

Homo Economicus, AIs, humans and rats: decision-making and economic welfare

Diane Coyle¹

Abstract

Critics of economics often highlight two related issues: the empirical falsity of the 'homo economicus' assumption of rational, self-interested maximization; and the ethical consequences of models based on this assumption. Yet many experiments in biology show non-human creatures often seem to behave as if they were rational maximisers, suggesting that context rather than cognitive capacity is important for determining behaviour. The critique of rational choice poses a less serious methodological challenge to economics than is sometimes thought. However, economists do need to respond to the ethical critique that decisions and policies based on the assumption of rational self-interested maximisation change the norms of individual behaviour for the worse. This paper argues that economics has become divorced from ethics because for a century it has dealt only with ordinal, not cardinal, welfare rankings and has thus ruled out interpersonal comparisons. While enabling economists to separate normative from positive analysis, this separation protocol has left welfare economics both internally contradictory and unable to address major societal decisions, even though welfare economics is used constantly in limited ways, such as cost-benefit analysis. This separation reflects empirically inaccurate assumptions concerning preference formation and the conditions of supply and demand (but not the rational choice assumption) in the foundational welfare economic theorems. Economics must urgently revisit welfare economics, particularly in the context of modern economies in which individuals are increasingly interdependent, and the assumptions required for the fundamental welfare theorems therefore increasingly invalid.

JEL codes: A13, D60, B00

Rationality in the wild

In one of its experiments, the artificial intelligence company Deep Mind set its AI agents competing for scarce resources in a game involving picking apples (Leibo et al 2017). The game aimed to explore rational choice maximisation subject to constraints, modelled as sequential social dilemmas: when would the agents cooperate and when would they free ride?

¹ University of Manchester (Diane.Coyle@manchester.ac.uk).

My thanks to participants in a seminar at All Souls, Oxford in March 2017, including Tim Besley, Cecelia Heyes, John Kay, Jesse Norman and Michael Sandel; to participants in the September 2017 conference of the International Network for Economic Method; and also to Ronald Noë, Nuno Palma, and Peter Sinclair. The usual disclaimer applies.

These AIs used ‘deep reinforcement learning’ meaning the algorithmic agent, “Must learn to maximize its cumulative long-term reward through trial-and-error interactions with its environment.” They were designed to make decisions like *homo economicus*, rational actors in a classic economic model of constrained optimization, and each formed part of the environment to which they had to respond.

All was co-operative harmony when apples were plentiful. When they became scarcer, the AIs became more aggressive, ultimately attacking each other. The more intelligent the AI, the more aggressive it became as the competition got tougher. This result from artificial intelligences programmed to act like *homo economicus* would seem to confirm the prejudice of those critics of economics who regard this key behavioural assumption, rational self-interest, as unrealistic, distasteful, or both. Many critiques of mainstream economics have focused on the construct of *homo economicus*, arguing that humans more often demonstrate altruism and mutual concern. For example, Bowles (2004) argues that an empirically adequate behavioural account of economic decision-making must incorporate rules-of-thumb. The expanding behavioural economics and psychology literature lends support to the argument that *homo economicus* is an inappropriate assumption about how humans take economic decisions. Whatever one’s views about the morality of maximising behaviour, there is convincing evidence that in many contexts it is empirically incorrect. Alternative assumptions have come to be widely used in applied economics (Pesendorfer 2006).

However, another large category of experiments, in biology, has looked at the behaviour of many types of creature or other entities such as bacteria in varying conditions of resource scarcity. Sometimes these demonstrate strong emotional reactions, such as sharing with a friend, or punishing cheats even at a cost to themselves. But they also often seem to act like self-interested maximisers, for example being willing to write off sunk costs, appearing to calculate conditional probabilities correctly, or deriving consistent exchange rates of grapes for cucumbers, as would be predicted by economic models of constrained optimisation (see for example De Waal, 2006; Hammerstein and Noë, 2016; Herbranson, 2010; Hurley & Nudds, 2006). The kind of trade that takes place in these biological markets is not well-characterised by simplistic economic models. However, they do appear to be consistent with models of constrained optimization in the presence of information

asymmetries and transactions costs, or in other words models that are everyday fare in modern mainstream economics.

