

PHILOSOPHY AND ECONOMIC THEORY

EDITED BY

FRANK HAHN AND
MARTIN HOLLIS

OXFORD READINGS IN PHILOSOPHY



PHILOSOPHY AND ECONOMIC THEORY

Edited by
FRANK HAHN AND MARTIN HOLLIS

OXFORD UNIVERSITY PRESS
1979

Oxford University Press, Walton Street, Oxford OX2 6DP

OXFORD LONDON GLASGOW
NEW YORK TORONTO MELBOURNE WELLINGTON
KUALA LUMPUR SINGAPORE JAKARTA HONG KONG TOKYO
DELHI BOMBAY CALCUTTA MADRAS KARACHI
NAIROBI DAR ES SALAAM CAPE TOWN

Published in the United States by Oxford University Press, New York

© *Oxford University Press* 1979

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of Oxford University Press

This book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser.

British Library Cataloguing in Publication Data

Philosophy and economic theory.—(Oxford readings in philosophy).

1. Economics—Addresses, essays, lectures

2. Philosophy—Addresses, essays, lectures

I. Hahn, Frank II. Hollis, Martin III. Series

330.1 HB72 79-41054

ISBN 0-19-875042-0

*Printed in Great Britain by
Fakenham Press Limited, London and Fakenham*

ACKNOWLEDGEMENTS

- I Friedman From *Essays in Positive Economics* by Milton Friedman (University of Chicago Press, 1953) part I, sections 1, 2, 3, and 6 (Conclusion). Copyright 1953 by The University of Chicago. All rights reserved. Reprinted by permission of the author and the publishers.
- II Robbins From *An Essay on the Nature and Significance of Economic Science* by L. Robbins, second edition (Macmillan, 1935) Chapter IV, pp. 72-9, 83-4, 86-9, 90-8, 100. Reprinted by permission of Macmillan, London and Basingstoke.
- III Hollis & Nell From *Rational Economic Man* by M. Hollis and E. J. Nell (Cambridge University Press, 1975) appendix to Chapter 7. Reprinted by permission of the publishers.
- IV von Mises From *Epistemological Problems of Economics* by L. von Mises translated by G. Reisman (D. Van Nostrand Co. Ltd., 1960) Chapter I, Sections 5 & 6. © 1960 by the William Volker Fund in the series Humane Studies. Reprinted by permission of the Institute for Humane Studies who will be publishing a new edition in 1979.
- V Simon From S. Latsis (ed.) *Method and Appraisal in Economics* (Cambridge University Press, 1976). Reprinted by permission of the publishers.
- VI Sen From H. Harris (ed.) *Scientific Models and Man. The Herbert Spencer Lectures 1976* (Oxford University Press, 1979). © Oxford University Press, 1979. Reprinted by permission of the publishers and the Herbert Spencer Fund.

- VII Arrow From P. Laslett and W. G. Runciman (eds.) *Philosophy, Politics and Society*, Third Series (Basil Blackwell & Mott, 1967). Reprinted by permission of the author.
- VIII Sen From *The Journal of Political Economy*, Vol. 78, 1970, pp. 152–7. © 1970 by The University of Chicago. All rights reserved. Reprinted by permission of the author and The University of Chicago Press.
- IX Varian From *Philosophy and Public Affairs*, Vol. 4, no. 3 (Spring 1975), pp. 223–47. Reprinted by permission of Princeton University Press.
- X Hammond From *Econometrica*, Vol. 44, 1976, pp. 793–800 (omitting technical appendix). Reprinted by permission of The Econometric Society (U.S.A.), and the author.
- XI Rawls From *A Theory of Justice* by John Rawls (Cambridge, Mass.: The Belknap Press of Harvard University Press and Oxford: The Clarendon Press). Copyright © 1971 by the President and Fellows of Harvard College. Section 41, pp. 258–65. Reprinted by permission of the publishers.

