Collective and Individual Rationality:

Some Episodes in the History of Economic Thought

Andrew Martin Paul DENIS

Thesis submitted in fulfilment of the requirements for PhD in Economics

City University, London

Department of Economics
School of Social and Human Sciences

December 2001

This thesis is dedicated to the memory of my father,

Paul Justin Denis (1924-1982).

Contents

Lists of tables and figures

Tables

Figures

Acknowledgements

I should like to thank my advisor, Geoffrey Kay, and my wife, Mary Denis, for their unfailing support and encouragement during the production of this thesis. The inspiration for the thesis originated in work completed in the late 1980s as a student on the excellent MA course in Political Economy at the then Middlesex Polytechnic; I should like to record my debt to the teaching staff on that course, in particular the late Geoffrey Pilling.

I should like to thank the following for their encouragement and/or for insightful criticism of various points in the thesis: Erik Angner, William Barber, Stephan Böhm, John Broome, Mario Bunge, Pete Clarke, Paul Coleshill, John Cowley, William Dixon, Sheila Dow, Denis Glycopantis, Denis Gray, Alfons Grieder, the late Peter Holl, Allan Isaac, Steve Miller, Simon Price, Joan Safran, Joe Sen, Adrian Seville, Ron Smith, Ian Steedman, Richard Sturn, Julian Ullman, Jack Vromen, and Rachael Walker.

Papers based on various parts of the thesis have been presented at the Economics Department and Interdisciplinary seminars in the School of Social and Human Sciences at City University, and conferences of the European Society for the History of Economic Thought (ESHET), the European Economics Association, the Association for Heterodox Economics (AHE), and the Leeds Postgraduate Economics Annual Conference. I should like to thank the many scholars, too numerous to list, who have participated in these discussions.

As is detailed at the appropriate points in the thesis, papers based on the material embodied in it have been submitted to various journals and I should particularly like to thank the editors, and several anonymous referees, for the *Journal of Socio-Economics*, *Constitutional Political Economy*, and *History of the Human Sciences*, for their helpful comments.

A debt of gratitude is owed to my students on the final year undergraduate option in History of Economic Thought at City University, London. Their responses to the material of this thesis, as it developed, have been extremely helpful: I should like to thank them for their enthusiasm, commitment and challenging criticism.

Finally, I should like to thank John Cowley and Jessica Holding for the invaluable gift of books which had belonged to the late Arthur Clegg and John Cameron, respectively.

The views expressed and errors committed here, are, of course, entirely my own and not to be associated with any of the above.

Declaration on consultation and copying

The following statement is included in accordance with the Regulations governing the 'Physical format, binding and retention of theses' (*Research Studies Handbook*, 2000: Appendix ii, (xxx), paragraph 5(e)):

> I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Andy Denis

Abstract

This thesis argues for the fundamental importance of the opposition between holistic and reductionistic world-views in economics. Both reductionism and holism may nevertheless underpin *laissez-faire* policy prescriptions. Scrutiny of the nature of the articulation between micro and macro levels in the writings of economists suggests that invisible hand theories play a key role in reconciling reductionist policy prescriptions with a holistic world.

An examination of the prisoners' dilemma in game theory and Arrow's impossibility theorem in social choice theory sets the scene. The prisoners' dilemma epitomises the collective irrationality coordination problems lead to. The source of the dilemma is identified as the combination of interdependence in content and independence in form of the decision making process. Arrovian impossibility has been perceived as challenging traditional views of the relationship between micro and macro levels in economics. Conservative arguments against the possibility in principle of a social welfare function are criticised here as depending on an illicit dualism.

The thesis then reviews the standpoints of Smith, Hayek and Keynes. For Smith, the social desirability of individual self-seeking activity is ensured by the 'invisible hand' of a god who has moulded us so to behave, that the quantity of happiness in the world is always maximised.

Hayek seeks to re-establish the invisible hand in a secular age, replacing the agency of a deity with an evolutionary mechanism. Hayek's evolutionary theory, criticised here as being based on the exploded notion of group selection, cannot underpin the desirability of spontaneous outcomes.

I conclude by arguing that Keynes shares the holistic approach of Smith and Hayek, but without their reliance on invisible hand mechanisms. If spontaneous processes cannot be relied upon to generate desirable social outcomes then we have to take responsibility for achieving this ourselves by establishing the appropriate institutional framework to eliminate macroeconomic prisoners' dilemmas.

Key to symbols and abbreviations used in the thesis

Symbols

The *delta* symbol (Δ): ΔX means the change in X, where $X = C$ or M.
The *prime* symbol ('): X' means X plus ΔX, where X means C or M.

