiveness of HIV due to myopia: weighing the short-term benefits more highly than the long-term costs. By favouring short-term physical pleasure, young people 'appear to understand their HIV risks and know how to behave to prevent transmission — yet they don't choose to act on that knowledge.' Finally, DellaVigna and Malmendier (2004: 394) demarcate 'sophisticated' *Homines Economici*, for whom 'market interaction with [...] firms enables the individuals to achieve the efficient consumption level' from 'naïve' boundedly-rational subjects, for whom 'firms design the pricing so as to take maximal advantage of the consumer overconfidence and underestimation of renewal'. Such labels denigrate boundedly-rationality individuals as deficient relative to their (hyper-)rational counterparts. Credence is thus buttressed in hyper-rational presuppositions as prescribing how a 'whole' subject should be configured.

Implications for policy

This subjectification of *Homo Economicus* has tangible consequences for boundedly-rational actors. Since they are conceptualised as deviating from *Homo Economicus*, they are positioned as legitimate targets of corrective *policy*. This logic mirrors that of the 'medical' model of disability identified by Handley (2003: 110), in which medical knowledge classifies 'disability' as 'the reduction or absence of an individual's physical, cognitive or sensory functions to the point that

“normal” functioning and capabilities are restricted or absent and that such states are entirely natural phenomena’. Actors are constructed in binary opposition by refracting conceptions of disability through the prism of assumptions concerning what constitutes ‘normalcy’, ‘able’ and ‘disabled’. Correspondingly, articulating ‘(hyper-)rational’ and ‘boundedly-rational’ actors as antinomies in BE presents the latter – like disabled individuals – as *deficient* relative to the ideal (Mehta 2013). Comprehending disability as actors’ functional limitations, policy prescriptions underscore convalescence and cure to reintegrate disabled individuals into social processes, rather than positing social reform accommodating them or questioning structural and cultural factors relating to disability. Analogously, by articulating bounded-rationality in terms of actors’ deficient cognitive systems, BE policy prescriptions aim to engender behaviour approximating hyper-rationality, rather than interrogating the values and concerns leading BE to pathologise bounded-rationality (McMahon 2015).

In addition to constituting the normative *ideal* for subjects, *Homo Economicus* is thus also subjectified as a state *potentially-realizable* through policy. BE policy initiatives are oriented around an imperfect and irrational subject requiring correction – whether overtly, such as through social marketing (Crawshaw 2013); implicitly, as in providing ‘cooling-off periods’ when purchasing insurance (Dolan *et al.* 2010); or circumventing the inexorable irrationality of individuals, such as through default pension enrolments (Benartzi and Thaler 2007). As Pykett (2012: 222) observes, ‘[p]eople who are not expert in managing their emotions, by implication, need the government to manage their emotions for them – through affective arrangements, support for mental short-cuts and education and training for the more reflexive aspects of the brain.’

Grounded in the logic of what Hausman (2012: 102) labels ‘preference purification’, BE seeks to investigate and reconstruct the preferences that *would* have informed the decision-making of hyper-rational individuals had her/his cognition not been obstructed by psychological factors, while establishing realisation of such reconstructed preferences as a normative benchmark for policy-making (Infante *et al.* 2016). That is, based on the dualistic model of individuals in Figure 1 above, BE attempts to recreate the preferences of the hyper-rational agent in the inner-core through isolating this norm from the distorting influences of the psychological outer-shell. In turn, it recommends policies enabling individuals to make decisions in accordance with these preferences *as if* they were *Homo*

Economicus (Thaler and Sunstein 2003).⁸ For instance, to enable overweight individuals to counter their willpower problems and consume healthier diets, O'Donoghue and Rabin (2006) recommend taxing potato chips and subsidising carrots. This formulation assumes hyper-rationality to be hindered by irrational biases favouring immediate (as opposed to long-term, rational) gratification (Loewenstein 1996). Striving for ideal subjectivity, therefore, engenders policies enabling suspension or moderation of emotion to complement the performance of its inferred opposite: hyper-rational decision-making based on constrained optimisation in which behaviour is finely attuned to pecuniary incentives. Consequently, individual bounded-rationality, presented as explaining why markets fail to operate as predicted by neoclassicism, engenders a functionalist rationalisation for corrective non-market institutional arrangements. By subjectifying *Homo Economicus* as the ideal, BE assumes that constructing institutional environments enabling individuals to formulate more 'rational' decisions – through designing default options curtailing inertia or conceiving information devices bolstering her/his capacity to select welfare-enhancing options – will resolve the theoretical problematic of mediating between individual and collective rationality. Whether grounded in discourse of libertarian-paternalism or 'nudging' (Thaler and Sunstein 2003, 2008; Sunstein 2015), debiasing through law (Jolls *et al.* 1998; Jolls and Sunstein 2006), or asymmetric paternalism (Camerer *et al.* 2003), BE policy seeks to implement modest interventions facilitating subjectivity approximating *Homo Economicus* and, thus, accord the microfoundation for markets to function effectively.⁹ By not restricting choice, such interventions claim to benefit