CONTENTS

INTRODUCTION	1
I. THE METHODOLOGY OF POSITIVE ECONOMICS by Milton Friedman	18
II. THE NATURE OF ECONOMIC GENERALIZATIONS by Lionel Robbins	36
III. TWO ECONOMISTS by Martin Hollis and Edward J. Nell	47
IV. THE SCIENCE OF HUMAN ACTION by Ludwig von Mises	57
V. FROM SUBSTANTIVE TO PROCEDURAL RATIONALITY by Herbert A. Simon	65
VI. RATIONAL FOOLS by Amartya K. Sen	87
VII. VALUES AND COLLECTIVE DECISION-MAKING by Kenneth J. Arrow	110
VIII. THE IMPOSSIBILITY OF A PARETIAN LIBERAL by Amartya K. Sen	127
IX. DISTRIBUTIVE JUSTICE, WELFARE ECONOMICS, AND THE THEORY OF FAIRNESS by Hal R. Varian	134
X. EQUITY, ARROW'S CONDITIONS, AND RAWLS' DIFFERENCE PRINCIPLE by Peter J. Hammond	155
XI. THE CONCEPT OF JUSTICE IN POLITICAL ECONOMY by John Rawls	164
NOTES ON THE CONTRIBUTORS	170
BIBLIOGRAPHY	171
INDEX OF NAMES	175

INTRODUCTION

WHEN English-speaking philosophers think of economics, they usually have a particular kind of pure theory in mind. This is the class of theories predominantly taught in western universities and often called neo-Classical. Purity here is a matter of conceiving *homo economicus* in abstraction from his social setting and, more excusably, of forswearing the attempt to make economics part (or all) of a general theory of society. By contrast, political economy, as the term is now used, is just such an attempt and its champions insist that no economic theory can be as pure as neo-Classicals pretend. Marxian economics is a leading example but has no monopoly and there are heretics within the temple too. The point is less one about the fact-value distinction, although purity is indeed hard to maintain in welfare economics, than one about the proper uses of abstraction. So it raises issues of method and metaphysics, which will be stressed as we proceed. But let us start in innocence, with the claim of pure theory to state the general principles of an autonomous science of economics.

The most effective and subtle of the social sciences offers a fine array of topics for philosophers. Yet there is oddly little commerce at present. There used to be more, when economics was turning itself into a closed, technical discipline with the aid of Logical Positivism and against the wishes of those who valued its leadership of the moral sciences. Since then economists, absorbed in technical issues, have for the most part written little to surprise philosophers and philosophers have tackled economics rarely, nervously, and by way of example only. On the economists' side, Popper's picture of science as conjectures and refutations has been widely admired and endorsed but not seen as a threat to the hopes of a positive economics grounded on orthodox Positivist tenets. The upheaval caused by Quine's pragmatism, Kuhn's paradigms, and other more recent *bouleversements* has yet to send more than a tremor through the temple.

The topic most freely discussed is that of causal laws. Debate is within a broad assumption that progress comes by testing hypotheses against experience. In bald summary of the line handed out to beginners, a natural

law is a regularity in nature holding in specifiable conditions; we have detected one when we have a well enough confirmed theory; a theory is a set of logically-linked, high-order generalizations; the only test of a theory is the success of its predictions; prediction and explanation are two sides of the same and only coin, in that explaining a fact is finding another from which it could have been predicted. Also, in keeping with this Positivist perspective, sciences are thought of as differing in subject-matter, not in method of validation, and there is a thorough distinction of 'is' from 'ought' (positive from normative). Admittedly so clear and simple a rubric is only for beginners and sits uneasily even in the introductory chapter to economics textbooks, whose later chapters are methodologically more involved and pragmatic. But economists are wont to blame their unsolved problems on the incompleteness of economic theory rather than on a philosophic virus and the search for empirical laws goes on.

One fierce, plainly philosophical argument is flourishing, however, and has long enlivened the pages of the *American Economic Review*. It was touched off by Milton Friedman's essay 'The Methodology of Positive Economics' and has to do with whether a good scientific theory needs realistic assumptions. Off hand one might suppose that theories of imperfect competition apply more widely and so are a clear advance on the theory of perfect competition as a guide to economic behaviour. To maintain the contrary, Friedman opens with a pithy and definitive statement of Positivism in economics and tries to use it to show that realism cannot be an independent criterion. He carves any sound economic theory into 'a language' and 'a body of substantive hypotheses'. The former is, he says, 'a set of tautologies' and 'its function is to act as a filing system'; the latter is 'designed to abstract essential features of a complex reality'. Since the only test of a theory is the success of its predictions, the theory of perfect competition enters the lists on equal terms and indeed proves the most useful. His critics have found it hard to escape the elenchus, since they too accept the picture of science, however much they wish to hymn the need for realism.

