

## *Introduction*

ADAM OLIVER

For much of the nineteenth century, a discursive, almost exclusively theory-driven style of political economy dominated the economics discourse. Among much else, for instance, David Ricardo had stated that the value of a good is proportional to the cost of the labour taken to produce it, Thomas Malthus believed that population growth would inevitably lead to famine, and J. S. Mill and Karl Marx warned that wages would never rise much above subsistence levels. However, in many cases, of course, brilliantly prosecuted, empirical analysis to test the often competing claims was somewhat lacking. Due to the absence of tools to provide seemingly objective answers, the perceived usefulness of political economy for informing and assessing policy was ultimately undermined.

Alfred Marshall aimed to address this state of affairs by making economics more scientific, that is, more testable and tested. Recognizing that the world is complex and changing, and that preconceptions may be less than valid, he undertook ground-breaking empirical work to demonstrate that, in contrast to the forewarnings of Mill and Marx, wages were in fact increasing over time, as a consequence of greater productivity necessitated by competition. Marshall may have paid insufficient attention to the possibility that not all workers receive an equitable share of the fruits of competition, but, by empirically addressing testable theoretical propositions, he revolutionized economics. From that point onwards, mathematical and applied empirical economics superseded political economy as the dominant force within the mainstream economics community.

The measurement of expected utility is an important branch of economics that benefited substantively from the introduction of mathematics. By the end of the nineteenth century, many economists had turned away from the attempt to derive cardinal measures of utility, and, in subsequent decades, built welfare economics instead upon Vilfredo Pareto's famous ordinal criterion, a development that Luigino

Bruni and Robert Sugden (2007) have coined the Paretian turn. However, between the 1920s and the early 1950s, a number of remarkable mathematicians and mathematically minded economists, including Frank Ramsey, Leonard Savage, John von Neumann, Oskar Morgenstern, Jacob Marschak and Paul Samuelson, contributed towards developing axiomatic systems of formal logic that prescribed how people ought to choose under risk and uncertainty if they are expected utility maximizers. This movement culminated in the specification of expected utility theory.

The standard gamble, a tool that theoretically enables the elicitation of an individual's cardinal utilities for goods, was developed from the axioms of expected utility theory, and thus appeared to provide a solution for the derivation of strength of preference indices. However, eliciting individuals' preferences relies principally upon how they do choose (i.e. on descriptive behaviour), rather than on the normative proposition of how they ought to choose if they are expected utility maximizers, and thus the validity of the theory for utility elicitation is informed by observed choice. The axioms of expected utility theory offered themselves to formal testing, and, as early as the 1950s, findings that undermined the descriptive validity of expected utility theory – perhaps the first empirical behavioural economics results – were published. The most frequently tested and contested assumption of the theory has proven to be the independence axiom, which essentially states that the subjective value that people attach to any particular outcome ought not to be affected by the probability with which that outcome occurs, and that their perception of any particular probability ought not to be influenced by the outcome that is attached to that probability. The implication of independence is that when an individual is presented with a choice of two or more options, he or she should deem any outcome common to the options as irrelevant to his or her decision.

The earliest and most famous violation of the independence axiom was postulated in print by Maurice Allais in 1953. To illustrate, consider the following four options, A, B, A\* and B\*:

A: \$1m for certain

B: 10% chance of \$5m; 89% chance of \$1m; 1% chance of \$0

A\*: 11% chance of \$1m; 89% chance of \$0

B\*: 10% chance of \$5m; 90% chance of \$0

Allais contended that if individuals were faced with two choice contexts – i.e. A or B and  $A^*$  or  $B^*$  – they would show a tendency to choose A and  $B^*$ . However, close examination reveals that A and B share the common outcome of an 89 per cent chance of \$1m, and therefore, according to the independence axiom, the chooser should consider this common consequence as irrelevant to his decision. Similarly,  $A^*$  and  $B^*$  share a common outcome of an 89 per cent chance of \$0. If common consequences are ignored in both choice tasks, A is identical to  $A^*$  and B is identical to  $B^*$ ; therefore expected utility theory requires the individual to choose A and  $A^*$  or B and  $B^*$ , or neither in both contexts. Allais's conjecture of frequent observance of people choosing A and  $B^*$  has been confirmed in much subsequent empirical testing, and has been labelled the common consequence effect.

