

1

Introduction

Bina Agarwal
University of Delhi, India

and

Alessandro Vercelli
University of Siena, Italy

Introduction

‘How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it.’

(Adam Smith, [1759] 1966: 3)

‘[T]he insistence on the pursuit of self-interest as an inescapable necessity for rationality subverts the “self” as a free, reasoning being, by overlooking the freedom to reason about what one should pursue.’

(Amartya Sen, 2002: 46)

Economics today is at an exciting stage of evolution, as it begins to reopen routes of interchange with other disciplines. In this interdisciplinary exchange, psychology stands closer to centre-stage than most other disciplines. It has provided the grist for challenging many standard economic assumptions and catalysed the rapidly growing fields of behavioural economics and experimental economics.

For most part of the twentieth century, and especially since the 1950s, the characterization of the human being as *Homo economicus* dominated economics. Strongly influenced by Newtonian Physics, to which is traced the formalization of modern economic theory, the underlying approach made the psychological characteristics of the economic agent largely irrelevant. Indeed, the approach acted as a protectionist barrier against insights both from other social sciences and the humanities, and from the observed complexity of human behaviour in real life. The last two decades, however, have brought an emerging recognition of the crucial role played

by the psychological attributes of economic agents in explaining economic behaviour.¹ This has paved the way for a less reductionist approach to the subject.

Under standard economic assumptions, human beings are presumed to be narrowly rational, motivated by individual self-interest, preoccupied with maximizing personal utility or satisfaction, driven by cold economic calculation without concern for others, capable of instantaneous learning, and so on. Real people are found to be more complex, driven not just by self-interest but also by altruism, guilt, liking, and other emotions;² caring not just about what they themselves get but also about what others get, and capable of forgoing personal payoffs for an equitable distribution of gains; wanting fair treatment and not simply higher incomes; influenced in their choice of workplace not just by considerations of profit or income maximization but also by emotions which tie them to a particular work environment and to fellow workers; learning slowly from mistakes rather than being able to self-correct instantaneously; affected by 'money illusion' which shapes their economic expectations;³ capable of self-deception about their abilities and exuding levels of confidence (or lack thereof) inconsistent with their real talents; and so on. In other words, the introduction of psychology into economic behaviour provides a tool for challenging a central assumption of standard economics – that of *Homo economicus* – in a way that strikes at the heart of microeconomic theory, where this assumption is embedded.

Indeed, as soon as we open the door of economics to psychology we introduce a breach in the traditional dyke that forestalls an analysis of the motivations and cognitive states of agents. (Or perhaps we should say, *reopen* the door to psychology, since in the nineteenth and early twentieth centuries a number of eminent economists – such as Nassau Senior, William Jevons, Irving Fisher, John Maynard Keynes, among others – did stress psychological factors in economic behaviour.⁴) Basically, psychology has pushed economists to recognize the heterogeneity of human responses and to examine a range of neglected factors that impinge on what motivates people and how individuals deal with complexity.

While many of these challenges can be explored theoretically, testing them empirically is equally important. Here too economists are now drawing on a tool which has long been standard in psychology – namely the use of controlled experiments to assess how people are likely to behave in given settings and contexts. Indeed, experimental economics has grown rapidly since the 1960s, and proved especially important in the development of behavioural economics. As is well-known, it is basically a tool to collect information by creating an artificial, controlled environment in a laboratory (typically a classroom) and using the tight experimental control to examine the relevance of psychological motives in economic decision-making. True, there are limitations to lab-based experiments, since the situations tested and responses obtained might deviate from the complexity of real-life situations

and responses in radical ways. But, arguably, this limitation can be overcome, at least in part, by setting up control experiments in the field among people facing similar real-life situations. This is now being done by some scholars (including Cardenas in this volume).

To these growing fields of behavioural economics and experimental economics, this book seeks to make a contribution. It brings together (in substantially revised form) a collection of papers that were presented as papers or as ideas at the Thirteenth World Congress of the International Economic Association, held in Lisbon in September 2002. All the chapters, in one way or other, focus on the insights that psychology provides in understanding economic behaviour, and the challenges it poses to standard economic assumptions.