⁸ For a comprehensive critical overview of this logic of 'preference purification', see Infante *et al.* (2016).

⁹ Behavioural experts have increasingly been invited by national and supranational public agencies to offer alternative policy analyses and recommendations. The former Obama administration in the United States, for example, implemented myriad BE-inspired reforms, such as the Making Work Pay tax cut in 2009, the individual mandate informing the 2010 health-care reform bill, various components of the Dodd-Frank financial reform bill, and assorted other regulations and policy measures. Embodied in the Behavioural Insights Team, the Conservative government in the United Kingdom similarly accorded increasing influence to BE, while the influence of behavioural techniques has also been noted in countries such as New Zealand, Brazil and France (for surveys, see Jones *et al.* 2013). Indeed, in the context of Australia, a Behavioural Economics Team (BETA) was also established in the Department of Prime Minister and Cabinet in 2015 (see <https://www.dpmc.gov.au/domestic-policy/behavioural-economics>).

boundedly-rational actors while imposing minimal harm on ‘rational’ actors (Sunstein 2015).

In this respect, by seeking to correct the marginal problems posed by individuals’ deviations from *Homo Economicus*, the objective of BE remains to improve the functioning or expansion of markets. That is, through circumspectly arranging each individuals’ choice architecture to enable her/him to act in accordance with her/his ‘true’ preferences (Hausman 2012), behavioural programs are deemed a means to secure the competitiveness and, indeed, legitimacy of markets (Santos and Rodrigues 2014). The result is that political questions concerning individual decision-making are confined to empirical and technical problems concerning economic actors and the market (McMahon 2015). In their discussion of mortgage markets, for instance, Thaler and Sunstein (2008) disregard the contribution of the global financial system in augmenting rates of household indebtedness, while restricting their focus to the complexity of individual financial decisions as the primary problem to be redressed through the appropriate choice architecture. In turn, choice architects may focus on formulating and implementing default options to shield actors from inertia or design information devices to bolster her/his ability to select welfare-enhancing options. In both cases, in accordance with the dualistic model of behaviour depicted in Figure 1, policy should seek to enable individuals to act *as if* s/he were *Homo Economicus* according to her/his underlying hyper-rational preferences, whilst simultaneously avoiding alteration of the extant structure of incentives defining a given problem-situation. That is, ‘[t]he central goal would be to inform consumers of fees rather than set prices’ (Thaler and Sunstein 2008: 83).

Conclusion

This looks like two steps forward and two back. BE seems to be an advance on neoclassical theory in two respects: i) it makes mainstream economics more ‘realistic’ by taking account of people’s decision-making behaviour in the real-world; and ii) it draws on literature from psychology, thereby giving the appearance of a more interdisciplinary method. However, as this article has demonstrated: iii) it retains the neoclassical focus on the individual and rational behaviour, but now treating it as a normative ideal rather than a description; and iv) it is used

as the basis for developing public policies that seek to ‘nudge’ people into behaving more like the hyper-rational ideal underpinning neoclassical theory. Arguably, this last feature makes BE more actively pernicious because it has become an active agent in trying to make real-world behaviour more like the *Homo Economicus* found in neoclassical economics textbooks.

David Primrose is a doctoral candidate in the Department of Political Economy at the University of Sydney. He would like to thank Frank Stiwell and the anonymous referees for their helpful advice on previous versions of this paper.