AD	aggregate demand	*MEC*	marginal efficiency of capital
C	commodity	*MPC*	marginal propensity to consume
i	the rate of interest		
M	money		

Abbreviations of source titles[1]

Astronomy	Smith 'The History of Astronomy', in EPS: 33-129	LLL	Hayek (1982)
		Mandeville	Hayek (1967b)
		NSP	Hayek (1978a)
COL	Hayek (1960)	RTS	Hayek (1944)
EPS	Smith (1980)	SIP	Hayek (1967a)
CRS	Hayek (1979)	Sup	Keynes (1979)
CWXIII	Keynes (1973b)	TBT	Hayek (1978b)
CWXX	Keynes (1981)	Times	Keynes (Keynes, 1937a, b, reprinted in Hutchison, 1977)
CWXXI	Keynes (1982)		
CWXXVII	Keynes (1980)		
EP	Keynes (1972a)	TM	Keynes (1971)
GT	Keynes (1973a)	TMS	Smith (1976/1759)
IEO	Hayek (1948)	TSO	Hayek (1952)
KES	Hayek (1983)	WN	Smith (1976/1776)

Other abbreviations

AI	artificial intelligence	GE	general equilibrium
ESS	evolutionarily stable strategy	PC	predicate calculus
		SWF	social welfare function

[1] For details of works by Keynes, Hayek and Adam Smith, please refer to the bibliographical notes attached to the relevant chapters: Section 4.2 for Adam Smith, and the appendices to Chapters 5 and 6 for Hayek and Keynes respectively.

Chapter 1 Introduction: Holism versus reductionism in economic thought[2]

1.1 Preamble

'What, if any, is the legitimate rôle of the state in the economy?' That is a fundamental question – perhaps *the* fundamental question – for economics. How do (micro level) agent interests and behaviours interact to generate (macro level) social outcomes? Are those outcomes desirable, or should society as a whole, in the form of the state, intervene to modify them? This thesis will investigate these questions and explore the answers that have been given by some characteristic economic thinkers. The thesis thus forms part of an investigation into the views of various writers on the articulation between micro and macro levels in economics, between individual actions and social outcomes, between individual and collective rationality.

The thesis begins, in this chapter, with an introduction to some of the fundamental methodological issues underlying the remainder of the work. A consideration of two problems in twentieth-century political economy – the prisoners' dilemma and Arrow's impossibility theorem – will then establish the currency of these themes in contemporary political economy. Attention will then turn to Adam Smith's 'invisible hand', and the group evolutionary theory of Friedrich Hayek. The question addressed is, What is the mechanism by which these writers supposed that individual (micro) rationality translated into collective (macro) rationality? In conclusion, an alternative twentieth-century response to this issue, that of John Maynard Keynes, will be considered.

This structure is dictated by the following considerations. Firstly, the issues of micro and macro, of disjuncture and emergence, of reductionism and holism, and of a providential

[2] An article based on this chapter is at point of writing under consideration with the Journal of Socio-Economics as Denis (2001a).

or indifferent world, are ones which continually re-emerge in political economy. The two issues with which I start the thesis establish that point by reference to two controversies of the second half of the 20th century. Of these, the prisoners' dilemma turns out to be of fundamental significance for the rest of the thesis. Arrovian impossibility is not directly of the same level of significance, but does play a useful role here, both in illustrating some of the themes of the thesis, and in highlighting just why the prisoners' dilemma, by contrast, is so important.

The bulk of the thesis, the most important two chapters, consider the invisible hand mechanisms of Smith and Hayek. Given the topic of the thesis, the articulation of micro actions and macro outcomes, it was essential to consider the invisible hand, firstly in its original, and secondly in its modern incarnations. In this way we may see what is enduring and what has changed in the presentation of this theme by providentialist political economists. An important finding of these chapters is that both Hayek and Smith reject the reductionist approach adopted, for example, by modern monetarist and new classical writers in favour of a holistic methodological approach. Finally, the chapter on Keynes is required in order to show what happens when a holistic methodological approach is accompanied by an explicit rejection of the invisible hand: in an indifferent, non-providential world we ourselves are obliged to take responsibility for the unintended consequences of our actions.

The present chapter introduces the thesis by means of a consideration of a recent paper by Mario Bunge (Bunge, 2000) 'Systemism: the alternative to individualism and holism'. Firstly, I argue, with Bunge, that schools of thought in economics may be characterised according to their stance on a key methodological opposition: that between holism and reductionism. Secondly, I argue that this choice of standpoint has important consequences for policy prescription. Arguing against a simplistic correlation of reductionism and *laissez faire*, the case is made for two kinds of methodological underpinning for *laissez-faire*: firstly, reductionist and, secondly, holist plus invisible hand mechanism. On the basis of these methodological preliminaries, the chapter concludes by outlining the subsequent structure of the thesis.