The book examines these interfaces from several angles, but two ideas are especially dominant and constitute running threads in the volume. These are rationality and altruism. The interchange between economics and psychology inevitably brings a focus on what we understand by rationality. The rationality entertained by mainstream economics is quite different from that entertained by psychology, and the tension between the two approaches was pointed out by some scholars a long while ago. Simon (1976), for instance, highlighted the divergence of vision and methodology between the two disciplines by characterizing the notion of rationality adopted in standard economics as 'substantive' rationality (emphasizing its unbounded character, its exclusive focus on optimizing equilibrium and self-interest maximization, and its narrow requirements of intertemporal coherence), and characterizing the notion of rationality adopted in standard psychology as 'procedural' (or 'bounded') rationality.

Another sharp critic of the narrow view of rationality found in mainstream economic theory – Amartya Sen (1977, 2002) – has emphasized the link between rationality and freedom. Rationality, he argues, depends on freedom not only because without some freedom of choice, the idea of rational choice becomes vacuous, but also because 'the concept of rationality must accommodate the diversity of reasons that may motivate choice' (2002: 5). Self-interest maximization can at best be seen as one among many goals that a person might choose to pursue, but a canonical selection of this one goal as an exclusive guide to rationality, and a rejection of all other motives and concerns that a person may have, would effectively involve, according to Sen (2002: 5, his emphasis) 'a basic denial of *freedom of thought*'.

Indeed, criticisms of narrow economic rationality have grown over the years, and the tension between the standard economics approach and the broader view of rationality emphasized in psychology, as well as in some heterodox contributions in economics, emerges in each of the chapters included here. Narrow rationality does not find corroboration in any of the analytical models and experiments contained in the book. Rather, several of the authors argue that apparently irrational behaviour can be seen as rational in the light

of the models suggested by them. Observed behaviour that appears to be inconsistent with a narrow view of rationality can still be considered rational from a broader and more comprehensive perspective. Extending the scope of what is deemed rational also helps extend the scope of economics itself, such as by relaxing the most demanding assumptions (mentioned above) that underlie the orthodox concept of rationality. In doing so, economics also reduces the gap between its own assumptions and methods and the assumptions and methods of psychology.

The second dominant idea that this volume explores is that of altruism. This can also be seen as related to the larger project of broadening what is deemed rational, beyond the single-minded pursuit of self-interest to the admission of other motivations, in particular 'other regarding' motivations. Several chapters emphasize the importance of altruism as one of the guiding forces of human behaviour – examining its origins, how it evolves, and what its implications are within the realm of the household, the workplace and the community. Altruism can impinge on many economic attitudes – 'the motivation to produce, the propensity to distribute, and the tendency to accumulate and transfer – within families, societies and across generations' (Stark and Y.Q. Wang, this volume). It can also impinge on relationships between co-workers and between workers and employers, on incentive schemes, and on work contracts. And in the context of communities, it can throw light on how institutions evolve or dissolve, whether or not people cooperate socially, and so on.

Discussions on these and related concepts are enriched by insights from other disciplines. In fact, the chapters in this book can be seen as a testament to the growing scientific exchange between economics and other disciplines and fields – the social sciences and humanities on the one hand, and the natural sciences (in particular, evolutionary biology) on the other. All the chapters draw important insights from psychology, but in addition some draw also on other fields such as cognitive sciences (Bénabou and Tirole), evolutionary biology (Stark, Y.Q. Wang and Y. Wang; Vasin), ethology (Vasin), epistemology (Fehr and Tyran; Vercelli), ethics (Jungeilges and Theisen), and political science (Cardenas). On the one hand, the chapters provide extensions, developments and new proofs of psychological ideas by using the tools of economics (game theory, principal-agent theory, intertemporal maximization, and so on). On the other hand, these contributions have the potential for influencing other disciplines, thus progressively broadening the scope of interdisciplinary exchange. Cases in point are the influence of microeconomics on evolutionary biology, apparent in the analogy being drawn between the selfish gene and *homo economicus* (Dawkins, 1990); and on neuroscience, apparent for instance in the argument that economic theory may provide an alternative to the classical Cartesian model of the brain and behaviour (Glimcher, 2003). In turn, an important theoretical insight in evolutionary biology, the 'Hamilton rule' (which relates to altruism between relatives),

is introduced in this volume by Stark and Y. Wang who clarify and extend its validity to economic issues by using a standard modelling device in economics, the prisoner's dilemma.⁵