david.primrose@sydney.edu.au

References

- Ackerman, F. (2004) Still Dead After All These Years: Interpreting the Failure of General Equilibrium Theory’, in F. Ackerman and A. Nadal (eds), *The Flawed Foundations of General Equilibrium*, Routledge, London and New York, pp. 14-32.
- Akerlof, G.A. and R.J. Shiller (2009) *Animal Spirits: How Human Psychology Drives the Economy and Why It Matters for Global Capitalism*, Princeton University Press, Princeton.
- Akerlof, G.A. and R.J. Shiller (2015) *Phishing for Phools: The Economics of Manipulation and Deception*, Princeton University Press, Princeton.
- Altman, M. (2005) Reconciling Altruistic, Moralistic and Ethical Behaviour with the Rational Economic Agent and Competitive Markets, *Journal of Economic Psychology*, Vol. 26, pp. 732-57.
- Altman, M. (2015) ‘Introduction’, in M. Altman (ed.), *Real World Decision Making: An Encyclopedia of Behavioral Economics*, Praeger, ABC-CLIO, New York, pp. xv-xxxi.
- Ariely, D. (2008) *Predictably Irrational*, HarperCollins, New York.
- Arrow, K.J. (1987) Rationality of Self and Others in an Economic System, in R. Hogarth and M. Reder (eds), *Rational Choice*, Chicago University Press, Chicago, pp. 201-16.
- Aydinonat, N.E. (2009) *The Invisible Hand in Economics: How Economists Explain Unintended Social Consequences*, Routledge, London and New York.
- Backhouse, R.E. (1998) *Explorations in Economic Methodology: From Lakatos to Empirical Philosophy*, Routledge, London and New York.
- Banerjee, A.V. (2007) *Making Aid Work*, MIT Press, Cambridge and London.
- Banerjee, A.V. and E. Duflo (2011) *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs, New York.
- Bell, D. and I. Kristol (eds) (1981) *The Crisis in Economic Theory*, Basic Books, New York.

- Benartzi, S. and R. Thaler (2007) Heuristics and Biases in Retirement Savings Behavior, *The Journal of Economic Perspectives*, Vol. 21, No. 3, pp. 81-104.
- Berg, N. and G. Gigerenzer (2010) As-If Behavioral Economics: Neoclassical Economics in Disguise?, *History of Economic Ideas*, Vol. XVIII, No. 1, pp. 133-66.
- Berndt, C. (2015) Behavioural Economics, Experimentalism and the Marketization of Development, *Economy and Society*, Vol. 44, No. 4, pp. 567-91.
- Beshears, J., J.J. Choi, D. Laibson and B.C. Madrian (2008) How Are Preferences Revealed?, *Journal of Public Economics*, Vol. 92, No. 8, pp. 1787-94.
- Blaug, M. (1992) *The Methodology of Economics: Or How Economists Explain*, Cambridge University Press, Cambridge.
- Bowles, S. and H. Gintis (1990) Contested Exchange: New Microfoundations for the Political Economy of Capitalism, *Politics and Society*, Vol. 18, No. 2, pp. 165-222.
- Bowles, S. and H. Gintis (2000) Walrasian Economics in Retrospect, *The Quarterly Journal of Economics*, Vol.115, No. 4, pp. 1411-39.
- Bruni, L. and R. Sugden (2007) The Road Not Taken: How Psychology Was Removed from Economics, and How it Might be Brought Back, *Economic Journal*, Vol. 117, pp. 146-73.
- Butler, J. (1995) Conscience Doth Make Subjects of Us All, *Yale French Studies*, Vol. 88, pp. 6-26.
- Caldwell, B.J. (1991) The Methodology of Scientific Research Programmes: Criticisms and Conjectures in G.K. Shaw (ed.), *Economics, Culture, and Education: Essays in Honor of Mark Blaug*, Elgar, Aldershot, pp. 95-107.
- Callari, A.G. (2011) Robinson Crusoe and the Subject of Economics, in U. Grapard and G. Hewitson (eds), *Robinson Crusoe's Economic Man: A Deconstruction and Reconstruction*, Routledge, London and New York, pp. 75-90.
- Camerer, C.F. (1999) Behavioral Economics: Reunifying Psychology and Economics, *Proceedings of the National Academy of Sciences*, Vol. 96, No. 19, pp. 10575-7.
- Camerer, C.F. and E. Fehr (2006) When Does "Economic Man" Dominate Social Behavior?, *Science*, Vol. 311, No. 5757, pp. 47-52.
- Camerer, C.F., S. Issacharoff, G. Loewenstein, T. O'Donoghue and M. Rabin (2003) Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism', *University of Pennsylvania Law Review*, Vol. 151, No. 1211, pp. 1211-54.
- Camerer, C.F. and G. Loewenstein (2004) Behavioral Economics: Past, Present, Future, in C. Camerer, G. Loewenstein and M. Rabin (eds.), *Advances in Behavioral Economics*, Princeton University Press, New Jersey, pp. 3-52.
- Camerer, C. and D. Lovallo (1999) Overconfidence and Excess Entry: An Experimental Approach, *The American Economic Review*, Vol. 89, No. 1, pp. 306-18.
- Chernomas, R. and I. Hudson (2016) *Economics in the Twenty-First Century: A Critical Perspective*, University of Toronto Press, Toronto.
- Coase, R.H. (1937) The Nature of the Firm, *Economica*, Vol. 4, pp. 386-405.