There is nothing peculiar to economics about this issue, when it is raised in a Positive setting, although the dispute about perfect competition is a notable example of it. But it also crops up elsewhere in economic theory, this time in a distinctive way. Pure theory is deeply committed to an assumption that economic behaviour is rational—not, one might think, a self-evidently realistic postulate to start from. Yet other social sciences are trying to borrow the assumption, together with its economic trappings, and it seems to us the current area of most philosophical interest. In what follows we shall introduce the economic notion of rationality, show how philosophically charged and perplexing it is, note its bearing on social choice and social

justice, and indicate its relevance beyond economics. Since the philosophical topics in Positive economics will be aired in a more general manner, we shall not be neglecting our editorial duties by concentrating on this central assumption of pure economic theory.

In macro- no less than in micro-economics pure theory rests on propositions about individual action. To generalize, it needs to take the individual agent as typical or representative of others. Generality is achieved by conceiving the individual as a rational economic man and asserting that mankind at large is as rational as he. For instance the analysis of investment, savings and liquidity preference which occupies most of Keynes's *General Theory* takes just this form. Before low wages can be explained through the self-interest of capitalists as a group in keeping them low, it must be shown that each capitalist is acting rationally and in a way which, when aggregated, produces this effect. Before trades unions or shareholders can be treated as agencies, they are analysed as coalitions of individuals each with an interest in subscribing to rules which bind them together and in acting jointly. Throughout pure theory macro-movements are thus explained as the collective work of rational individuals and the pedagogic reasons why the pure economist starts with micro-analysis also reflect his deepest ontological commitments.

It is all too easy to suppose that pure theory equates rationality with self-interested action and hence applies only to rational egoists. There is every excuse. Edgeworth roundly declared, 'the first principle of economics is that every agent is actuated only by self-interest'¹ and countless economic models have been built on this principle. The stock distinction between economic and sociological theories of social action has been one between men as rational egoists versus men obedient to norms. Nevertheless egoism is only a special interpretation of the pure notion of rationality. To bring the point out and, at the same time, to guard against other ways in which the theory can easily be misunderstood, we have thought it best to give a technical (albeit incomplete) sketch. Despite the use of symbols, the careful novice should be able to follow.

The pure theory characterizes the agent independently of his environment. It is assumed that he has preferences and the characterization is by the kind of preferences he has. The domain of his preferences is the set of consequences of his possible actions. Let C , with elements c, c', c'' , etc., be the domain of a given agent's preferences and let R be a relation among the elements of C , so that cRc' is read as ' c is at least as good as c' '. The theory assigns three properties to R :

¹ F. Y. Edgeworth, *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences* (Kegan Paul, London, 1881), p.16.

- (1) *Completeness*. For all pairs (c, c') in C , either cRc' or $c'Rc$ or both. (When cRc' and not $c'Rc$, economists sometimes write cPc' (i.e. c is preferred to c'); when cRc' and $c'Rc$, they sometimes write cIc' (i.e. c is indifferent to c').)
- (2) *Reflexivity*. For all c in C , cRc .
- (3) *Transitivity*. For every triple (c, c', c'') in C , if cRc' and $c'Rc''$ then cRc'' .

(We stress that this is not a complete account but it will serve.)

The interpretation of the theory depends on the domain specified. Where C is taken as the vector space an element of which gives the bundle of goods consumed by the agent himself, it is usual (although not obligatory) to regard this as postulating purely selfish preferences. Where C is taken as the vector space an element of which gives the consumption bundle of every agent in the economy, the matter is more open and the agent's preferences on C may or may not be selfish. If his ordering R can be fully specified independently of the consumption by other agents, we might want to call him *prima facie* selfish; if not, then his preferences may be selfless or they may be envious. Hence cPc' can mean ' c is more in the self-interest of the agent than c' ' but it need not. The postulate that an agent is characterized by preferences rules out neither the saint nor Genghis Khan.