The Allais paradox is attributed to the certainty effect – the high weight that individuals tend to attach to certainty, higher than that allowed by expected utility theory. Ambiguity, where the exact probability of an uncertain event is unknown, is even less certain than risk (where the probability is clearly specified), and ambiguity aversion can cause further violations of independence. Daniel Ellsberg's classic thought experiment, published in 1961, is an example of this. Ellsberg posed a scenario where an individual is presented with an urn containing 30 black balls and 60 red and blue balls of an unspecified mix and is requested to pick one ball from the urn. Before picking a ball, the individual is asked to consider the following two choices, A or B and  $A^*$  or  $B^*$ , where:

A: \$100 if the ball is black, \$0 if red or blue

B: \$100 if red, \$0 if black or blue

$A^*$ : \$100 if black or blue, \$0 if red

$B^*$ : \$100 if red or blue, \$0 if black

In the choice between A and B, A therefore offers a definite one-third chance of winning \$100 (30 of the 90 balls will be black); B an ambiguous zero to two-thirds chance of winning \$100 (anything between 0 and 60 of the 90 balls will be red). As to  $A^*$  versus  $B^*$ ,  $A^*$  offers an ambiguous one-third to certain chance of winning \$100;  $B^*$  a definite two-thirds chance of winning £100. If a blue ball is picked in A or B, \$0 is won, and thus, according to the independence axiom,

the chooser should perceive the possibility of picking a blue ball as irrelevant to their decision in the first choice context. The same applies in the second choice context, where a blue ball in both  $A^*$  and  $B^*$  offers a common consequence of \$100. If consideration of the blue ball is eliminated from both choice contexts, then  $A$  is identical to  $A^*$ , and  $B$  is identical to  $B^*$ ; thus, according to expected utility theory, the individual should choose  $A$  and  $A^*$  or  $B$  and  $B^*$ , or consistently neither. However, Ellsberg hypothesized that many people will be ambiguity averse, and will therefore demonstrate a systematic preference for the risky but unambiguous options,  $A$  and  $B^*$ , a conjecture that has since been amply confirmed. In this volume, I offer a chapter that discusses ambiguity aversion in a little more detail, and which uses the concept as an explanation for the UK government's arguably excessive response to the 2009 swine flu outbreak.

During a period concurrent with publication of the early literature that demonstrated specific violations of the axioms of expected utility theory, Herbert Simon (1956) offered a challenge to the broader underlying assumption, common to welfare economics and much of rational choice theory, that people are optimizers who always seek to maximize utility. Simon argued that, due to bounds on their rationality and time, people do not expend the effort required to optimize, and instead use relatively simple rules of thumb, or heuristics, to help them reach decisions. In short, they 'satisfice', or make do with something that is good enough. Psychologists have since reported a great many heuristics that drive decision-making in particular contexts. Some of the most prominent of these rules of thumb include the representativeness, availability and anchoring heuristics. Representativeness, for instance, is the finding that people may overlook the objective probability of something happening, and resort more to preconceptions or stereotyping; availability is the tendency for people to assess the probability of an event by the ease with which similar instances can be brought to mind; and anchoring is the finding that people often place a heavy emphasis on particular salient features of a decision context – or even on entirely irrelevant prompts – and insufficiently account for less prominent characteristics.