1.2 Reductionism and holism: a response to Mario Bunge

This section and the next constitute a response to Mario Bunge (2000) 'Systemism: the alternative to individualism and holism'. Bunge's paper is an extremely interesting article which makes a number of telling points and is evidence of a growing discomfort with the reductionism of the neoclassical school currently hegemonic within the discipline of economics. However, there are two major points which need to be made with respect to the article. Firstly, some translation is required, since the terms 'holism' and 'reductionism' are used differently in his article and in the present thesis. The second point is that the relation between policy prescription and philosophical standpoint in political economy is both more complex and more fascinating than Bunge's account would lead one to believe.

The first of these points – that concerning translation between Bunge's terminology and mine – is considered in this section, which then proceeds to make some initial comments on the significance for political economy of the opposition between reductionism and holism. The next section looks in more detail at the second question, that of the relationship between methodological standpoint and policy prescription in political economy.

Bunge's thesis is that there are three fundamental research approaches in the social sciences: the two most influential, *individualism* and *holism*, being fatally flawed, with only the minority approach of systemism offering a viable way forward. The first two are inadequate – each managing to avoid the other's error only by committing an opposite error of its own, while the third, *systemism*, manages to synthesise the other two, accepting the criticism each makes of the other.

For Bunge, 'holism' is inadequate 'because there are no relations without relata'; and 'individualism' so 'because all individuals are interrelated'(147[3]):

> "Neither of the two most influential approaches to the study and management of social affairs is completely adequate ... individualism is deficient because it underrates or even overlooks the bonds among people, and holism, because it plays down or even enslaves individual action. By contrast, systemism makes room for both agency and structure." (156-157)

[3] Unqualified page numbers in this section and the next refer to Bunge (2000).

Systemism, apart from being defined negatively with respect to individualism and holism, is characterised by viewing the world from a systems perspective: 'everything is either a system or a component of a system and every system has peculiar (emergent) properties that its components lack' (147).

Now, there is nothing in these formulations which is contentious[4], and the points made are valuable ones. However, I think greater clarity can be obtained by stating the matter slightly differently. We are fully justified in endorsing Bunge's rejection of (what he calls) 'holism' as ignoring the fact, that the relations, of which our systems of relations consist, depend fundamentally upon the properties of the *relata*, the substrate-level entities which are doing the relating. What Bunge refers to as 'holism' makes the properties of an entity independent of the properties of its material substrate. This, however, is not the way that the term *holism* will be employed in this thesis.

We are equally justified in complaining, with Bunge, that what he refers to as 'individualism' fails to take into consideration the *relationships* between agents, the fact that individuals are only nodes in *systems* of such relationships, and the *emergence* of properties at the macro, or system level. But, when we do so, we are taking a position on an opposition which has implicitly or explicitly underlain a vast amount of methodological discourse in economics and elsewhere: that between holism and reductionism. In defining 'individualism', which I refer to as *reductionism*, Bunge is implicitly defining a non- or anti-individualist approach (which includes both 'systemism' and 'holism' in his terms), which I refer to as *holism*. Essentially, *reductionism* involves a strategy of interpreting the things we see in the world, and, in particular, economic phenomena, as congeries of substrate-level entities; and *holism* is the attempt to understand these phenomena as a whole or system, with emergent properties not enjoyed by constituent components. The terms *holism* and *reductionism* as used in this thesis may be defined as follows:

Reductionism: the view that an entity at one level can be understood as a congeries, an aggregate of entities at a lower, substrate level, that the properties and behaviour of higher level entities can be

[4] Apart, perhaps, from a certain tendency to hyperbole: that 'everything' is a system, or part of one, is both debatable and an unnecessary claim for the point being made.

understood in terms of the properties and behaviour of its constituent lower level parts, *taken in isolation*.

Holism: the view that phenomena at one level can be understood as emergent at that level, that a higher level entity can be understood as a product of the *interrelationships* between its component parts.

Correspondence relating to previous manifestations of this chapter suggests that some discussion is in order at this point. Firstly, the question has been raised (Bunge, personal communication) as to whether 'reductionism' and 'holism' so defined are ontological or epistemological standpoints. As stated, both definitions are epistemological – 'an entity … can be understood as …'. A logical further step, adopting a realist standpoint, would imply that a phenomenon can be understood in a particular way precisely because that is the way it is. In that vein, as well as being explicitly epistemological, the definitions would be implicitly ontological. However, that further step is unnecessary for present purposes.

Secondly, the exact definitions of the terms are as controversial as everything else in the debate between the supporters of each point of view. The above definitions are my own, and are as unlikely to please as any others[5]. Correspondents, responding to earlier papers on this theme, have criticised this or that statement about reductionism or holism – but, because their criticism was on the basis of other definitions than those set out above, they failed to engage with the points that I am making[6]. I submit that what is critical here is that holism and reductionism so defined can be shown to characterise two living trends in economic thought and to throw up interesting and enlightening questions about the nature of those trends.