These results from the three sets of experimental results, in AI, biology, and in human psychology, raise the obvious question: given that all these types of entity can sometimes solve certain economic tasks as if rationally, what is the role of assumptions about cognitive capacities in describing the behaviour of these different entities – AIs whose cognitive capacity is beyond doubt, living entities including fungi and bacteria with no neurons and therefore no cognitive capacities, and humans? This is not a question about similarities or differences in the ‘deep plumbing’ of the minds of different kinds of creature (or algorithms); similar behaviours clearly cannot imply the same kinds of cognitive processes. The answer must be linked to the evolutionary process of specialisation and exchange by individual entities, subject to resource constraints, in a range of contexts. Cognitive mechanisms cannot match the results of natural selection over long periods in the case of recurring problems or choice contexts.

In addition, individual human choices are socially as well as environmentally contextualised. There is more social complexity among humans than among many other biological creatures, or AI agents. And as Leibo et al (2017) conclude: “The complexity of learning how to implement effective cooperation and defection policies may not be equal. One or the other might be significantly easier to learn.” Cooperation is highly complex, computationally; self-interested maximisation is easy. Scarce resources make co-operation more costly. Context is everything.

This has been acknowledged on occasion in the economics literature. One example is Becker (1962), demonstrating that market outcomes consistent with the constrained optimisation rational choice theories can come about even when individual choices are wholly ‘irrational’, either making random choices or being inert. With no assumed set of preferences or utility function, changes in the opportunity set deliver a classical downward sloping demand curve. “The market would act as if ‘it’ were rational not only when households were rational but also when they were inert, impulsive, or otherwise irrational.” The market outcome is the result of changes in the opportunity set and relative prices. The individual cognitive process is not relevant.

We have much to learn about human decision making in economic domain. This research is under way in labs and field experiments, in a thriving research programme. Economics is steadily incorporating 'behavioural' assumptions into its models, economists on the whole content or even eager to accept alternative motivations where these are empirically useful. There is immense interest in the field of behavioural economics, although this too may turn out to be a red herring if we take seriously the insight from both biological markets theory and information theory (Smith, 2017) that context rather than cognition is important in determining behaviour, and the challenge for economists (and others) is therefore identifying when decision-making contexts are comparable.

This enthusiasm for lessons from other fields is in line with a broader turn in the past two decades away from theory toward empirics, particularly among younger researchers, according to a review of the leading economic journals (Hamermesh, 2013). Several long-neglected or sidelined fields such as economic history and geography, and the study of institutions' role in economic outcomes, have seen significant revivals. Much of the interesting work in economics, attracting young researchers and graduate students, is on the borders with a range of neighbouring disciplines, not just psychology and cognitive science, but also history, geography, information theory, evolutionary biology, complexity, political economy. So not only is the critique of the *homo economicus* construct not as straightforward as an empirical matter as it seems, but it is also the case that economists are deeply interested in the empirical realities of economic decision-making. Few are perturbed by the assaults on *homo economicus*. On the contrary, many economists see this as an exciting research agenda.

The separation protocol: is and ought in economics

Yet the fact that the terrain of human decision-making is complicated, and that sometimes the rationality assumption will be valid and sometimes not, leaves the related frequent criticism of economics, which is that rational self-interest is – at least sometimes – unethical, and people *ought not* make their decisions that way. Furthermore, it is argued, the assumption of calculating self-interest encourages people to act in unethical ways as it gives them a justification for such behaviour, or a

social signal that it is acceptable: it is 'efficient', or it is just the way things are in the 'real world'. Bowles (2016) and Sandel (2012) offer examples of policies that are based on the assumption of self-interested maximisation inducing people to behave in more self-interested ways than they otherwise would. This line of criticism has drawn relatively little response from economists, and I argue here that the reason lies in the weak foundations of welfare economics – although they are weak for other reasons. Although welfare economics necessarily underpins policy evaluation, and is widely used in practical contexts (such as cost benefit analysis and competition assessments), economists have paid almost no attention to its theoretical justification or empirical validity since the 1960s.