- Coase, R.H. (1960) The Problem of Social Cost, *Journal of Law and Economics*, Vol. 3, No. 4, pp. 1-44.
- Colander, D. (2000) The Death of Neoclassical Economics, *Journal of the History of Economic Thought*, Vol. 22, No. 2, pp. 127-43.
- Colander, D. (2006) Post Walrasian Economics: Some Historical Links, in D. Colander (ed.), *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge University Press, Cambridge, pp. 46-69.
- Colander, D., R.P.F. Holt and J.B. Rosser (2004) The Changing Face of Mainstream Economics, *Review of Political Economy*, Vol. 16, No. 4, pp. 485-99.
- Conly, S. (2012) *Against Autonomy: Justifying Coercive Paternalism*, Cambridge University Press, Cambridge University Press, Cambridge.
- Copson, A. (2015) What Is Humanism?, in A. Copson and A.C. Grayling (eds), *The Wiley Blackwell Handbook of Humanism*, Wiley Blackwell, Chichester, pp. 1-33.
- Crawshaw, P. (2013) Public Health Policy and the Behavioural Turn: The Case of Social Marketing, *Critical Social Policy*, Vol. 33, No. 4, pp. 616-37.
- Crespo, R.F. (2017) *Economics and Other Disciplines: Assessing New Economic Currents*, Routledge, London and New York.
- Damásio, A.R. (1994) *Descartes' Error: Emotion, Reason and the Human Brain*, Picador, London.
- Datta, S. and S. Mullainathan (2014) Behavioral Design: A New Approach to Development Policy, *Review of Income and Wealth*, Vol. 60, No. 1, pp. 7-35.
- Davis, J.B. (1998) Normative and Positive Economics, in P.A. O'Hara (ed.), *Encyclopedia of Political Economy*, Taylor and Francis (Routledge), London and New York, pp. 804-7
- Davis, J.B. (2006) The Turn in Economics: Neoclassical Dominance to Mainstream Pluralism?, *Journal of Institutional Economics*, Vol. 2, No. 1, pp. 1-20.
- Davis, J.B. (2011) *Individuals and Identity in Economics*, Cambridge University Press, Cambridge.
- Davis, J.B. (2015) Bounded Rationality and Bounded Individuality, *Research in the History of Economics and Methodology*, Vol. 33, pp. 75-93.
- Della Vigna, S. (2009) Psychology and Economics: Evidence from the Field, *Journal of Economic Literature*, Vol. 47, No. 2, pp. 315-72.
- Della Vigna, S. and U. Malmendier (2004) Contract Design and Self-Control: Theory and Evidence, *Quarterly Journal of Economics*, Vol. 119, No. 2, pp. 353-402.
- De Walque, D., W.H. Dow, C. Medlin and R. Nathan (2012) Stimulating Demand for AIDS Prevention: Lessons from the RESPECT Trail, *National Bureau of Economic Research (NBER) Working Paper 17865*, NBER, Cambridge, available online: <http://www.nber.org/papers/w17865> (accessed: 7 January, 2016).
- Diamond, P.A. (2008) Behavioural Economics, *Journal of Public Economics*, Vol. 92, pp. 1858-62.
- Dow, S. (2013) Mark Blaug on Formalism and Reality: The Case of Behavioural Economics, *Erasmus Journal for Philosophy and Economics*, Vol. 6, No. 3, pp. 26-43.