In an abstract way we may think of an agent as having available a set of mutually exclusive actions. The choice of one of these will have consequences, which are what interests the agent. In general a consequence need not be uniquely defined by an action but, whenever it is, a preference ordering over actions can be regarded as one over the consequences of actions and, for ease of exposition, we shall assume unique consequences of action.

The pure economist's definition of rational choice is now this: Given the set of available actions, the agent chooses rationally if there is no other action available to him the consequence of which he prefers to that of the chosen action. On the face of it, it looks very simple—even plausible; yet it has striking implications, as we shall next show by examples.

So let us consider an economy with only two goods—a restriction made for purposes of exposition only. Take an agent who owns some of both goods before the story starts and can trade at given prices. His action-set consists of all trades which he can make at these prices, given that (a) he can sell no more of any good than he owns and (b) that his receipts from a sale must not be less than the amount required for a purchase. The agent's preferences are defined over all possible consumption bundles of the two goods (all non-negative pairs).

Proposition A: The choice of the rational agent will be independent of the level of the two prices (i.e. will depend only on their ratio).

Proof: Consider two price situations which differ only in that each price in situation one is a positive multiple of the price in situation two. From the definition of the action-set of the agent it is clear that the sets are the same in both situations. But then the choice which was rational in one situation must also be rational in the other. For, if not, then one choice must be preferred to the other; but both are available in both situations.

This simple proposition plays an important role in the theory of inflation.

Proposition B: A rational agent will prefer to pay a given amount to the Government by means of an *ad valorem* tax to paying the same amount by means of a lump sum (poll) tax.

Proof: (a) Suppose the *ad valorem* tax on good one yields £T from the agent, good two being untaxed. Remove the *ad valorem* tax, thus lowering the price of good one, relative to good two, by the tax per unit. Impose a poll-tax of £T, thus making the agent pay £T whatever his choice. Check that the agent can now, if he wishes, buy as much of each good as he did when there was an *ad valorem* tax. Removing the *ad valorem* tax saves him the tax per unit of good one times the amount which he was buying; and that is exactly what is taken away from him by the poll-tax. Hence the *ad valorem* tax choice is in the action-set available under the poll-tax.

(b) If the agent now chooses a different action, it must, by the postulate of rational choice, be true that he prefers it to his *ad valorem* tax choice.

This proposition is closely related to the economist's prescription that price should equal marginal cost and is of wide practical interest.

Proposition C: Suppose there are n agents. Define a *competitive equilibrium* as a price system and a pair of trades, one for each agent, such that (a) at these prices, given preferences and the ownership of goods, the trade of each agent is a rational choice, (b) the total amount of each good supplied is equal to the total amount demanded. Assume that these prices are both positive and assume that each agent in this equilibrium would prefer to have more of any good (if possible) than he has—i.e. agents are not satiated. Then there is no way for a central authority to reallocate the existing stock of goods so as to ensure that each agent obtains an allocation at least as highly preferred as

his equilibrium allocation and at least one agent receives an allocation which he prefers to his equilibrium allocation. The equilibrium is Pareto-efficient.

Proof: Suppose the proposition is false. Consider agent i , who obtains an allocation which he prefers to his equilibrium allocation. Since he was choosing rationally in the equilibrium, the newer allocation was not in his action-set at equilibrium prices. It must then be that at these prices the new allocation would violate the condition that receipts from sales should equal amount spent. (Non-satiation ensures that in equilibrium the agent will not spend less than he receives.) On the other hand all those agents to whom the central authority allocates a bundle which they consider indifferent to their equilibrium bundle will find that at equilibrium prices the centrally allocated bundle either violates or just meets the condition that receipts equal outgoings. For, if receipts were less than outgoings, when the central allocation is indifferent to the equilibrium one, the agent was not rational in his choice of equilibrium trades. That is because he could have traded to get a bundle as good as his equilibrium one and still have something left over to buy more or sell less. But non-satiation then leads to the result that the equilibrium choice was not rational—a contradiction.