The early concerns with expected utility theory were not initially treated with a great deal of respect by the mainstream economics community. Leonard Savage, on violating his own sure thing principle (essentially the independence axiom under conditions of uncertainty),

attributed his choices to mere error, rather than a deliberate, systematic preference pattern. Nonetheless, during the 1960s and 1970s, psychologists continued to compile a growing body of evidence that placed serious question marks against, at the very least, the descriptive validity of the theory. Prominent among this evidence was the reporting of preference reversals, which, in their classic form, involve the offering of two bets, referred to as the *P-bet* and the *\$-bet*. The *P-bet* offers a high probability of winning a modest amount, the *\$-bet* offers a modest probability of winning a relatively large amount, and the two bets have similar expected values. People are asked to choose directly between the *P-bet* and the *\$-bet*, and are also asked for the value, in the form of approximate certainty equivalents, that they place on each of the two bets. It has been frequently reported that a substantial percentage – often a majority – of respondents choose the *P-bet* over the *\$-bet*, but value the *\$-bet* higher than the *P-bet*. For illustrative purposes, consider the following two bets, taken from the classic 1971 study by Sarah Lichtenstein and Paul Slovic:

*P-bet*: (\$4, 35/36; −\$1, 1/36)

*\$-bet*: (\$16, 11/36; −\$1.50, 25/36)

Thus, the *P-bet* offers a 35/36 chance of winning \$4 and a 1/36 chance of losing \$1. The *\$-bet* can be similarly read. In three tests, Lichtenstein and Slovic observed systematic preference reversals, in the direction of people choosing the *P-bet* but valuing the *\$-bet* higher, in something between 50 and 80 per cent of their respondents. It is unlikely that such a substantial, systematic pattern can be attributed to random error.

Several explanations for preference reversals have been suggested, but the most likely cause is that people use different heuristics across elicitation procedures. Choice tasks might encourage greater focus on the probability of winning, which favours the *P-bet*, while valuation tasks may tend to focus attention on the payoffs, which of course favours the *\$-bet*. Specifically, when valuing the *\$-bet*, people often anchor on its best outcome, but then fail to adjust the overall value of the bet downwards sufficiently to take account of its other less favourable attributes. This notion of anchoring and insufficient adjustment to account for the less desirable features of a choice option is, as earlier noted, a key finding in the behavioural economics literature, and

resonates to some extent in the chapter in this volume by Sunita Sah, Daylian M. Cain and George Loewenstein, where they argue that if financial advisers exaggerate, or further exaggerate, the usefulness of their advice due to having to disclose any conflicts of interest, their clients might anchor on the biased advice and insufficiently adjust their decisions to account for the bias. More generally, psychologists, in contrast to standard economic theory, became increasingly convinced that preferences are not always fixed and stable, but are often constructed in response to how they are elicited and how choice contexts are framed. Paul Slovic today still continues to study the psychological processes that can cause people to violate the assumptions of rational economic man, and in this volume, together with Daniel Västfjäll, he provides a chapter that partially attributes the insufficient sensitivity that most people feel towards large numbers of statistical deaths to psychic numbing, i.e. the observation that it is impossible to multiply the intensity of our feelings towards good or bad events by very large magnitudes.

Scepticism towards the preference reversal findings by members of the economics community somewhat ironically led to a degree of acceptance that phenomena that cannot be explained by mainstream theory are, indeed, genuine. In 1979, the economists David Grether and Charles Plott reported a study that controlled or corrected for many of the problems that they, possibly somewhat unfairly, perceived as inherent to the psychologists' work on preference reversals. For example, responding to what they saw as flaws in the methods of prior studies, Grether and Plott included real financial incentives (although Sarah Lichtenstein and Paul Slovic had actually already included financial incentives in field experiments on gamblers in Las Vegas), refrained from forcing people to make a choice between bets by allowing them to express indifference and attempted to remove the motivation for respondents to hedge their bets, by informing them that at the end of the experiment only one of the questions, chosen randomly, would be played out for real. Preference reversals were as significant and systematic as they were in the earlier psychologists' work, which convinced at least some mainstream economists that the phenomena could not be casually brushed aside.