[5] An excellent discussion with some initial references can be found in the entry "'Emergence' and 'reduction' in explanations" in Gregory (1987: 217-218). See also the entries on 'reductionism' and 'holism' in Honderich (1995: 750-52, 371-72).

[6] It would, of course, be open to critics to object to my fallible attempts to set out a logically coherent opposition in sympathy with the literature I had consulted.

Keynes, for example, in the 'Preface' to the French edition of 1939 of the *General Theory*, clearly advocates a holistic approach:

> "I have called my theory a *general* theory. I mean by this that I am chiefly concerned with the behaviour of the economic system as a whole .... And I argue that important mistakes have been made through extending to the system as a whole conclusions which have been correctly arrived at in respect of a part of it taken in isolation." (**GT**: xxxii[7])

Keynes sets out very clearly here what he takes to be the distinguishing feature of the two approaches: that, on the one hand, we can derive **correct** conclusions from the study of *micro*economic phenomena 'taken in isolation', but that to extend those conclusions to *macro*economic phenomena leads to error, and, on the other, that the correct approach is (what we would now call) a systems approach, aiming to examine the behaviour of 'the economic system as a whole'. The fallacy of composition involved in the approach he is criticising here, derives from a difference, between the two levels, in what it is legitimate to take as parametric or given (Keynes, 1973: 293).

Robert Lucas, on the contrary, is a very clear spokesman for the trend in economics which favours a reductionist methodology. The following is taken from the final paragraph of his *Models of Business Cycles*:

> "The most interesting recent developments in macroeconomic theory seem to me describable as the reincorporation of aggregative problems such as inflation and the business cycle within the general framework of 'microeconomic' theory. If these developments succeed, the term 'macro-economic' will simply disappear from use and the modifier 'micro' will become superfluous. We will simply speak, as did Smith, Ricardo, Marshall and Walras, of *economic* theory. If we are honest, we will have to face the fact that at any given time there will be phenomena that are well-understood from the point of view of the economic theory we have, and other phenomena that are not. We will be tempted, I am sure, to relieve the discomfort induced by discrepancies between theory and facts by saying that the ill-understood facts are the province of some other, different kind of economic theory. Keynesian 'macroeconomics' was, I think, a surrender (under great duress) to this temptation. It led to the abandonment ... of the use of the only 'engine for the discovery of truth' that we have in economics" (Lucas, 1987: 107).

Here we have a clear expression of the desire to reduce macroeconomics to microeconomics, and a characterisation of the Keynesian approach as an illegitimate 'surrender to temptation'.[8]

---

[7] Throughout the present thesis, emphasis in passages cited is as in the source unless its addition is explicitly noted.

[8] Simon Price (personal communication) has argued vigorously against labelling Lucas a reductionist.

Both the passages cited occur in contexts – a preface, and the concluding paragraph of a book – where the authors are standing back from the detail of the theories that they are presenting, and indicating what they regard as the underlying general features of their approaches. What they choose to highlight in both cases is their selection of a holist or reductionist approach. This, I think, establishes, at least a *prima facie* case, that the issue is worth looking at and potentially useful in understanding the controversy between various schools of thought in the history of economics.

* * *

Some examples can be given to show how this controversy continually emerges in economics:

Dore in an extended review of Binmore (1994), comments thus on the latter's deployment of evolutionary game theory: 'with [Binmore's] new definition of self-interest, methodological individualism, a hallmark of neoclassical economics, is abandoned. In its place is a more holistic and integrated view of society as an integral organism' (Dore, 1997: 227). According to Lawrence Boland, 'Demonstrating the dependence of all macroeconomics on microeconomic principles is essential for the fulfillment of the (methodological) individualist requirements of neoclassical economics' (cited in Nelson, 1984: 576). So Dore and Boland agree that the reductionist standpoint – 'methodological individualism' – is an essential component of neoclassical economics. In a similar vein, Schotter, right at the beginning of a book on *Free Market Economics*, says that in the libertarian, or classical liberal view, 'society is nothing more than the aggregate of the individuals composing it' (Schotter, 1985: 2). That is, before anything else, Schotter feels it important to establish the reductionist standpoint of libertarianism.

Identifying partial and general equilibrium (GE) approaches with reductionism and holism, respectively, he points out that Lucas has been extremely active in importing GE thinking into macroeconomics. While the latter point is undeniable, the identification between GE and holism is far more questionable. A GE model incorporating the salient points of analysis explored in Chapter 6, below, would indeed justify the designation of 'holist' – but such a model is, of course, very far indeed from what Lucas is trying to develop.