All the chapters have worked to substitute narrow economic assumptions with more comprehensive ones. Some of the authors view this ongoing process of innovation in economic theory as a mere progression of standard theory; others see it as a paradigmatic shift. For instance, many chapters in this volume have modified standard utility functions by adding a term that takes account of inequality aversion, altruism and/or other social and ethical propensities. This modification could be interpreted as a generalization of traditional utilitarianism; or, as some authors contend, it could be seen as a paradigm shift away from the pure consequentialism and selfishness of traditional utilitarianism to a different vision, one that is open to social motivations and deontological principles. We leave our readers to judge for themselves which view they are inclined towards, after reading the chapters.

The book is divided into two parts. Part I explores the interface of economics and psychology theoretically, in terms of analytical models and methodological issues, while Part II uses the technique of experiment economics to explore this interface empirically.

Part I Analytical models and methodological issues

The volume opens with Bénabou and Tirole's chapter on 'Self-Confidence and Personal Motivation', which touches on many strands of the new literature that links economics and cognitive psychology. In particular the authors focus on the psychological trait of self-confidence and its effects on the behaviour and performance of economic agents. They argue that high self-confidence can have at least three types of values: a consumption value (a favourable view of oneself makes a person happier and so enhances her/his utility); a signalling value (self-confidence makes it easier to convince others that you have the qualities you believe you have); and a motivation value (it improves the individual's motivation and morale to persevere with her/his goals and overcome setbacks, thus improving performance). The authors note, however, that in many circumstances over-confidence can also damage performance.

These issues are clarified through a general economic model that seeks to explain why people value their self-image, and how they attempt to enhance or preserve it through various types of seemingly irrational behaviour, from self-handicapping to self-deception through selective memory or awareness management. As the authors show, self-deception in fact serves a number of rational functions. The suggested model of self-deception through endogenous memory reconciles to some extent the motivated ('hot') and rational ('cold') features of human cognition. By opening the door

to a wealth of problems and concepts typical of cognitive psychology, and adding new insights on both the psychological and economic features of human behaviour, the chapter extends the realm of economic analysis.

It does so through the judicious relaxation of a few crucial characteristics of *homo economicus* rationality. A case in point is intertemporal coherence – a crucial requirement of traditional (‘substantive’) economic rationality – which must be relaxed to study the game of strategic communication between an individual’s temporal selves that may explain self-deception. This approach is inconsistent with traditional economic rationality but is consistent with a broader view of rationality that underlies the solution of the game. The insights so obtained could be applied to a wealth of unsettled questions in economics, such as the role of ‘animal spirits’ in investment, and (more or less optimistic) expectations formation in micro- and macroeconomics.

Alessandro Vercelli’s chapter on ‘Rationality, Learning and Complexity’ also focuses on the impact of cognitive psychology on economic behaviour, providing a broad survey of emerging issues. Both casual observation and experimental research suggest that cognitive psychology significantly affects expectations and learning, which in turn play a crucial role in economic decisions. However, standard economics conceives of expectations and learning in a way that makes cognitive psychology irrelevant. The author clarifies the reasons for this neglect and seeks to specify the conditions under which the chasm between economics and cognitive psychology may be bridged. He argues that the crucial obstacle to closing the gap is the narrowness of substantive rationality, with its restrictive notions of expectations formation and learning. The chapter seeks to identify the main assumptions underlying the standard paradigm and to classify deviations from it in coherent alternative paradigms.

Vercelli makes a distinction between ontological complexity with respect to the properties of the economic system, and epistemic complexity with respect to the formal property of the model that represents it. He emphasizes the need to introduce more epistemic complexity, which he sees not as a virtue but as a necessity. In addition, he analyses the implications of different concepts of rationality in defining the theoretical and empirical scope of economic models. He claims that substantive rationality forbids any consideration of the subjective features of economic agents. Rather the concept applies only to an optimizing equilibrium characterized by stability, certainty or ‘soft’ uncertainty, and perfect reversibility of time in a closed and stationary ‘world’. In particular the author shows that the standard assumption of rational expectations is inescapable in substantive-rationality models under uncertainty, but suffers from the same limitations. To take into account the complexity of the real economic world that is often characterized by disequilibrium dynamics, multiple equilibria, structural and dynamic instability and non-stationarity, it is necessary to introduce a more encompassing

notion of rationality. Such a notion would be one that enables, among other conceptual shifts, a more comprehensive hypothesis of expectations formation and takes into account the crucial role played by the cognitive and motivational features of economic agents. Only in this way is it possible to analyse the psychological determinants of economic behaviour, and make possible cross-fertilization between economics, psychology and the cognitive sciences.