- Fafchamps, M. (2009) 'Vulnerability, Risk Management and Agricultural Development', paper prepared for the *AERC Conference on Agriculture and Development*, Mombasa, available online: <http://www.web.stanford.edu/~fafchamp/vulneb.pdf> (accessed: 21 December 2016).
- Fehr, E., U. Fischbacher and S. Gächter (2002) 'Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms', *Human Nature*, Vol. 13, No. 1, pp. 1-25.
- Fehr, E. and K. Schmidt (1999) 'A Theory of Fairness, Competition and Cooperation', *Quarterly Journal of Economics*, Vol. 114, No. 3, pp. 817-68.
- Fine, B. and D. Milonakis (2009) *From Economics Imperialism to Freakeconomics: The Shifting Boundaries between Economics and Other Social Science*, Routledge, London and New York.
- Foucault, M. (1972) *The Order of Things: An Archaeology of the Human Sciences*, Vintage/Random House, New York.
- Foucault, M. (1980) *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*, Harvester Wheatsheaf, London.
- Foucault, M. (1982) 'The Subject and Power', *Critical Inquiry*, Vol. 8, No. 4, pp. 777-95.
- Foucault, M. (1991) *Discipline and Punish: The Birth of the Prison*, translated by Alan Sheridan, Penguin Books, New York.
- Foucault, M. (2003) *Society Must Be Defended: Lectures at the Collège de France 1975-1976*, Allen Lane, London.
- Fowler, B. and M. Brand (2011) 'Pathways Out of Poverty: Applying Key Principles of the Value Chain Approach to Reach the Very Poor', *USAID Discussion Paper Microreport No. 173*, available online: <http://microlinks.kdid.org/library/pathways-out-poverty-applying-key-principles-value-chain-approach-reach-very-poor> (accessed 19 November 2016).
- Friedman, M. (1953) 'The Methodology of Positive Economics', in *Essays in Positive Economics*, University of Chicago Press, Chicago.
- Fullbrook, E. (2001) 'Real Science is Pluralist', *Post-Autistic Economics Newsletter*, Vol. 5 March, Article 5.
- Gächter, S., H. Orzen, E. Renner and C. Starmer (2009) 'Are Experimental Economists Prone to Framing Effects? A Natural Field Experiment', *Journal of Economic Behavior and Organization*, Vol. 70, pp. 443-46.
- Garnett Jr., R.F. (1994) *Value, Man and Markets in Modern Economic Discourse*, unpublished doctoral dissertation, University of Massachusetts, Michigan.
- Garnett Jr, R.F. (2006) 'Paradigms and Pluralism in Heterodox Economics', *Review of Political Economy*, Vol. 18, No. 4, 521-46.
- Gigerenzer, G. (2000) *Adaptive Thinking: Rationality in the Real World*, Oxford University Press, New York.
- Gigerenzer, G. (2015) 'On the Supposed Evidence for Libertarian Paternalism', *Review of Philosophy and Psychology*, Vol. 6, No. 3, pp. 361-83.
- Gilovich, T., D. Griffin and D. Kahneman (eds) (2002) *Heuristics and Biases: The Psychology of Intuitive Judgement*, Cambridge University Press, Cambridge.

- Gladwell, M. (2005) *Blink: The Power of Thinking Without Thinking*, Penguin Books, London.
- Goldstein, D. and D. Gigerenzer (2002) Models of Ecological Rationality: The Recognition Heuristic', *Psychological Review*, Vol. 109, pp. 75-90.
- Griffin, D. and A. Tversky (1992) The Weighing of Evidence and the Determinants of Confidence, *Cognitive Psychology*, Vol. 24, No. 3, pp. 411-35.
- Hahn, F. (1984) *Equilibrium and Macroeconomics*, MIT Press, Cambridge.
- Handley, P. (2003) Theorising Disability: Beyond 'Common Sense', *Politics*, Vol. 23, No. 3, pp. 109-18.
- Hausman, D.M. (2012) *Preference, Value, Choice, and Welfare*, Cambridge University Press, Cambridge.
- Heidl, S. (2016) *Philosophical Problems of Behavioural Economics*, Routledge, London.
- Heilmann, C. (2014) 'Success Conditions for Nudges: A Methodological Critique of Libertarian Paternalism', *European Journal for Philosophy of Science*, Vol. 4, No. 1, pp. 75-94.
- Hennings, K. and W.J. Samuels (eds) (1990) *Neoclassical Economic Theory: 1870 to 1930*, Kluwer Academic Publishers, Boston.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, N.S. Henrich, K. Hill, F. Gil-White, M. Gurven, F.W. Marlowe, J.Q. Patton and D. Tracer (2005) 'Economic Man' in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies, *Behavioral and Brain Sciences*, Vol. 28, No. 6, pp. 795-815.
- Heukelom, F. (2011) Behavioral Economics, in J.B. Davis and D. Wade-Hands (eds), *The Elgar Companion to Recent Economic Methodology*, Edward Elgar, Cheltenham, pp. 19-38.
- Heukelom, F. (2012) A Sense of Mission: The Alfred P. Sloan and Russell Sage Foundations' Behavioural Economics Program, 1984-1992, *Science in Context*, Vol. 25, No. 2, pp. 263-86.
- Heukelom, F. (2014) *Behavioural Economics: A History*, Cambridge University Press, Cambridge.
- Hewitson, G. (1999) *Feminist Economics: Interrogating the Masculinity of Rational Economic Man*, Edward Elgar, Cheltenham.
- Hollis, M. and E.J. Nell (1975) *Rational Economic Man: A Philosophical Critique of Neo-Classical Economics*, Cambridge University Press, London.
- Infante, G., G. Lecouteux and R. Sugden (2016) Preference Purification and the Inner Rational Agent: A Critique of the Conventional Wisdom of Behavioural Welfare Economics, *Journal of Economic Methodology*, Vol. 23, No. 1, pp. 1-25.
- Ingrao, B. and G. Israel (1990) *The Invisible Hand: Economic Theory in the History of Science*, MIT Press, Cambridge.
- Jolls, C. and C.R. Sunstein (2006) Debiasing Through Law, *Journal of Legal Studies*, Vol. 35, pp. 199-241.