Kevin Hoover, on the other hand, in a special edition of *The Monist* on 'The Metaphysics of Economics', argues that

> "The idea that macroeconomics stands in need of a microfoundational base is a commonplace among economists … [but] ontological reduction of macroeconomics to microeconomics is untenable … while the program of microfoundations may illuminate macroeconomics in various ways, it cannot succeed in its goal of replacing macroeconomics." (Hoover, 1995)

Feeling it necessary to point out that an example she wants to use for purposes of illustration, although couched in reductionist terms, could be replaced by a more holistic version, Nancy Cartwright remarks that 'We could attempt to explain macroeconomic regularities by reference to capacities and relations that can only sensibly be attributed to institutions or to the economy as a whole, with no promise of reduction to features of individuals' (Cartwright, 1995: 278).

Nelson (1984) is an interesting case since at first he seems enthusiastic about the possibility of reducing macro to micro in economics:

> "The possibility of reducing at least parts of macroeconomic theory to microeconomic theory is especially exciting …. [I]t seems that the large–scale phenomena dealt with in macroeconomics must be the results of the total effects of the small scale phenomena dealt with in microeconomics. Therefore, one might expect that bridges could be built by merely adding up the microeconomic laws describing the microphenomena to obtain the macroeconomic laws describing macrophenomena …. [I]t does seem that a reduction of macroeconomics to microeconomics would not be plagued with the kind of ontological difficulties which might attend [other reduction] schemes" (Nelson, 1984: 573-74).

His conclusion, however, is pessimistic. There is, he concludes, 'little prospect of providing anything which deserves to be called microfoundations … it is unlikely that we will see any progress in reducing macroeconomics to microeconomics' (Nelson, 1984: 593). Cross and Strachan agree that 'The way forward … is to eschew the reductionist search for finer-grain microfoundations and instead [to] study how complex economic systems can emerge from the interactions between agents' (Cross and Strachan, 1997: 565).

Finally, Hayek, who spent much of his life criticising the linked errors of macroeconomics, economic statistics and socialism, was amongst the most trenchant of those wishing to deny the legitimacy of macroeconomics: 'I recognize … microeconomic theory as the only legitimate economic theory' (**KES**: 22); 'I believe it is only

microeconomics which enables us to understand the crucial functions of the market process' (**KES**: 27); '[M]y disagreement with that book [sc Keynes's *General Theory*] did not refer so much to any detail of the analysis as the general approach followed in the whole work. The real issue was the validity of ... macro-analysis'. (**TBT**: 100). As we shall see, however, the citation of these passages by no means settles the question of Hayek's stance on the methodological issue of reductionism versus holism.

* * *

This section has, firstly, argued that while Bunge's triad individualism-systemism-holism is a profound one, the opposition between individualist, on the one hand, and non- or anti-individualist approaches, on the other, is of more fundamental interest for the purposes of this thesis, and, secondly, that for the remainder of this thesis the 'individualist' approach will be referred to as *reductionism*, and the non-individualist approach as *holism*. Further, I have proposed definitions of reduction and holism and tried to show that the opposition between them is a living issue in economics. The next section takes up the issue of the association, which can be seen to emerge here, and which is asserted by Bunge, between methodological premises and policy prescription.

1.3 Policy prescription and social philosophy: reducibility and the invisible hand

A pattern seems to emerge from the examples cited: apparently there is a tendency for libertarians and those on the right of the spectrum of policy views within economics, such as Lucas, to appeal to reductionist methodological premises, while those on the left, those like Keynes adopting a more interventionist stance, are more likely to invoke holistic underpinnings for their theoretical pronouncements. Bunge takes up this point, speaking of the 'obsolete individualism of … Smith … [and] the neoclassical economists' (147); 'the cultural policy of liberalism, which is based on individualism, is one of benign neglect. By contrast, the totalitarian cultural policy, which is based on holism, is one of censorship' (150-151); 'all market worshippers espouse individualism' (151). 'The radical individualists oppose all social planning in the name of individual liberties … holists swear by top-down planning … they are likely to ignore their aspirations and rights [sc those of the common people]. In either case, the powerless individual, whether forsaken or corralled, has nothing to gain' (153). By contrast, 'systemism takes into

account social values (ignored by individualism) as well as individual values (ignored by holism)' (157).

But there is a paradox[9] here. We have just seen that there *seems* to be an association between a *laissez-faire* policy prescription and a reductionist methodology. Nevertheless, starting with Adam Smith{ XE "Smith, Adam" }, a profoundly influential trend, epitomised by writers such as Friedrich Hayek and Armen Alchian, has proposed what has been called an 'invisible hand theorem'[10,11] (De Vany, 1996: 427), in which order in human affairs is 'emergent', 'resulting from human action but not design'[12]. This emergence seems flatly to contradict the association between *laissez-faire* and reductionism just noted. This section will explore that puzzle.