Several of the chapters that follow, focus on a different psychological trait – altruism – and its importance in understanding economic behaviour. In their chapter, Stark, Y.Q. Wang, and Y. Wang explore both where altruism comes from and what its repercussions might be. This chapter is constituted of two papers that have emerged from the same research project, and which are presented here as Parts 1 and 2 of a single piece. Part 1 by Stark and Y.Q. Wang – ‘On the Evolutionary Edge of Altruism’ – focuses on how altruism evolves. The authors use the family, in particular siblings, as their starting point, arguing that the emergence of altruism within families can be seen as an important step for explaining the emergence and spread of altruism in society at large. They note that it is more likely for altruism to pervade large groupings if it evolves between siblings than if it fails to establish itself even within families. Accordingly, the authors discuss the evolutionary foundations of the emergence of altruism between siblings based on the ‘Hamilton rule’ suggested by evolutionary biology. According to this rule, altruism is likely to spread in a population if the benefit obtained from giving, times the coefficient of relationship, exceeds the cost of giving. This maximizes the replication opportunities of common genes, since the coefficient of relationship measures the probability of the genes being the same. Within a family, altruism would thus evolve if the benefit to one sibling from receiving help exceeds twice the cost of providing help borne by the other sibling, given that the coefficient of the relationship between two siblings is one half.

Part 2 of this chapter by Stark and Y. Wang – ‘The Intergenerational Overlap and Human Capital Formation’ – already assumes that altruism exists (rather than explaining why it evolves), and the focus is on the economic consequences of altruism, a crucial repercussion being the level of human-capital formation. The authors explain the strong positive correlation between the formation of human capital and life expectancy on the basis of parental altruism and the duration of the intergenerational overlap. Since education costs less if it is financed by parents than by market borrowing, the longer altruistic parents live the more will be the children’s human capital investment. An extended overlap entails the formation of a larger quantity of human capital. This also explains the positive correlation between education and health. However, the authors note, this effect is separate from the returns to human capital – a higher life expectancy increases the period over which the returns can be reaped.

Although in this research Stark and Y. Wang primarily trace the repercussion of parental altruism on human-capital formation, the altruistic trait can also have other implications. For instance, it can impinge on intergenerational transfers of income and resources, as well as on intergenerational transfers of the altruistic trait itself or of substitutes for altruism through the 'demonstration effect' – the provision by adults of care and attention to their own parents, aimed at instilling appropriate preferences in their children (Stark, 1995). Of interest, too, is the possible extension of altruistic behaviour beyond the family to the larger society, in whether individuals cooperate or defect when interacting in non-familial groups; or in the formation of long-term time preferences and the determination of optimal consumption (Falk and Stark, 2000). These aspects have a central bearing on contemporary concerns such as environmental preservation (for example, do you use up most of a forest now or save most of it for your children?), and have promoted other fruitful interdisciplinary exchanges, such as between economics and political science.⁶

The next chapter, 'Human Reproduction and Utility Functions' by Vasin, like the work of Stark and Y.Q. Wang (Part 1 of Chapter 4), also draws upon evolutionary biology, but for a different purpose. He criticizes the standard assumptions of utility functions in game theory, in particular the *homo economicus* model. Among the questions he asks are: why are people willing to work for lower wages than they can earn elsewhere? Or, putting it differently, why do people deviate from what we would expect under standard economic assumptions, namely maximizing individual economic payoffs? He explains this in terms of non-economic motivations, such as people finding their current jobs more interesting or more useful to society, or feeling that their relationships with colleagues substitute for the family.