Although for the critics of the subject, the *is* and the *ought* are inseparable, economics has insisted for more than 80 years – since Lionel Robbins' famous (1932) description of the subject – on a strict separation between positive and normative. Pigou (1908) represented the older tradition, writing: "Ethics and economics are mutually dependent." Indeed, Adam Smith made the same link. While realistic about human nature, he observed that everyone appreciates that, "[H]is own interest is connected with the prosperity of society." (Smith (2000, 1759); see also Rothschild (2001)). In a departure from this tradition, the positivist movement spearheaded in economics by Robbins and others led the discipline to rule out cardinal measurements of welfare and inter-personal welfare comparisons. Robbins (1932) claimed that economics and ethics were on 'different planes':

"Economics is neutral as between ends. Economics cannot pronounce on the validity of ultimate judgments of value."

Modern economics has assumed since then that utility is ordinal, not cardinal, and is therefore inherently not comparable between individuals. This separation protocol is still widely adhered to in the economics profession. For example, in a well-known essay, Milton Friedman (1953) wrote:

"Positive economics is in principle independent of any particular ethical position or normative judgments. ... Its task is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances. Its performance is to be

judged by the precision, scope, and conformity with experience of the predictions it yields. In short, positive economics is, or can be, an 'objective' science, in precisely the same sense as any of the physical sciences."

(See also Blaug, 1992). Many economists continue to see their discipline as largely contributing 'positive' insights, even as they also accept that many economic decisions or policy choices involve value judgements. There is certainly great merit in the implied effort to distinguish scientific knowledge from value-driven conclusions; without honourable efforts of this kind, there would be no room for policy debates seeking to find common ground between competing values and interests. Economists are able to derive solid policy-relevant empirical results, particularly in the realm of applied microeconomics. This includes evidence about 'behavioural' choices, where insights from economic research indicate policies that seem likely to improve individuals' economic welfare under any reasonable ethical framework (for example, Chetty, 2015). If everything is political, the floodgates are open to ignoring such evidence.

However, the separation protocol hamstring welfare economics, and therefore calls into question the validity of many practical applications of economics to policy choices. The protocol is manifested in the Pareto improvement criterion: a policy can only be said to increase aggregate social welfare if it makes at least one person better off, and no-one worse off. This is so obviously restrictive that economists have often argued (following Hicks (1939) and Kaldor (1939)) that a policy is welfare improving if the winners can (at least in theory) compensate the losers. This is sometimes called the potential Pareto improvement criterion. However, Scitovsky (1941) not long after, and subsequent authors (for example, Baumol, 1952; Roberts, 1980), have demonstrated that the Kaldor-Hicks compensation criterion is logically inconsistent unless unreasonable assumptions are made about individuals' preferences. For any intervention will change relative prices, and therefore the optimal quantities of goods for each individual, making it possible for both a given policy and its reversal to be Pareto improving if individual preferences for goods differ – a logical inconsistency. As Baumol put it, the Hicks-Kaldor criterion:

“[H]as not eliminated the problem of interpersonal comparison of utility. It has only subjected utility to the measuring rod of money, a measuring rod which bends and stretches and ultimately falls to pieces in our hands.” (p89)

The problem of assessing the economic welfare impact of policies gets thornier still when moving from the two-person two-good world to aggregate outcomes for society. Welfare economists had long noted that any aggregation involves an implicit value judgement about distribution (Graaff, 1957). In principle, the concept of a social welfare function (SWF) (Bergson, 1938; Samuelson, 1947) reintroduced normative judgements explicitly. The economist or policymaker can specify an objective function – say equal outcomes, or the maximin criterion – and aggregate individual utilities with appropriate weights. However, in his famous Theorem, Kenneth Arrow (1950) established that there is no aggregation procedure from individual utility to social welfare that will satisfy the Pareto criterion, and a small number of other seemingly reasonable assumptions. Although Sen (1970, 2017) has shown that aggregate social welfare can be defined when the assumptions (particularly unrestricted domain) are relaxed, Arrow’s theorem is a formal statement of the obvious truth that there are unavoidable conflicts of interest or dilemmas in society. The ‘impossibility’ is the result of a clash between trying to make a value judgement in terms of a SWF while excluding the possibility of interpersonal welfare comparisons.

Economics students are nevertheless socialised to glide from the basic microeconomics of individual utility maximisation to a somewhat vague sense of maximising social welfare, looking for Pareto improvements. Arrow is mentioned reverentially, perhaps even proven, only to then sit quietly on his pedestal. The second best theorem (Lipsey and Lancaster 1956) is mentioned, then quietly set to one side. Young economists are taught through the first and second welfare theorems that, under certain assumptions, the competitive market equilibrium is Pareto efficient; and that given any initial allocation of resources, a Pareto efficient outcome can be reached through exchange. The point that, given any initial allocation the domain of potential Pareto improvements will be small, is overlooked.