Perhaps the canonical statement of the 'invisible hand theorem' is Smith's statement that

> "It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities, but of their advantages." (**WN** I.ii.2[13])

Even here, in this well-known and apparently simple statement, there is something mysterious about the relation between micro and macro levels. The butcher, brewer and baker do not care about the dinners they provide, so that, in some sense, the desirable social outcome of feeding the members of society is achieved *in spite of* rather than *because of* the motives and behaviours of the food providers. The articulation between

---

[9] I recognise, of course, that the 'paradox' will not seem very paradoxical if one starts out with different expectations about the association, or otherwise, between reductionism and holism, on the one hand, and the various possible policy prescriptions, on the other.

[10] Mario Bunge (personal communication) correctly points out that the invisible hand is not a theorem but a postulate: citing works in which it is referred to as a theorem is not to be taken as endorsement of that usage.

[11] De Vany implies that a 'weak invisible hand theorem', what he calls 'little more than a proposition about the computation of a price vector', would consist in an assertion of the Pareto optimality of a general equilibrium, subject to all the usual caveats of general equilibrium theory (De Vany, 1996: 427). See also Mirrlees (1997: 1311-1312).

[12] The reference is to Adam Ferguson (1767) *An Essay on the History of Civil Society* p187.

[13] That is, in the accepted mode of reference to Adam Smith's works, Book I, chapter ii, paragraph 2 of the *Wealth of Nations*).

the motivation and behaviour of agents at the micro level and macro outcomes is not even at first blush a trivial or straightforward question. It seems odd that the trend which is associated with this 'invisible hand theorem', the tendency to support a *laissez-faire* policy prescription, which as we have seen seems itself to be associated with reductionism, should refer to the 'emergent' properties of human order, when emergence is precisely what distinguishes the holist from the reductionist approach.

The question therefore arises, whether the invisible hand theorem is consistent with the reductionist or the holist approach. Haworth (1994 Ch 4, Section 1 'A logical dilemma', 32-35) performs an invaluable service by addressing precisely this issue. Haworth's procedure – part of his philosophical critique of libertarian thinking – is as follows. Firstly, he identifies two 'theses', implicit in libertarian thought, which he defines as follows, each illustrated by a statement from a libertarian[14] source. Then the libertarians' logical dilemma arises from the mutual incompatibility of the two theses:

The reducibility thesis: the fully developed market economy can be understood as the sum or aggregate of its discrete components, the individual bilateral exchanges at the micro level. Sir Keith Joseph: 'Since inequality arises from the operation of innumerable preferences, it cannot be evil unless those preferences are themselves evil.'[15]

The invisible hand thesis: the market is a 'paradigmatic exemplar' of want-satisfaction (ie, unrestricted market forces leave agents better off than any alternative economic environment) because an invisible hand transmutes our self-interested behaviour into socially desirable outcomes. Mandeville: 'the grand principle that makes us social creatures, the solid basis, the life and support of all trade and employment without exception is *evil*.'[16]

---

[14] Or at least libertarian-approved: Mandeville has been claimed by Hayek as a libertarian thinker (Hayek, 1966).

[15] The reference is to Keith Joseph and Jonathan Sumption (1979) *Equality* London: John Murray, p 78. Colleagues commenting on previous versions have expressed surprise that I allow such an 'outrageous', 'outstandingly silly' statement to pass (Mario Bunge, Richard Sturn, personal communications). Again, to cite an instance of a claim is not in any way to endorse that claim: I wholly share those colleagues' distaste.

[16] The reference is to Bernard de Mandeville in Philip Harth (ed) (1970) *Bernard Mandeville: The Fable of the Bees* London: Penguin. This contains *The Fable of the Bees or Private Vices Made Publick*

So reductionism says that evil only comes of evil, while the job of the invisible hand is specifically to transmute evil into good[17]. For Sir Keith the aggregate outcome cannot be evil as long as the preferences it is based on are innocent; for Mandeville, on the contrary, the aggregate outcome cannot be good unless the preferences underpinning it are evil, vicious, selfish. Thus Haworth is able to conclude that the libertarians cannot have it both ways: 'libertarianism is seriously broken backed in the sense that it must abandon one of its central theses.' (Haworth, 1994: 34)

My own approach is slightly different from Haworth's. Firstly, we need to pay more attention to the association, alluded to earlier, between methodological standpoint and policy prescription. There are policy implications of the choice between reductionism and holism. And, indeed, the consequences for policy implied by the approach selected, so far from being a mere scholium, are the tail which wags the methodological dog[18]. If one adopts the systems approach and recognises that the unintended collective outcomes of an unplanned, uncoordinated mass of individual actions may have far from desirable features, then the obvious implication is to see whether there is anything we can do about it. The absence of an invisible hand invites the intervention of the very visible hand of state intervention. The reductionist approach, on the contrary, says that, assuming individuals can be counted on to do the best they can for themselves given the constraints

---

Benefits (1705 and various dates under various names with accreted text) itself, and the 'Vindication' (1724) Mandeville published after the *Fable* was arraigned before the Grand Jury of Middlesex as a public nuisance.