So now we have shown that to obtain the centrally allocated bundle at equilibrium prices each agent would spend as much as or more than he receives and at least one agent would have to spend more than he receives. So if agent i is allocated the consumption (c_1^i, c_2^i) and owns (x_1^i, x_2^i) , we have, when (p_1, p_2) are equilibrium prices:

$$p_1(c_1^i - x_1^i) + p_2(c_2^i - x_2^i) \geq 0, \quad \text{for all } i.$$

$$p_1(c_1^i - x_1^i) + p_2(c_2^i - x_2^i) > 0, \quad \text{for some } i.$$

Adding over agents and recalling that $p_1 > 0$, $p_2 > 0$, we conclude that the allocation would use more of at least one good than is available:

$$\begin{aligned} \text{i.e.} \quad \sum_i c_j^i &> \sum_i x_j^i \\ &\text{for either } j = 1 \text{ or } j = 2 \text{ or both.} \end{aligned}$$

Hence Proposition C is not false.

This proposition has played a central role in economics and it can be made more significant by its 'dual', which asserts (roughly speaking) that any

Pareto-efficient allocation can, under certain conditions, be decentralized into a competitive equilibrium. But this and the many other applications would take us too far afield. The examples already given suffice to show that the rational-choice postulate is indeed powerful and useful.

To complete our brief account, and because some of the contributors later are concerned with it, we shall now sketch the notion of collective rational action.

Consider an economy of n agents and let there be a number of possible alternatives. Let \mathcal{R} be the set of all possible orderings of the alternatives, where the orderings are complete, reflexive, and transitive. For instance, if there are two alternatives a and b , then either a is preferred to b , or b is preferred to a , or a and b are indifferent. The problem posed by Arrow² is this: does there exist an ordering which is complete, reflexive, and transitive (i.e. in \mathcal{R}) which is derived (a map) from any preference profile $R^1 \cdots R^n$ of the n agents with R^i in \mathcal{R} , for all i ? That is, can we get an ordering of the alternatives which depends on the orderings of the agents?

The answer is of course yes. We need only to make the ordering of any one individual into our social ordering, so that his preferences *are* the social preferences. But that is too easy and also undesirable. So we exclude this solution by imposing a *non-dictatorship* condition.

Next it seems reasonable to require that, if alternative b is higher in the preferences of every agent than is alternative a , then the social order should put b as socially preferred to a . So we write this in by imposing the *Pareto-condition* on the social order which we will accept.

Lastly one imposes the plausible, but none the less problematic, condition that social preferences between b and a should depend only on agents' preferences between b and a and not at all on their preferences between any other pair. This is the condition of *Independence of Irrelevant Alternatives*.

Arrow proved a remarkable theorem: if we allow any preferences in \mathcal{R} (unrestricted domain) and require non-dictatorship, the Pareto-condition, and Independence of Irrelevant Alternatives, then a social preference cannot be constructed from agents' preferences.

A vast literature has grown up on this theorem. Without trying to pass any judgements, we would like four points to be noted. Firstly, the problem has been formulated at a high level of generality. In particular we are asked to find a derivation or 'map' which will work for *any* admissible preference profile ('unrestricted domain'). Thus the agnosticism concerning the origin or social determination of preferences is carried very far. Secondly, in conformity with the theory of rational choice of the agent only orderings of

² K. J. Arrow, *Social Choice and Individual Values* (New York, 1951). See also his article in the present volume.

alternatives are considered. In particular we are not allowed to consider intensities of preferences. If we were allowed to do so, then in deriving a social order we could either take account of them or not. If we do not, nothing is gained. If we do, it implies that we can compare intensities. Thirdly, the Pareto property may not in fact be desirable. It may conflict with the view that there are 'private domains' for each agent over which his ordering should be decisive. Fourthly, the requirement of the independence of irrelevant alternatives rules out considerations of intensity of preferences and is not a self-evidently acceptable axiom. The large literature notwithstanding, however, Arrow's result shows that there are great difficulties in extending the notion of individual rationality found in economic theory to one of collective rationality.³

This concludes the brief exposition and we next consider two obvious objections to the postulate in a summary way. One is to using unexplained preferences as explanans, the other to supposing that agents do or can act rationally on the preferences ascribed. Both objections look formidable but, we think, neither is decisive.