Vasin also examines whether it is possible to endogenize utility functions and identify how they evolve. Evolutionary game theory indicates that evolutionary stable strategies in self-reproducing populations maximize the fitness of individuals and, as a result, also the individual reproduction rate. Like Stark and Y.Q. Wang, Vasin focuses on altruistic and cooperative behaviour between relatives. He shows that such behaviour is evolutionarily stable if it maximizes the total fitness of the family. However, he points out that in both human and non-human populations there are factors that limit the prevalence of altruistic and cooperative behaviour. In particular this behaviour is not protected from the invasion of selfish agents.

In any case, in his opinion, demographic data show that modern human populations maximize neither individual, nor family, nor population fitness. In fact, he suggests that the typical payoff functions of individuals are based on auxiliary utility functions (affected by feelings of pleasure and by consumerism) that maximize the fitness of 'superindividuals' (corporations,

organizations and institutions) who use a given human population as a resource for reproduction, and, to this end, manipulate people's utility functions.

In the next two chapters – by Englmaier and Chillemi respectively – the impact of altruism and social preferences is examined from a somewhat different, but related, angle, and in a different setting – that of labour markets and industrial relations. A central question facing economists has been how to design the right incentive to ensure that workers perform as the Principal desires. Englmaier in his chapter, 'Moral Hazard, Contracts and Social Preferences' provides a survey of recent contributions in the emerging field of behavioural contract theory that try to incorporate social preferences into the analysis of optimal contracts in situations of moral hazard.

Real-world contracts seldom follow economists' theoretical predictions that are based on the assumption that the agent is entirely self-interested. This view misses out on important factors which affect people's workplace choices and the contracts they sign – in particular their social preferences. Social ties in the workplace, altruistic relations with co-workers, team spirit and work morale, ideas about fairness (which play out differently in relation to fellow workers as versus employers), all matter. For instance, patterns of reciprocity and notions of fairness play an important role in human interactions and especially in labour markets, where people work closely together. People might care not only about their own payoff but also how payoffs are distributed amongst their fellow workers. They may have a social preference for equality among co-workers and might rather forego profits than accept inequitable distributions, since inequality causes them disutility (that is, they may have inequality aversion).⁷ The preferences of agents can also exhibit inequality aversion when they compare themselves to the Principal. Agents may suffer a utility loss if they fail to get their fair share of the output, if the allocation is seen as being 'inequitable'. Taking such social preferences into account can explain behaviour that would appear irrational within the standard economic framework.

Basically, the survey shows how incorporating social preferences in economic models can enhance our understanding of relationships in the industrial workplace and add valuable insights to the analysis of incentive provisions. It also shows how these social preferences can be modelled. These aspects are examined both for the standard one-agent-one-principal problem *and* for multiple-agent settings and team production problems. In these models, a utility function is specified such that a separable term is added to standard utility derived from one's own income, to capture the disutility experienced when others get unequal incomes (that is, to capture relative income comparisons). Recently, experimental and field evidence has also helped amend standard utility functions to take account of social preferences.

Overall the survey shows that social preferences interact in non-trivial ways with incentives and alter the structure of optimal compensation schemes, sometimes drastically. So far the results are inconclusive with regard to the question: under what circumstances is a fair-minded workforce desirable (from the employer's viewpoint). The insights gained from choosing an optimal structure of incentives are still ambiguous, even in the settings of the tournament and team models worked out in this literature. According to Englmaier, so far the main advantage of these new research contributions is in their opening the door to a fruitful dialogue with researchers in the field of human-resource management. This, he argues, can provide a promising avenue for future research. Related issues that he also highlights as worthy of further investigation are the implications of social preferences for structuring work teams, the production process and the information environment.

Chillemi's chapter – 'Mutual Concern, Workplace Relationships and Pay Scales' – also focuses on the impact of altruism in the work place. He examines how altruism among co-workers affects the performance of effort-enhancing incentive schemes and the firm's profits, based on Rotemberg's notion of *trusty altruism*. He outlines Rotemberg's investigation into whether friendly relations in the workplace can induce altruistic feelings among co-workers, thus helping to solve the free-rider problem in team production. Each worker chooses his degree of altruism with the intent of maximizing his own material surplus. Chillemi notes that crucial to Rotemberg's results is good fellowship that allows each worker to recognize the true attitude of his fellow workers, and also makes commitment possible.