[17] A correspondent finds this statement confusing: I seem to him to be conflating positive and normative issues. This confusion can be avoided by mentally substituting 'possession of property *x*' for 'evil'. The moral value, if any, of property *x* is irrelevant when considering whether the possession of property *x* by some macro-level entity or phenomenon requires or contradicts the possession of property *x* by the substrate-level entities. It just so happens that 'evil' is the property considered by the representatives of the two theses adduced by Howarth. So Joseph says that the macro level outcomes have exactly he same character as the substrate they are based in, while Mandeville says they have exactly the opposite character.

[18] Of course, this is not a tight, one-to-one relationship: as Ian Steedman (personal communication) points out, very different policy prescriptions may in various ways be made consistent with similar methodological standpoints.

they face, the aggregate outcome of those individual actions will also be the best available: state intervention in the economy is nugatory.

There are two possibilities: we could be living in a world where the reductionists are right or one where the holists are right. Needless to say, I think we reside in the latter. If we lived in the former, the macro level would simply reflect the micro level. There would be nothing for an invisible hand to do. The individual would be directly social, or, what comes to the same thing, there would be no separate category of the social. Individual utility maximisation would directly be social welfare maximisation: the distinction between them would be meaningless. Likewise, macro irrationality would be just a summary of micro irrationality: unemployment would either be a product of irrational behaviour by workers, such as 'pricing themselves out of jobs', or it would be the product of a rational desire for leisure[19], and, hence, itself rational. In general, individuals could with confidence be left to get on with it without supervision or intervention. A reductionist world would be a *laissez-faire* world.

If, on the other hand, we were to inhabit, as in my opinion we do, a holistic world[20], then reductionists would (and do) have a problem. It is fairly obvious that higher level entities are not simply aggregates of their micro components: water does not behave as an aggregate of hydrogen and oxygen; steam, liquid water, and ice do not consist of tiny gaseous, liquid and solid molecules; nor do chairs consist of hard, green, ugly or uncomfortable molecules (Haworth, 1994: 35). As Ken Binmore says, 'As we all know, you would be very lucky if a small change in your program led only to a small change on your screen. Tiny programming errors typically lead to wild and unpredictable results when the program is run' (Binmore, 1994: 213). All these properties emerge at higher levels. The problem faced by the reductionist is how to reconcile this fact – of an obvious disjuncture between levels – with the reductionist *laissez-faire* policy prescription. Libertarians face severe difficulties sustaining a logically consistent reductionism in a holistic world.

[19] As the Duke of Edinburgh memorably asserted on the Jimmy Young Programme, and as Robert Lucas still believes (1987: 66-67)

[20] A correspondent points out that we cannot inhabit a 'holistic world' as holism is a doctrine, not a kind of thing. Clearly this is expression is a figure of speech, shorthand for a world in which a holistic standpoint would be appropriate.

The invisible hand is one potential solution to this problem. There are two possibilities. Either one can ignore the disjuncture between levels, and adopt a thoroughgoing reductionist methodology and policy stance – this seems to be line taken by Joseph, Lucas and Friedman[21] – or with Hayek and Adam Smith one can accept that disjuncture, and so adopt a methodological holism, but at the same time postulate a mechanism reconciling that methodological holism with a *laissez-faire* policy reductionism. Such a mechanism is an invisible hand mechanism. The invisible hand allows us to say, granted that social outcomes are not *logically* bound to reflect individual behaviour in an aggregative, summary manner, nevertheless a mechanism exists which ensures that *in practice* they do so. The invisible hand is what allows us to think, and act, in a reductionist way in a holistic world: it underpins reductionism by tacitly conceding holism. *Laissez-faire* is vindicated, we are inveigled into tying the visible hand behind our back, if we can be persuaded that the invisible hand will do its job instead, and do it better.