Firstly, then, it seems plain that preferences are determined in part by the very process which they are meant to explain, and not given independently as the theory requires; nor, granted that they are not all of equal permanence, importance, or worth, can they be relied on to yield any social optimum when aggregated.

As far as explanation goes, the pure economist replies that preferences, although mutable, change slowly in relation to the endogenous variables of economic models. Hence it does no harm to the rest of the analysis to treat them as exogenous (and unexplained). That they do change more slowly is a falsifiable hypothesis which is confirmed by experience. For instance much, in our view vulgar, criticism of the theory has turned on the effects of advertising, by which firms are allegedly able to engineer whatever preferences suit them. The pure economist might, of course, retort by disputing the power of advertising to affect tastes directly. But, more subtly, he will probably answer by pointing out that the domain of preferences may very well not be the space of goods but the space of characteristics of goods—the speed and comfort of travel, for example, rather than the motor cars which

³ It is plain that this account leaves much of interest undiscussed. In particular it is not made clear what interpretation to put on agents' preferences. Do they reflect just choice or judgement or consideration of their own welfare? Certainly these are matters for investigation (witness the work of A. K. Sen) but they do not, in our view, much affect the formal apparatus. For it would seem on close examination that it is satisfactory to endow agents with single preferences but to make their domain large enough. For instance the domain includes not only butter but also 'freedom' and so forth, and the agents' preferences give us trade-offs between all the elements.

embody these characteristics in varying proportions. Advertising, he will argue, is more to do with getting us to ascribe these virtues to particular cars than with changing what we ask of a car. By affecting the (perceived) map from actions to consequences, it can change actions without changing tastes. In general to perceive something under a fresh description is not necessarily to alter the criteria for evaluating it against other things and hence it is no blow to the theory to recognize that preferences alter. Whether he is right is an empirical matter but he does have at least some answers to the objection.

This retort is too quick, however, when it comes to aggregating preferences for purposes of Welfare economics. If we want to order social states by the relation ' a is socially no worse than b ', we must know why preferences of agents have a claim on our attention at all and which preferences are to count for how much. Whether or not we can significantly ask if A 's preferences are rational, we can certainly ask if they are good or indeed if they are no worse than B 's.

The general answer is that pure economists are democrats enough to try for a measure of welfare which respects what people actually want. But, believing that some wants are more transitory or antisocial than others, they are also ready to weight preferences in arriving at social preferences. By adopting some moralist's preferences over preferences, they can assign zero or even negative weights in some cases. If this seems to sully the purity of a theory of rational action, it at least permits them to escape Arrow's theorem by introducing intensity of preferences and a suitable notion of comparability.⁴ Also they do not propose their own moral calculus and have been content to axiomatize various moral ideas supplied by others. For instance Rawls's theory of justice as fairness has given rise to much recent work in economics, one result of which has been to convert the criterion of ordering social states by the welfare of the worst-off person in each into a lexicographical maximin on the grounds that, if a and b leave the worst-off equally well-off, while b does better for everyone else, b should be socially preferred to a . Yet the economist thinks of himself in transactions like these not as solving problems but as helping to clarify possible solutions. (We shall ask presently whether he is being disingenuous.)

Where preferences are being weighted by their social worth, economists can try leaving the choice of weights to others. Where weights are used because preferences are transitory they cannot. The fact that the agent's preferences will change with learning, ageing and circumstance should reasonably limit what he can prefer today, if he is not to regret it tomorrow. The point has been studied recently and a new, but not deeply different,

⁴ A. K. Sen, *Collective Choice and Social Welfare* (Oliver and Boyd, 1970).

axiomatization of rationality has been proposed to allow for it.⁵ But more than technicality is involved, since orthodox pure theory cannot admit any serious sense in which preferences may be irrational. Indeed there is difficulty even in the unproblematic thought that it is irrational to hold inconsistent preferences or to be unwarrantably mistaken about the consequences of enacting a preference (for instance that one will not regret doing so). For the theory takes preferences to range over well-defined states (consequences of actions) and it is not obvious where to locate these possibilities.