Resting on the scaffolding of utilitarianism and general equilibrium theory, these assumptions are:

- Consumers and producers are rational and self-interested
- Individuals have fixed preferences (over all current and future goods)
- Individuals have full information
- There is competition with no barriers to entry (or exit) (and therefore no increasing returns to scale)
- Private and social costs and benefits are equal (there are no externalities)
- Goods are rival – if I use it, you can't
- There are complete markets (including for all future goods)
- Goods are owned and able to be exchanged – there are property rights and the law enables and enforces transactions
- Individuals maximise their utility/profit
- Individual utilities can be aggregated

None of these assumptions is empirically valid (although, as noted earlier, the work on biological markets indicates the one most often criticised by others, the rational choice assumption, may be less troubling than the others at least in the context of collective outcomes). However, the most significant divergence between assumptions and reality, threaded through several on the list, is the separation between individual and social: the absence of externalities and increasing returns, for example, the assumption of fixed (rather than socially influenced) preferences, and the assumption of full information (rather than asymmetric information where the asymmetries may be determined by other people).

Nevertheless, the idea of competitive market equilibrium as the optimum outcome is a powerful benchmark. It was cemented into place by the co-evolution of events, political developments and economic ideas in the 1970s and 80s. Margaret Thatcher and Ronald Reagan implemented in their policies and philosophies a version of economics nurtured in think tanks and some universities during the whole of the post-war period (Stedman-Jones, 2012). In the evident macroeconomic failures of the 1970s, and the collapse of the centrally planned economies in 1989, the course of events seemed to validate this shift in public philosophy. Academic economics in turn embraced rational expectations, public choice and real business cycle theory, in a market turn that (I argue elsewhere) reached its high water mark around 1990 and has since been ebbing slowly (Coyle, 2007).

Whereas Pigou, the key early codifier of welfare economics, interpreted the presence of extensive market failures as a rationale for government intervention to achieve desired social outcomes, including for example redistribution to people on low incomes (Kumekawa, 2017), in the late 20th century the presumption was reversed: governments should only be intervening to fix specific, identified market failures. Indeed, 'government failure' was invoked in the public choice literature as at least an equally significant pitfall for the policymaker (Le Grand, 1991). Yet, as Baumol (1952, p165) noted, the conclusion that the market knows best derives entirely from the list of assumptions meaning there is no inter-dependence in individuals' preferences and utility outcomes. The reasoning is circular. The assumptions are not neutral.

Thanks to the separation protocol, though, economists' technocratic instinct remains strong. The task of the economist interested in public policy or social outcomes is technical: identify the relevant market failures, and appropriate correctives, leaving value judgements to others, the philosophers or politicians. Economics is concerned with normative questions, and economists have values and views, but the specific disciplinary contribution is technocratic. And this is surely right up to a point. Given a preferred outcome, inevitably based on ethical criteria, economics *does* provide theoretical and empirical tools to analyse how it might be attained. The subject has a tradition from Smith (1759) to Sen (2009) of explicitly adopting the perspective of the 'impartial spectator', taking into account other people's perspectives, in public reasoning. The majority of economists are therefore comfortable with their practice of attempting to separate value judgments, including their own, from an analysis of what *is* – and indeed most attempt to observe that separation honourably. This is easier in some fields, such as many areas of applied microeconomics, than in others, particularly macroeconomics, where the assumption that the economy can be characterised by a 'representative agent' has been both widely used and powerfully critiqued (Jackson and Yariv, 2017).

Implications for economic policy

The separation protocol severely hampers the economic analysis of public policy questions, where the underlying issue is precisely how to organise the collective use and allocation of resources. This is not to say that welfare economics is not used. On

the contrary, economists constantly make normative judgements. One instance is cost-benefit analysis (CBA), a widely implemented policy application. CBA tries to evade the issue of cardinal utility measurement by putting all costs and benefits into monetary terms, ideally using a competitive market price, although in practice often using a range of methods for deriving prices at which to calculate and compare costs and benefits. In arguing for the use of market prices, Harberger (1971) did observe that this ruled out certain dimensions people might value:

“These elements-which surely include the income-distributional and national-defense aspects of any project or program, and probably its natural-beauty aspects as well-may be exceedingly important, perhaps even the dominant factors governing any policy decision, but they are not a part of that package of expertise that distinguishes the professional economist from the rest of humanity.”