Drawing on Rotemberg's work, Chillemi outlines a model to explain the fact that firms rarely adopt pay schemes based on worker competition. This model is then used to characterize the most profitable incentive scheme for maximizing the workers' efforts. Chillemi finds that under reasonable circumstances the firm's surplus increases with worker altruism. An interesting issue discussed in this chapter is whether altruism is consistent with (substantive) rationality. More specifically do workers choose their altruism parameters in order to maximize their material surplus? When workers choose these parameters cooperatively a strictly positive level of altruism emerges in equilibrium. In the case of endogenous altruism, becoming altruistic does not appear to be the best choice, but the scheme of incentives can be modified so that both the Principal and the workers gain.

The six chapters in the first part of this volume thus interweave psychology and economics theoretically, to challenge many of the assumptions and formulations of standard economics. The four chapters which follow in the second part of the volume carry forward this interdisciplinary exchange between psychology, economics and the cognitive sciences by using laboratory and field experiments, again to broaden our notions of rationality, and to take account of altruism as an important constituent of human motivation.

Part II Laboratory and field experiments

Fehr and Tyran in their chapter on ‘Expectations and the Effects of Money Illusion’, examine the nexus between cognitive psychology and rationality also discussed by Vercelli in Part I. In particular, they focus on the effects of money illusion. The authors argue that until recently money illusion was anathema to macroeconomists who tended to dismiss the ‘psychological’ explanation especially for two reasons: first because it contradicted the basic assumption of (narrow) rationality in economics (the argument being that rational human beings do not exhibit illusions, and if by assumption people behave rationally, there is little to study!). On this ground, money illusion stands rejected *a priori*. Second, macroeconomists rejected money illusion on the grounds that it was neither relevant nor backed by convincing evidence. It was seen as irrelevant on the argument that those suffering from such an illusion would lose economically, and this would provide a strong incentive to take illusion-free decisions. Fehr and Tyran emphasize that this argument is seriously flawed since it neglects the indirect effects of money illusion in shaping expectations, even if the individual-level effects are small and transitory.

The authors design experiments to investigate whether money illusion causes nominal inertia. Their results show that money illusion can have massive aggregate effects under conditions of strategic complementarity. Two types of aggregate effects are demonstrated. First, the authors show that money illusion is the cause of nominal inertia after an anticipated monetary shock in an economy with a unique equilibrium. Second, money illusion can have permanent effects by coordinating individuals on inferior equilibria. The use of the experimental method also makes it possible to precisely identify the conditions under which rational expectation models are correct, and the conditions under which they fail to capture important economic facts and forces.

The results obtained by Fehr and Tyran are similar to those highlighted by Vercelli in his chapter through a different approach. When the environment is sufficiently complex, characterized by a multiplicity of equilibria and strategic complementarity, money illusion (understood as a deviation from the rational-expectations equilibrium) may have large and permanent effects, despite limited individual-level deviations. The findings from experiments thus corroborate those obtained via cognitive psychology through questionnaires, namely that money illusion has a ubiquitous framing effect since the nominal representation of economic processes is often simpler and more salient. A fully rational agent expects a certain degree of money illusion from at least some of the other agents, and this is enough to produce macroeconomic inertia, even if individual money illusion is minute or non-existent.

Laboratory experiments also provide insights in Kritikos and Bolle’s chapter on ‘Utility-Based Altruism’. The chapter focuses on what we had noted to be

one of the central themes of this volume, namely the existence and nature of altruism and its impact on economic behaviour. The authors argue that one of the most prominent experiments used by economists for testing the existence of altruism is the dictator game. Based on this game, Kritikos and Bolle seek to provide evidence on economic approaches to the study of altruism, where the recipient's information status is variable. The standard dictator game serves as their benchmark to which they compare a modified game, where the recipient has incomplete information about the size of the pie.

The authors confirm the relevance of altruism in understanding economic behaviour, but note that its features depend on environmental conditions and subjective framing. In particular they explore the influence of cognitive psychology on the existence and degree of altruism. To this end the experiments compare the willingness of 'dictators' to make more or less altruistic offers to an anonymous recipient whose information status varies. The experiments show that it matters what degree of information the 'dictator' attributes to the recipient. This cannot be explained by the usual income-based approach favoured in economics, but can be explained easily by the utility-based approach favoured in psychology (and first applied to this problem in economics by Gary Becker), provided that a fairness component is built into the utility function.