- Jolls, C., C.R. Sunstein and R. Thaler (1998) A Behavioral Approach to Law and Economics, *Stanford Law Review*, Vol. 50, No. 5, pp. 1593-1608.
- Jones, R., J. Pykett and M. Whitehead (2013) *Changing Behaviours: On the Rise of the Psychological State*, Edward Elgar, Cheltenham.
- Kahneman, D. (2003a) Maps of Bounded Rationality: Psychology for Behavioral Economics (Nobel Speech), *American Economic Review*, Vol. 93, No. 5, pp. 1449-75.
- Kahneman, D. (2003b) A Psychological Perspective on Economics, *American Economic Review*, Vol. 93, No. 2, pp. 162-8.
- Kahneman, D. (2011) *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York.
- Kahneman, D. and A. Tversky (1979) Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, Vol. 47, No. 2, pp. 263-91.
- Katzner, D. (1999) Methodological Individualism and the Walrasian Tâtonnement, *Journal of Economic and Social Research*, Vol. 1, pp. 5-33.
- Katzner, D. (2010) The Current Non-Status of General Equilibrium Theory, *Review of Economic Design*, Vol. 14, No. 1-2, pp. 203-219.
- Keynes, J.N. (1973) *The Scope and Method of Political Economy*, Transaction Publishers, New Brunswick and London.
- Koppl, R. (1992) Invisible-Hand Explanations and Neoclassical Economics: Toward a Post Marginalist Economics, *Journal of Institutional and Theoretical Economics*, Vol. 148, No. 2, pp. 292-313.
- Kőszegi, B. and M. Rabin (2008) Choices, Situations, and Happiness, *Journal of Public Economics*, Vol. 92, pp. 1821-32.
- Laibson, D.I. (1997) Golden Eggs and Hyperbolic Discounting, *Quarterly Journal of Economics*, Vol. 112, No. 2, pp. 443-77.
- Laibson, D.I. (2002) Bounded Rationality in Economics, PowerPoint slides, University of California Berkeley 2002 Summer Institute on Behavioral Economics website, available online: <http://elsa.berkeley.edu/symposia/sage02/slides/laibson3.pdf> (accessed: 21 January 2016).
- Laibson, D.I. and J.A. List (2015) Principles of (Behavioral) Economics, *American Economic Review*, Vol. 105, No. 5, pp. 385-90.
- Leggett, W. (2014) The Politics of Behaviour Change: Nudge, Neoliberalism and the State, *Policy and Politics*, Vol. 42, No. 1, pp. 3-19.
- Liu, E.M. and J-K. Huang (2013) Risk Preferences and Pesticide Use by Cotton Farmers in China, *Journal of Development Economics*, Vol. 103, pp. 202-15.
- Loewenstein, G. (1996) Out of Control: Visceral Influences on Behavior, *Organizational Behavior and Human Decision Processes*, Vol. 65, No. 3, pp. 272-92.
- Madra, Y.M. (2016) *Late Neoclassical Economics: The Restoration of Theoretical Humanism in Contemporary Economic Theory*, Routledge, London and New York.
- McMahon, J. (2015) Behavioral Economics as Neoliberalism: Producing and Governing *Homo Economicus*, *Contemporary Political Theory*, Vol. 14, No. 2, pp. 137-58.