Here too, then, is an appeal to the separation protocol. It is not surprising that as the practice of CBA has developed, there have been increasing efforts to take into account ‘wider impacts’ including environmental externalities, for example. CBA has been criticised in particular from the perspective of environmentalists for trying to put a monetary figure on intrinsic value, on what many people consider to be inherently priceless (Kelman 1981). Economists have tried to address these challenges by extending CBA to a full general equilibrium context (Drèze and Stern, 1987) and by taking account of non-linear impacts such as environmental tipping points (Dietz and Hepburn, 2013). However, the technique simply makes implicit rather than explicit the normative judgements in any policy decision based on CBA by its use of the money metric as the yardstick for comparison. The question of whose benefits and costs is not addressed – this, again, is delegated by economists to the political process. Above all, the method assumes there is no difference between sum of the values of costs or benefits to an individual and the social costs and benefits. Interaction and social influence are assumed away. No wonder many decisions supposedly based on the technical exercise of a CBA have the air of having the result massaged to give the desired outcome.

Many other areas of applied economics also involve an explicit but limited welfare assessment. Examples include competition assessments, where the criteria refer to

the total or consumer surplus in a merger or market dominance context and its counterfactual, or any policy debate concerned about increasing productivity or living standards in the context of endogenous growth or agglomeration economies.

The challenge to economics from the economy

The circumstances in which markets fail are exactly the circumstances in which governments fail too, because they are when private and collective interests diverge the most. What's more, these divergences are growing thanks to the changing character of the economy. Technological innovation has increased the extent and significance of social spillovers, for example in the form of network effects in digital markets or agglomeration economies in economic geography.

Technology has always been social. Even an old technology such as electricity took half a century to manifest itself in higher productivity because of the need to make many complementary investments and to rearrange the organisation of work and home (David, 1990). Even now is difficult for many low- and middle-income countries to deliver a consistent electricity supply. For that matter, power blackouts are not uncommon even now in leading economies such as Denmark, or the United States, when the political and social conditions for generation, distribution and transmission go awry. Complex goods and services at the technological frontier involve large-scale co-operative activities, extensive communication and knowledge, and major investments in tangible and intangible infrastructure. New goods and services are often non-rival and therefore have public goods characteristics. Increasing returns to scale, externalities, and non-rival goods are pervasive in modern knowledge economies. When there is rapid technical change and rapid diffusion, as now, fixed preferences are even less likely to exist than in stable times when it is simply fashion or social influence or learning that change individual preferences. Some of the biggest digital businesses now are even called 'social networks'. Government co-ordination in terms of public-good research, technical standards, skills and so on is essential for markets to come into existence. The century old 'states' vs 'markets' dichotomy is not a valid empirical description of modern economies. Although there never was a market without the state (or vice versa), their mutual interdependence has increased greatly with the levels of economic complexity and living standards.

Similarly, consider the everyday use of GDP growth as a metric of the success of economic management. While economists will be aware of some well-known failings of GDP to measure economic welfare (such as environmental externalities, or the exclusion of consumer surplus from innovations), the presumption is that higher real GDP means higher social welfare (Coyle, 2014). Yet as any aggregation of individual outcomes involves a judgement (implicit or explicit) about distributional weights, an increase in GDP can only be evaluated at all in terms of economic welfare by taking into account distributional changes (Jorgenson, 2017). Though appreciated in theory since at least Samuelson (1947) or Graaff (1957), this had been long forgotten, until Piketty (2014) and his colleagues brought the generalised increase in inequality since 1980 to public and policy attention.