Jungeilges and Theisen in their chapter, 'Equity Judgements Elicited through Experiments', also generate data from experiments in order to examine whether individual decisions are consistent with the Rawlsian second principle of justice.⁸ According to the utilitarian school, welfare judgments should be based on how policies affect the sum of individual utilities, but according to the Rawlsian school welfare judgments should be based on how policies affect the utility of the worst-off individual in society.

Earlier empirical research has indicated that actual choices are determined by a mix of ethical and selfish considerations that are context-dependent, since they are affected by the constraints under which choices are made, and by strategic factors. In order to recover the underlying ethical principles from observed choices, Jungeilges and Theisen use experiments for eliciting the principles that guide individuals, when prioritizing on behalf of society. Specifically they test whether or not individuals make decisions in accordance with Rawls' second principle. They do so by asking students their responses to six different contexts of choice, each of which has a distributional consequence. The subjects are also asked about their demographic characteristics, parental employment background, and so on, and are selected from among business administration students at two stages of their education.

The authors apply complex statistical and econometric techniques to the data generated by their experiments in order to extract the maximum possible information in the most reliable way. Their results show that the support for Rawlsianism declines with changes in a single factor in each choice situation. Utilitarian logic could be used to explain such a decline, but this is by

no means the only possible rationale. Equity judgments typically influence decisions according to the background of the decision maker. In the binary response models worked out by the authors, gender, parental background and education have a statistically significant effect. They find that women's behaviour is closer to Rawlsian principles than men's behaviour, and that subjects who list their parental background as self-employment tend to display a more selfish attitude than those with other parental backgrounds. However, the effect of different educational levels, while found to be relevant, is unclear and needs to be explored further.

The last chapter in this volume – 'Groups, Commons and Regulations' – by Juan Camilo Cardenas, uses both laboratory and field experiments to examine the impact of regulation, in a situation characterized by a negative externality due to the excessive exploitation of a natural resource. Most experimental studies are done in classrooms and in developed-country contexts. Cardenas departs substantially from this in conducting comparative experiments both in the field and in the classroom, in 10 different sites across rural Columbia. The field experiments involve villagers who have joint access to the natural resource. The focus of Cardenas' experimental research is also on the less-studied coordination problem, namely the management of common pool resources, critical for shedding light on questions of environmental sustainability.

The chapter explores the choices individuals make when an external regulation is introduced to solve the coordination problem. It examines how individuals vote when asked their preference regarding the application of such regulations by an external regulator, such as the state. The results suggest that even if a majority of players vote against the externally-imposed regulations, they are still willing to cooperate and reduce over-extraction. However, the players do not respond substantially to changes in the penalty size. Indeed players seem to cooperate 'too much' under a low penalty, and free-ride 'too much' under a high penalty. The results confirm that neither students nor villagers take decisions according to the canonical *homo economicus* model, nor achieve the socially optimum condition. These deviations from standard economic theory may be accounted for only by considering psychological and cultural factors that the author has sought to elicit through personal interviews.

A comparison between the responses of students and villagers is also of interest. Cardenas finds similarities on some counts and differences on others. On the differences, for instance, the villagers are found to be more opposed to external regulatory interventions than the students and to reject such regulations more often than the students. But the villagers are more inclined to cooperate under a non-binding setting. If the experiments had been conducted with students alone, without replication in the field with subjects who have practical familiarity with the problem, the research would have missed relevant information and provided fewer insights. The behavioural

differences found between college students and villagers would not only be of general interest to those using experimental methods, but would also caution against light-heartedly extending experimental results obtained with students, using the standard experimental economics method, to the population at large, or to different sections of it.

* * *

In sum, the chapters in this book contribute in varied ways to an important and innovative stream of research in economics, and we hope our readers will find them stimulating. As the chapters show, economics can considerably extend its theoretical and empirical scope by incorporating insights from psychology and other disciplines and challenging standard economic assumptions. And therein is likely to lie both the continued relevance of economics and its long-term evolution.