Also in 1979 came the publication of a monumental work in the history of behavioural economics, namely Daniel Kahneman and Amos Tversky's prospect theory. In prospect theory, Kahneman

and Tversky proposed two major modifications to expected utility theory. The first is that in prospect theory, the carriers of value are gains and losses around a reference point, rather than final assets, and losses are weighted substantially more – at least twice as much – as gains of the same magnitude. Second, people transform probabilities in an inverse S-shaped pattern, such that they overweight small probabilities, underweight large probabilities and perceive subjective probability to equal objective probability at approximately 0.4. Perhaps rather ambitiously, Kahneman and Tversky claimed that prospect theory could account for most of the major violations of expected utility theory. The latter nonetheless remains the dominant theory of rational choice, a state of affairs that in 1990 caused Maurice Allais to complain that for ‘nearly forty years the supporters of [expected utility theory] have exerted a dogmatic and intolerant, powerful and tyrannical domination over the academic world; only in very recent years has a growing reaction begun to appear. This is not the first example of the opposition of the “establishments” of any kind to scientific progress, nor will it be the last’ (Allais, 1990, p. 8). For a discipline that had for a century been developed with scientific credentials in mind, the cursory dismissal by the majority of mainstream economists of empirical evidence that falsified the axioms on which much of standard theory is based was not particularly edifying. Subsequent to Allais’s statement, however, behavioural economics has gained more acceptance as a sub-branch of economics, culminating in Kahneman being awarded the 2002 Nobel Memorial Prize in Economic Sciences.

The growing ‘reaction’, which gathered momentum during the 1980s, produced a cottage industry of behavioural economics, encompassing theory-driven work, such as the development of several alternatives to expected utility theory, and research that further tested and questioned the axioms of standard theory. Some of the alternatives to expected utility theory were mathematically motivated in that they directly weakened the axioms underlying the theory so as to allow a wider array of preference patterns; others, such as prospect theory, were modifications based on observed psychological processes, arising from the view that the logic-based axioms of standard theory overlooked the fact that people are human. Of all the alternatives, prospect theory, although still deep in the shadow of expected utility theory, is the most influential, at least partly because Kahneman and Tversky offered numerical parameters for loss aversion and probability

weighting, which gives the theory predictive power, and partly because the theory does indeed resonate with how people often appear to reach their decisions. A substantial literature on the anomalies to standard economic theory was published during this time, many of which were reviewed by Richard Thaler in a series of articles in the *Journal of Economic Perspectives*, later collectively printed in his book, *The Winner's Curse* (Thaler, 1994). Thaler was an early pioneer in empirical behavioural economics, and worked with Kahneman and Tversky on experiments on loss aversion, for example. He paved the way for other luminaries in the field, including George Loewenstein, Drazen Prelec and, later, Matthew Rabin, all of whom contribute, or contribute to, chapters in this volume.

During the 1990s, George Loewenstein wrote extensively on people's attitudes towards discounting time. One particular aspect of this general area of interest that has attracted much attention in the behavioural economics discourse is present bias, otherwise known as hyperbolic discounting. Present bias is the observation that people attach an enormous weight to the immediate moment, irrespective of whether that moment is pleasurable or painful, and consequently underplay the importance of all subsequent moments. Present bias may well explain why many people consume too much tasty but unhealthy food, drink excessive quantities of alcohol, refrain from exercise and accumulate substantial debts, overlooking the long-term consequences of their actions. In his chapter, Matthew Rabin argues that a combination of present bias and projection bias, where people do not rationally predict future tastes and overlook the extent to which current behaviours are likely to become habits, can cause people to embark upon and maintain unhealthy lifestyles.