In this actual economic context, of a growing gap between the assumptions of individual preferences and action in the welfare theorems and the inter-dependence of individuals in actual economies, the welfare economic foundations for analysing public policy are increasingly inadequate. Atkinson (2001) noted that students had stopped studying welfare economics by the 1960s, even though economics was increasingly full of value statements: “[D]espite the prevalence of welfare statements in modern economics, we are no longer subjecting them to critical analysis.” This may have been overstating the case. Nevertheless, standard textbooks from the 1970s onward contain technical discussions of Pareto efficiency, the welfare theorems, and consumer and producer surplus; but lack anything resembling the thoughtful discussion in Samuelson (1947). Atkinson argued that economists must consider explicitly the moral consequences of particular models, as the scrutiny might lead them to re-examine their models. Recent changes in the economy make it all the more important to consider revive welfare economics. The economic analysis of public policy choices ultimately is not and cannot be technocratic. The fundamental reason is that individual and collective interests will often diverge. Conflicts of interest and hence interpersonal welfare comparisons are unavoidable, and the Pareto criterion is of relatively little practical use. The fundamental welfare theorems should perhaps be read as a lesson in how *unlikely* unco-ordinated individual exchange alone is to achieve Pareto efficiency.

How economics has responded

Thus the assumptions underpinning the standard welfare economic framework are decreasingly realistic. Yet in other fields of economics, the importance of interdependency rather than individualism is widely albeit implicitly acknowledged. The co-evolution of events, political trends and economic thought is apparent. Economic thinking has been responding to financial crises in 1997/98, 2001 and 2008, to the voter backlash against globalisation and 'elites', as well as to longstanding failures such as the inability of economists to explain growth or its absence.

For instance, in addition to the embrace of behavioural economics, already noted, the importance of institutions is widely acknowledged. Old habits die hard and so this can be quite a reductive approach too, but it is a great step forward: institutions by definition involve more than one individual, and are located in place and time. Governments, public bodies, universities, firms, co-ops, charities, temples, unions, families are recognised as alternative means of reaching collective decisions about the use and allocation of resources. Institutions are shaped by asymmetric information and transactions costs – which are standard in modern economics – and also by social preference formation, which is not (Bowles, 2004).

Interdependence is by definition recognised (in some defined ways) in game theory, which has wide influence in economics. The market design literature is similarly inherently concerned with interdependence. In the context of digital markets and financial markets, network theory is of growing importance, where the existence and even the identity of other individuals is core. Environmental economics focuses on externalities, as do studies of network markets. Modern growth theory makes growth depend on knowledge spillovers, people learning from each other, and spillovers are generally seen to be pervasive in advanced economies. There is a growing interest in economic applications of the science of complexity (see for example Colander and Kupers, 2009) and in the lessons of evolutionary theory (see for example Lo, 2017). In short, compared to the 1980s and 90s, the research frontier in economics is transformed (Coyle, 2007).

Yet the technocratic instinct and the legacy of free market politics leave economic policies still stranded in a narrower and less empirical version of economics than is actually practiced in the academy now. Economic policy makers who were students in earlier decades have firmly internalised the free market framing. Active researchers are engaged in a richer and more realistic economics, in terms of behavioural assumptions and institutional context for example, but have given too little thought to the welfare economics framework, which inevitably constrains the application of their findings.

Nor has there been any rush by the profession to address the critique that economic policies resting on the assumption of rational self-interest actually change individual behaviour. Similarly, the wider possibility that economics is performative does not resonate with economists (Coyle, 2012). Although the practice of economics has changed substantially over recent decades, reflecting an implicit recognition of interdependence rather than individualism, the one area that has not changed is welfare economics. Unless economists revisit the foundations of welfare economics, our ability to respond to today's policy questions will remain limited.

Conclusions

This paper has discussed two common criticisms of economics. One relates to the hypothetical *homo economicus*, a self-interested, rational maximiser. One is that this assumed character is not realistic. The other is that economics can be confined to technical analysis separate from ethical issues.

I have argued that the first of these criticisms is overstated. It is not obvious where biology, neuroscience and psychology will take us: AIs, humans and pigeons (or fungi) can act like rational, self-interested individuals, yet their cognitive capacities clearly differ. We do not yet understand which aspects of context determine when people (or other entities) act in the individual rational choice mode or making 'behavioural' decisions shaped by social or psychological factors, or by rules of thumb.

What's more, economists are broadly interested and open to this research agenda. To the extent that models make a rational choice assumption, this is increasingly seen as a starting point. Applied economists, who make up the majority of the profession by numbers, are often more than willing to test alternative models of decision-making. This research agenda is congruent with economists' self-image as empiricists and technocrats. Of course, there are exceptions, but testing different models of individual decision-making poses no fundamental issues for policy economists interested in 'what works'.