- Medin, D. and M. Bazerman (1999) Broadening Behavioral Decision Research: Multiple Levels of Cognitive Processing, *Psychonomic Bulletin and Review*, Vol. 6, No. 4, pp. 533–46.
- Mehta, J. (2013) The Discourse of Bounded Rationality in Academic and Policy Arenas: Pathologising the Errant Consumer, *Cambridge Journal of Economics*, Vol. 37, No. 6, pp. 1243–61.
- Milonakis, D. and B. Fine (2008) *From Political Economy to Economics: Method, the Social and the Historical in the Evolution of Economic Theory*, Routledge, London.
- Mirowski, P. and E. Nik-Khah (2017) *The Knowledge We Have Lost in Information: The History of Information in Modern Economics*, Oxford University Press, New York.
- Mullainathan, S. (2005) Development Economics Through the Lens of Psychology, in F. Bourguignon and B. Pleskovic (eds), *Annual World Bank Conference in Development Economics 2005: Lessons of Experience*, The World Bank/Oxford University Press, New York, pp. 45–70.
- Mullainathan, S. and R. Thaler (2001) Behavioural Economics, in N. Smelser and P. Baltes (eds), *International Encyclopedia of the Social and Behavioural Sciences*, Pergamon, Oxford, pp. 1094–100.
- Mumby, D.K. and L.L. Putnam (1992) The Politics of Emotion: A Feminist Reading of Bounded Rationality, *Academy of Management Review*, Vol. 17, No. 3, pp. 465–86.
- Nelson, J.A. (1995) Feminism and Economics, *Journal of Economic Perspectives*, Vol. 9, No. 2, pp. 131–48.
- O'Donoghue, T. and M. Rabin (1999) Doing It Now or Later, *American Economic Review*, Vol. 89, No. 1, pp. 103–24.
- O'Donoghue T. and M. Rabin (2006) Optimal Sin Taxes, *Journal of Public Economics*, Vol. 90, No. 10–11, pp. 1825–49.
- Pesendorfer, W. (2006) Behavioral Economics Comes of Age: A Review Essay on 'Advances in Behavioral Economics', *Journal of Economic Literature*, Vol. 44, No. 3, pp. 712–21.
- Pykett, J. (2012) The New Maternal State: The Gendered Politics of Governing Through Behaviour Change, *Antipode*, Vol. 44, No. 1, pp. 217–38.
- Rabil Jr., A. (ed.) (1988a) *Renaissance Humanism, Volume 1: Foundations, Forms, and Legacy*, University of Pennsylvania Press, Philadelphia.
- Rabil Jr., A. (ed.) (1988b) *Renaissance Humanism, Volume 2: Foundations, Forms, and Legacy*, University of Pennsylvania Press, Philadelphia.
- Rabin, M. (1998) Psychology and Economics, *Journal of Economic Literature*, Vol. 36, No. 1, pp. 11–46.
- Rabin, M. (2002) A Perspective On Psychology and Economics, *European Economic Review*, Vol. 46, No. 4, pp. 657–85.
- Rabin, M. (2013) Incorporating Limited Rationality into Economics, *Journal of Economic Literature*, Vol. 51, No. 2, pp. 528–43.
- Robbins, L. (1935) *An Essay on the Nature and Significance of Economic Science*, Macmillan, London.