That brings us to the second blunt objection, which doubts whether agents do or can act rationally on the preferences ascribed. It is all too plain that no one in fact has a complete ordering of relevant alternative states for every, or indeed any, occasion—nor could he have one without spending his life in hypothetical comparisons. Nor will it do to postulate consistency without ascribing to everyone a degree of computational skill which no one in fact possesses. We are thus led back to the question left hanging at the start—whether a theory needs realistic assumptions if it is to have explanatory merit. A possible, if blunt, answer is that pure economic theory works well enough when it is used for rational choice over a more or less small subset of available choices. Certainly there is no denying that modest empirical studies have been successful. But, finding this a craven reply and the issues it shirks interesting for all the social sciences, we urge even the pure to have the courage of its assumptions.

Perhaps, then, it is not irrational to act on incomplete preferences, even when the agent makes a choice which, had he made the effort, he would have discovered to be inferior to another choice available. After all, the exploration of his preferences or environment takes time, trouble, and expense (buying *Which?*, for example) and so is an activity which should itself feature among his preferences. Admittedly this thought threatens to make nonsense of the very idea of complete preferences, since the agent cannot know whether he rationally prefers to rest content with what he knows already unless he knows already what he would discover if he continued searching. But it suggests that it can be rational to act on a consistent subset of one's preferences even if it would be found inconsistent with some larger set. Thus Simon has long advocated an approach through 'bounded rationality', where agents behave as in the theory which we have outlined but only over a subset of alternatives and of their environment.⁶ The subset is determined by the agent's 'aspiration level', which in turn has social and

⁵ B. Peleg and M. Yaari, 'On the Existence of a Consistent Course of Action when Tastes are Changing', *Review of Economic Studies* (1973).

⁶ H. Simon, *Models of Man* (J. Wiley and Sons, 1957).

psychological determinants. It is only when outcomes fall below aspiration levels that the agent will institute search of his environment or of his preferences.

This way out is descriptively plausible but has not so far proved theoretically useful, since the aspiration levels and the search activities are ill-defined.⁷ But it offers a tempting answer to the question whether it is preferences and correct computation which constitute rationality. The answer is, in effect, that there is no 'correct' definition of rationality but that different applications of the theory call for different definitions, each leaving a fringe of actions whose rationality is undecidable.

At any rate we think this a much better line than that of restoring completeness by making the consequence space merely subjective. It may seem promising to say that, even if the agent would have changed his mind, had he discovered more, nevertheless he chose rationally what he perceived to be best at the time. Indeed it is hard to deny that a man can act rationally on false beliefs about consequences, including those of satisfying his present desires. Hence it looks as if the objective notion of rationality which we have used in sketching the theory should be recast in terms of how an agent's world looks and works from within. The change would be more than local. For instance it would no longer follow that a competitive equilibrium is Pareto-efficient relative to the true possibility of the economy; Welfare economics would have to judge social states on the basis not just of preferences but also of varying perceptions and computational skills, with great loss of persuasiveness. But it looks as if it could and indeed should be done. In truth, however, pure economic theory has found it hard to make the change. Pure theory rests on an assumption that R is a fixed relation, present in all economic agents. The typical or representative individual, with whom we began, has to have complete, reflexive, transitive preferences, if the first, essential theorems are to sustain a scheme of explanation. The theory has had to ignore the source of preferences⁸ and to construct everything from their implications. It is not enough that each man does what seems to him required by what he thinks his preferences are in what he takes to be the circumstances. Yet plainly there is an unrealism about the universal assumption of R and, we submit, a notion of 'bounded' or restricted, but still objective, rationality is the best way to face up to it.

That, we agree, does not finally dispose of the doubt that whether the rationality postulate is realistic enough. Nor does it prove that pure theory

⁷ R. Radner, 'Aspirations, Bounded Rationality and Control', Berkeley Discussion Paper, 1975.

⁸ But see C. C. von Weizsäcker, 'Notes on Endogenous Changes in Taste', *Journal of Economic Theory*, iii (1971).