The second criticism poses a more fundamental challenge to economics. It goes to some deep issues in welfare economics, a field little debated in recent times although making welfare judgements is nevertheless an everyday part of practical economics. Although a sequence of highly distinguished economists from the 1930s on have pointed out a number of logical inconsistencies in the positivist framework that has dominated economics for nearly a century, nevertheless economists have continued to insist on the separability of the positive and the normative, and on the possibility of evaluating policies in a technocratic manner while ruling out interpersonal utility comparisons and relying on the Pareto criterion.

A consequence of ruling out by assumption interdependencies such as social influences on preferences, economies of scale and external costs and benefits has been the embodiment in much economic advice of the conclusions of the first and second welfare theorems: in short, that 'the market', free exchange, will deliver optimal outcomes; and that specific market failures can be identified and corrected. Although government failure certainly exists, it tends to be in the same instances as market failure and for the same reason: because the individual and collective interest diverge.

There were other drivers of the public policy emphasis on 'free markets', political and economic, as well as economic thinking. However, the changing character of modern economies, as well as the tide of events, means it is important to revisit this benchmark. In many digital markets, for instance, there are large externalities and economies of scale, knowledge spillovers are pervasive, and goods are non-rival. Economists need to provide the theoretical and empirical tools for public policy makers to assess these markets.

Economists have implicitly accepted this in some recent trends in the subject toward the study of interdependencies, such as institutional economics and market design. However, in contrast to the 'is' agenda of understanding how humans make choices in comparable contexts, where critics also have something to learn from research in other disciplines, economists have not engaged with the 'ought' critique. It goes to the fundamentals of economics but this makes engagement all the more important.

REFERENCES

- Arrow, K, 'A Difficulty in the Concept of Social Welfare', *Journal of Political Economy*, Vol. 58, No. 4. (Aug., 1950), pp. 328-346.
- Atkinson, A, 'The Strange Disappearance of Welfare Economics', *Kyklos*, 2001, Vol 54, pp 193-206.
- Baumol, W.J. 'Community Indifference', *The Review of Economic Studies*, Vol. 14, No. 1. (1946 - 1947), pp. 44-48.
- Baumol, W.J., *Welfare Economics and the Theory of the State*, The London School of Economics and Political Science/ Longmans, Green & Co., 1952.
- Becker, G.S. 'Irrational Behavior and Economic Theory', *Journal of Political Economy*, Vol 70, No 1, (Feb 1962), pp1-13.
- Bergson, A, 'A Reformulation of Certain Aspects of Welfare', *The Quarterly Journal of Economics*, Vol. 52, No. 2. (Feb., 1938), pp. 310-334.
- Blaug, M. *The Methodology of Economics: Or How Economists Explain* 5th ed 1996 (first pub. 1962) Cambridge University Press.
- Bowles, S, *Microeconomics: Behavior, Institutions and Evolution*, Princeton University Press, 2004.
- Bowles, S, *The Moral Economy*, Yale University Press (2016).
- Chetty, Raj. 2015. "Behavioral Economics and Public Policy: A Pragmatic Perspective." *American Economic Review*, 105(5): 1-33.
- Colander, D and R Kupers, *Complexity and the Art of Public Policy*, Princeton University Press (2012).
- Coyle, D, *The Soulful Science*, Princeton University Press (2007).
- Coyle, D, 'The Public Responsibilities of the Economist', Tanner Lectures in Human Values, Oxford University, 2012.
- Coyle, D, *GDP: A Brief But Affectionate History*, Princeton University Press (2014).
- David, P.A., 'The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox', *American Economic Review*, 80 (2), 1990, 355-361.
- De Waal, F, *Primates and Philosophers: How Morality Evolved*, Princeton University Press, 2006.
- Dietz, S & Hepburn, Cameron. 'Benefit-cost analysis of non-marginal climate and energy projects' *Energy Economics*, Elsevier, vol. 40(C), 2013, pages 61-71.

Dreze, J and Stern, Nicholas, The theory of cost-benefit analysis, ch. 14, p. 909-989 in Auerbach, A. J. and Feldstein, M. eds., *Handbook of Public Economics*, vol. 2, Elsevier, 1987.