- Roberto, C.A. and I. Kawachi (2016) An Introduction to Behavioral Economics and Public Health', in C.A. Roberto and I. Kawachi (eds), *Behavioral Economics and Public Health*, Oxford University Press, New York, pp. 1-26.
- Rodrik, D. (2015) *Economic Rules: The Rights and Wrongs of the Dismal Science*, W.W. Norton, New York.
- Roncaglia, A. (2005) *The Wealth of Ideas: A History of Economic Thought*, Cambridge University Press, Cambridge.
- Ruccio, D.F. and J. Amariglio (2003) *Postmodern Moments in Modern Economics*, Princeton University Press, New Jersey.
- Santos, A. and J. Rodrigues (2014) Neoliberalism in the Laboratory? Experimental Economics on Markets and their Limits, *New Political Economy*, Vol. 19, No. 4, pp. 507-33.
- Screpanti, E. (2000) The Postmodern Crisis in Economics and the Revolution Against Modernism', *Rethinking Marxism*, Vol. 12, No. 1, pp. 87-111.
- Selten, R. (1998) Features of Experimentally Observed Bounded Rationality, *European Economic Review*, Vol. 42, No. 3, 413-36.
- Sen, A. (1977) Rational Fools: A Critique of the Behavioural Foundations of Economic Theory, *Philosophy and Public Affairs*, Vol. 6, No. 4, pp. 317-44.
- Sen, A. (2002) *Rationality and Freedom*, The Belknap Press of Harvard University Press, Cambridge and London.
- Senstat, J. and G. Constantine (1975) A Critique of the Foundations of Utility Theory, *Science and Society*, Vol. 39, pp. 157-79.
- Sent, E-M. (1998) *The Evolving Rationality of Rational Expectations: An Assessment of Thomas Sargent's Achievements*, Cambridge University Press, Cambridge.
- Shaikh, A. (2016) *Capitalism: Competition, Conflict, Crises*, Oxford University Press, New York.
- Shapiro, C. and J.E. Stiglitz (1984) Equilibrium Unemployment as a Worker Discipline Device, *American Economic Review*, Vol. 74, pp. 433-44.
- Simon, H.A. (1955) A Behavioral Model of Rational Choice, *Quarterly Journal of Economics*, Vol. 69, No. 1, pp. 99-118.
- Simon, H.A. (1956) Rational Choice and the Structure of the Environment, *Psychological Review*, Vol. 63, pp. 129-38.
- Simon, H.A. (1959) Theories of Decision-Making in Economics and Behavioral Science, *American Economic Review*, Vol. 49, No. 3, pp. 253-83.
- Sloman, S.A. (1996) The Empirical Case for Two Systems of Reasoning, *Psychological Bulletin*, Vol. 119, No. 1, pp. 3-22.
- Sugden, R. (2015) Looking for a Psychology for the Inner Rational Agent, *Social Theory and Practice*, Vol. 41, No. 4, pp. 579-98.
- Sunstein, C.R. (2014) *Why Nudge? The Politics of Libertarian Paternalism*, Yale University Press, New Haven.
- Thaler, R. (1987) Anomalies: The January Effect, *Journal of Economic Perspectives*, Vol. 1, No. 1, pp. 197-201.

- Thaler, R. (2000) From Homo Economics to Homo Sapiens, *Journal of Economic Perspectives*, Vol. 14, No. 1, pp. 133-41.
- Thaler, R. (2015) *Misbehaving: The Making of Behavioural Economics*, WW Norton, New York.
- Thaler, R. and C.R. Sunstein (2003) Libertarian Paternalism, *American Economic Review*, Vol. 93, No. 2, pp. 175-9, Papers and Proceedings.
- Thaler, R. and C.R. Sunstein (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press, New Haven.
- Thrift, N.J. (2006) Re-Inventing Invention: New Tendencies in Capitalist Commodification, *Economy and Society*, Vol. 35, No. 2, pp. 279-306.
- Tversky, A. and D. Kahneman (1973) Availability: A Heuristic for Judging Frequency and Probability, *Cognitive Psychology*, Vol. 5, pp. 207-32.
- Tversky, A. and D. Kahneman (1974) Judgement Under Uncertainty: Heuristics and Biases, *Science*, Vol. 185, No. 4157, pp. 1124-31.
- Tversky, A. and D. Kahneman (1986) Rational Choice and the Framing of Decisions, *Journal of Business*, Vol. 59, No. 4, Part 2, pp. S251-S78.
- Tversky, A. and D. Kahneman (1991) Loss Aversion in Riskless Choice: A Reference-Dependent Model, *The Quarterly Journal of Economics*, Vol. 106, No. 4, pp. 1039-61.
- Veblen, T. (1898) Why Is Economics Not An Evolutionary Science?, *The Quarterly Journal of Economics*, Vol. 12, No. 4, pp. 373-97.
- Velupillai, V. and Y.F. Kao (2016) Herbert Alexander Simon, in G. Faccarello and H. Kurz, *Handbook of the History of Economic Analysis, Volume 1: Great Economists Since Petty and Boisguilbert*, Edward Elgar, Cheltenham, pp. 669-74..
- Von Neumann, J. and O. Morgenstern (2007) *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.
- Wilkins, A. (2013) Libertarian Paternalism: Policy and Everyday Translations of the Rational and the Emotional, *Critical Policy Studies*, Vol. 7, No. 4, pp. 395-406.
- Wolff, R.D. and S.A. Resnick (2012) *Contending Economic Theories: Neoclassical, Keynesian and Marxian*, MIT Press, Cambridge.
- World Bank (2015) *The World Development Report 2015: Mind, Society, and Behaviour*, World Bank, Washington.