A further issue that has attracted much attention in the behavioural economics community is the notion of motivational crowding out, an area over which Bruno Frey has been a leading researcher since at least the 1990s. Motivational crowding out is the observation that paying people to do something may crowd out their intrinsic or altruistic motivation to do that very thing. For example, to allude to Richard Titmuss's classic 1970 study, *The Gift Relationship*, if a person is altruistically motivated to donate blood voluntarily, that motivation may be eroded if he or she is offered monetary recompense for the donation, because the act would no longer be altruistic, and would more closely resemble a market exchange. For such acts, therefore, the



offer of money could lead to less action. This is potentially at odds with the relative price mechanism of standard economic theory, which implies that offering people money, or more money, to do something will increase the likelihood that they will do it. In this volume, Kate Disney, Julian Le Grand and Giles Atkinson argue that an external reward may crowd in altruistic behaviour if it is perceived as supporting or reinforcing an autonomous action, but crowd out altruism if the reward is perceived as controlling. By analysing data in relation to a small charge – i.e. a negative payment – for plastic bags imposed on shoppers in a supermarket, they claim that the charge reinforced a moral norm that people should be reusing their shopping bags, which, they argue, may rebut the idea that charging for environmental goods crowds out intrinsic concern and creates a society of selfish individuals who have no social awareness. Bruno Frey also offers a chapter in this volume, and argues that non-monetary incentives, such as awards and honours, may often be a more powerful tool for motivating people than money, because such incentives might strengthen people's identity with the organization for which they work. Awards may therefore crowd in intrinsic motivation.

From the 1990s to the present, a number of behavioural economists have come onto the scene, many inspired no doubt by those who are these days thought of as the old masters, Kahneman, Tversky and Thaler. Among these are Dan Ariely, David Laibson and Sendhil Mullainathan, but there are also of course a host of others doing interesting and potentially important work in the area. Not all of these are based at US institutions, and some have been undertaking work in behavioural economics for almost as long as the pioneering figures. For example, Ernst Fehr in Switzerland and Robert Sugden in the UK, to name but two, have made significant contributions to behavioural economics over the last three decades. A thorough review of behavioural economics over even the last ten years is beyond my scope here, but three developments will be highlighted. The first relates to some extent to Bruno Frey's chapter mentioned above, namely George Akerlof and Rachel Kranton's work on identity economics in the 2010 book of the same name. Akerlof and Kranton propose that people will experience positive utility from working for an organization with which they identify, and negative utility if they perceive themselves to be outsiders, and yet this utility, experienced from feeling that one belongs or otherwise, is not incorporated into standard

economic theory. According to Akerlof and Kranton, our identity defines who we are – our social category – and will influence our behaviours, because different behavioural norms are associated with different social categories. If we take firms, for example, good managers, according to the theory of identity utility, will want their workers to be motivated insiders who identify with the goals of the firm, rather than alienated outsiders. If this is achieved, then employees will want to work enthusiastically towards the objectives of the firm, irrespective of additional personal financial rewards, because they will intrinsically support the firm's mission. This theory can be applied across all conceivable public and private organizations, and may, in monetary terms, offer an inexpensive way to motivate individuals.

A second notable development, or, rather, event, of the past decade was the publication of Daniel Kahneman's intellectual autobiography in 2011. There will of course be many who dispute the general importance of Kahneman's many contributions, but there can be little argument that he is the greatest living behavioural economist, all the more remarkable since he openly admits to having a limited knowledge of economic theory. His substantial impact is and will remain beyond doubt, from his development of prospect theory with Amos Tversky, to his work with a number of colleagues on the gestalt characteristics, or in other words his observations that people's remembered and decision utility often fail to correspond in systematic ways to the utility that they experience. This latter observation cuts a chink in the armour of welfare economics, and has informed, for good or ill, the development of the new economics of happiness. At my own institution, the London School of Economics and Political Science, Kahneman's book is required reading on a number of courses, and everyone who wants to engage with behavioural economics ought to read it.

However, perhaps the most significant development in behavioural economics since 2000 has been the increasing efforts to apply approximately three decades of behavioural economics observations to practical policy concerns. These policy efforts have been principally underpinned by philosophical frameworks – essentially, soft forms of paternalism – developed by some of the world's leading behavioural economists, and include asymmetric paternalism, formulated by Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue and Matthew Rabin in 2003. The most famous and influential of these frameworks is libertarian paternalism, developed by Richard Thaler