Friedman, M, 'The Methodology of Positive Economics', reprinted in *Essays In Positive Economics* University of Chicago Press, 1966.

Graaff, J de V, *Theoretical Welfare Economics*, Cambridge University Press, 1971.

Hamermesh, D S. 'Six Decades of Top Economics Publishing: Who and How?' *Journal of Economic Literature* 5(1) (2013) pp162-72.

Hammerstein, Peter & Ronald Noë, 'Biological Trade and Markets', *Philosophical Transactions of the Royal Society B*, 2016 371 20150101; DOI: 10.1098/rstb.2015.0101. Published 4 January 2016

Harberger, A.C., 'Three Basic Postulates for Welfare Economics: An Interpretive Essay', *Journal of Economic Literature*, Vol. 9, No. 3 (Sep., 1971), pp. 785-797.

Herbranson, W., & Schroeder, J. (2010). 'Are birds smarter than mathematicians? Pigeons (*Columba livia*) perform optimally on a version of the Monty Hall Dilemma.' *Journal of Comparative Psychology*, 124 (1), 1-13.

Hicks, J.R., 'The Foundations of Welfare Economics' *The Economic Journal*, Vol. 49, No. 196. (Dec., 1939), pp. 696-712.

Hurley, S., & Nudds, M. (2006). *Rational Animals?* Oxford: Oxford University Press.

Jackson, Matthew O. and Yariv, Leeat, The Non-Existence of Representative Agents (September 7, 2017). Available at: <https://ssrn.com/abstract=2684776> or <http://dx.doi.org/10.2139/ssrn.2684776>

Kaldor, N., 'Welfare Propositions of Economics and Interpersonal Comparisons of Utility', *The Economic Journal*, Vol. 49, No. 195. (Sep., 1939), pp. 549-552.

Kelman, S, 'Cost Benefit Analysis: An Ethical Critique', *Regulation*, January/February 1981, pp33-40.

Kumekawa, I., *The First Serious Optimist*, Princeton University Press, 2017.

Le Grand, J. 'The Theory of Government Failure', *British Journal of Political Science*, (1991) 21(4), 423-442.

Lo, A, *Adaptive Markets: Financial Evolution at the Speed of Thought*, Princeton University Press, 2017.

Joel Z Leibo et al, 'Multi-agent Reinforcement Learning in Sequential Social Dilemmas', in Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, Sao Paulo, Brazil, <https://storage.googleapis.com/deepmind-media/papers/multi-agent-rl-in-ssd.pdf>

Lipsey, R.G. and Kelvin Lancaster, 'The General Theory of Second Best', *The Review of Economic Studies*, Vol 24,1, 1956, pp 11-32.

Pesendorfer, W. 'Behavioral Economics Comes of Age: A Review Essay on Advances in Behavioral Economics.' *Journal of Economic Literature*, 44(3) (2006): 712-721.

A C Pigou, *Economic Science in Relation to Practice*, Macmillan, London (1908).

Piketty, T, *Capital in the 21st Century*, Harvard University Press (2014).

Roberts, K., 'Price-independent welfare prescriptions', *Journal of Public Economics*, Volume 13, Issue 3, 1980, Pages 277-297.

Robbins, L., *The Nature and Significance of Economic Science*, Macmillan, London (1932).

Rothschild, E, *Economic Sentiments: Adam Smith, Condorcet, and the Enlightenment*, Harvard University Press, 2001.

Samuelson, P.A., *Foundations of Economic Analysis*, Chapter VIII, 'Welfare Economics', (1947, 1983)

Sandel, M.J., *What Money Can't Buy: The Moral Limits of Markets*, 2012.

Scitovszky, T de, 'A Note on Welfare Propositions in Economics. *The Review of Economic Studies*, Volume 9, Issue 1 (Nov., 1941), 77-88.

Sen, A., *Collective Choice and Social Welfare*, Penguin, expanded edition 2017 (first published 1970)

Sen, A., *The Idea of Justice*, Allen Lane, 2009.

Smith, A. *The Theory of Moral Sentiments*, Prometheus Books, New York, 2000, first publ. 1759

Smith, J, *A Random Physicist Takes on Economics*, 2017.
www.arandomphysicist.com