Finally, we thank all the authors of this volume for their rich contributions and for their patient receptivity to our suggestions for revision. And we owe a very special thanks to Michael Kaser, IEA General Editor, who has been the hidden third in our selection of chapters, in our pursuit of authors, and in keeping us on schedule.

Notes

- 1 See also Rabin's (1998) review of psychological findings that are of particular relevance to economics.
- 2 On the potential effect of emotions on economic behaviour, see especially Elster (1998), who also emphasizes the dearth of work on how emotions actually influence behaviour.
- 3 The term 'money illusion' has been used by scholars in different ways, but broadly it relates to the tendency to think in terms of nominal rather than real monetary values.
- 4 See e.g. the review by Loewenstein (1992); see also Thaler (2000).
- 5 Interested readers might also see Rilling *et al.* (2002) who use the iterated prisoner's dilemma game in an experiment to investigate the neurobiological basis of cooperative social behaviour, such as reciprocal altruism.
- 6 See especially, Ostrom *et al.* (1994), and the survey in Baland and Platteau (1996).
- 7 On this, see also, Englmaier and Wambach (2002). Among other things, they show that inequity aversion among agents gives a plausible explanation for the predominance of linear wage schemes in real labour markets.
- 8 Rawls (1997: 302) spells out his second principle as below:

'Social and economic inequalities are to be arranged so that they are both: (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and (b) attached to offices and positions open to all under conditions of fair equality of opportunity.'

Jungeilges and Theisen's chapter focuses on (a).

References

- Baland, J.-M. and J.F. Platteau (1996) *Halting Degradation of Natural Resources: Is there a Role for Rural Communities* (Oxford and New York: Oxford University Press).
- Dawkins, R. (1990) *The Selfish Gene*, 2nd edn (Oxford: Oxford University Press).
- Elster, J. (1998) 'Emotions and Economic Theory', *Journal of Economic Literature*, vol. 36(1), pp. 47–74.
- Englmaier, F. and A. Wambach (2002) 'Contracts and Inequity Aversion', CESIFO Working Paper no. 809, University of Munich.
- Falk, I and O. Stark (2001) 'Dynasties and Destiny: On the Roles of Altruism and Impatience in the Evolution of Consumption and Bequests', *Economica*, vol. 68, pp. 505–18.
- Glimcher, P.W. (2003) *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics* (Cambridge, MA: MIT Press).
- Loewenstein, G. (1992) 'The Fall and Rise of Psychological Explanations in the Economics of Intertemporal Choice', in G. Loewenstein, and J. Elster (eds), *Choice over Time* (New York: Russell Sage Foundation), pp. 3–34.
- Ostrom, E., J. Walker and R. Gardner (1994) *Rules, Games and Common-Pool Resources* (Ann Arbor MI: University of Michigan Press).
- Rabin, M. (1998) 'Psychology and Economics', *Journal of Economic Literature*, vol. 36(1), pp. 11–46.
- Rawls, J. (1997) *A Theory of Justice* (Cambridge, MA: The Belknap Press of Harvard University Press), twenty-second printing.
- Rilling, J.K., D.A. Gutman, T.R. Zeh, G. Pagnoni, G.S. Berns and C.D. Kitts (2002) 'A Neural Basis for Social Cooperation', *Neuron*, vol. 35, pp. 395–405.
- Sen, A.K. (1977) 'Rational Fools: A Critique of the Behavioural Foundations of Economic Theory', *Philosophy and Public Affairs*, vol. 6, pp. 317–44.
- Sen, A.K. (2002) *Rationality and Freedom* (Cambridge MA: Harvard University Press).
- Simon, H.A. (1982) 'From Substantive to Procedural Rationality', in H.A. Simon, *Models of Bounded Rationality* (Cambridge, MA: MIT Press).
- Smith, A. [1759] (1966) *The Theory of Moral Sentiments*, Reprints of Economic Classics (New York: Augustus M. Kelley).
- Stark, O. (1995) *Altruism and Beyond: An Economic Analysis of Transfers and Exchanges Within Families and Groups* (Cambridge: Cambridge University Press).
- Thaler, R. (2000) 'From Homo Economicus to Homo Sapiens', *Journal of Economic Perspectives*, vol. 14(1), pp. 133–41.