What I am suggesting, therefore, is the following: the *laissez-faire* policy prescription does, indeed, embody a reductionist standpoint. However, there is more than one way of sustaining that standpoint methodologically. One can believe, or at least act as if one believes, that the world truly is reductionist in relevant ways and that supposed macro-level pathology is simply the summation of micro-level behaviour which may or may not be pathological. *Laissez-faire* is a reductionist policy prescription in the sense that it issues from a reductionist methodological standpoint. Or one can accept that the world is holistic and hence that macro-level pathologies might in principle be emergent at that level, but postulate the existence of an invisible hand mechanism which ensures that the reductionist policy prescription of *laissez-faire* is nevertheless valid. The latter strategy combines methodological holism with policy reductionism.

* * *

[21] Although few of us attain to consistency, and it is always possible to find holistic-sounding formulations in reductionist writers – the difficulty lies in interpreting them.

Bunge's criticisms of the 'individualism' (for which, following the terminology I adopt, read *reductionism*) of the neoclassical economists are well taken[22]. But the simple, one-to-one relationship between this reductionist standpoint and a *laissez-faire* policy prescription, which Bunge asserts, simply doesn't exist. Compare the standpoints of Hayek and Friedman. For Friedman, economics is based on the study of 'a number of independent households – a collection of Robinson Crusoes' (Friedman, 1962: 13; see also Haworth, 1994: 8). For Hayek, on the contrary, 'individuals are merely the *foci* in the network of relationships' (**CRS**: 59). So, on the definitions proposed earlier, Hayek subscribes to a very clearly holistic, and Friedman to an equally clearly reductionist methodological standpoint. Yet they still both endorse the same basic framework for policy prescription: *laissez-faire*. And in bracketing Smith with the neoclassicals in the reductionist camp (148), Bunge is simply in error – as we shall see in Chapter 4.

Writers such as Smith and Hayek are methodologically very distant from the crude reductionism of Joseph, Lucas and Friedman. They tacitly recognise a holistic world by invoking invisible hand mechanisms. For Smith the invisible hand is literally the hand of an omniscient and omnipotent deity desiring nothing other than the maximisation of human welfare. For Hayek, writing in a more secular age, the invisible hand mechanism takes the form of an evolutionary process based on the exploded group selection theory of VC Wynne-Edwards. Hayek's attempts to distinguish his own stance from that of Keynes, his anxiety to head off a line of thought leading from the holistic or systems thinking premises he shared with Keynes to an interventionist policy prescription, lead him to make the crudely reductionistic statements about macroeconomics and microeconomics which we noted earlier.

The resolution, then, to the puzzle identified at the beginning of this section is as follows. We can understand the contradictory association between an appeal to reductionist methodological underpinnings and an assertion of a holistic invisible hand mechanism on the basis of a dual foundation. The first part is a factual hypothesis about the nature of the world we actually inhabit, namely, the systems view of the world, the hypothesis that the world is a holistic one with all that this implies about the scope and potential for collective intervention in the economy. The second part concerns the *laissez-faire* policy

[22] Although it may well be a rhetorical error blandly to dismiss monetarist macroeconomics as akin to

prescription of the reductionist camp, a policy prescription which both depends on and supports the reductionist methodological stance. Taking the two points together, we can see that some mechanism has to be introduced to mediate between a holist world and a reductionist policy prescription: the invisible hand does the job. The invisible hand, whatever its precise content, whether it comprises the hand of a deity as in Smith, or the result of an evolutionary process as in Hayek, allows us to assert that, in practice, the world can be treated *as if it were reductionist*. The logical inconsistency which Howarth has correctly identified is an internalisation of the inconsistency between reductionist *laissez-faire* programme and holistic world. It is illogical because it is attempting to do what is ultimately impossible, namely to reconcile the irreconcilable.

The alternative to both of these approaches is to combine recognition of the holistic nature of the world we live in with acceptance that there is no invisible hand. In this view, rational individual self-seeking behaviour is by no means either the necessary or the sufficient micro substrate for the desirability of social outcomes. Rather, behaviour must be *directly* social if desirable social outcomes are to be obtained. According to Keynes, for example, egotistical activity uncoordinated by the state may lead to inefficient outcomes. The price system aggregates rational individual actions but the aggregate is an unintended outcome as far as those individuals are concerned. There is no particular reason why unintended outcomes should necessarily be desirable and often they are not. Individuals take responsibility for maximising their own welfare, given what everyone else is doing, but somebody[23] has to take responsibility for organising the aggregate outcome if undesirable aggregate outcomes are to be avoided: 'there is no design but our own ... the invisible hand is merely our own bleeding feet moving through pain and loss to an uncertain … destination' (**CWXX**: 474).

Marx, on the other hand, takes the argument a stage further by arguing, on the contrary, that there is, indeed, a design which is not our own, a design without a designer. Like Hayek, Marx believes this design to be 'emergent' at the macro level, but for Marx, unlike Hayek, *because* it is not our own design, it is alien to us. In the absence of directly social activity, atomistic behaviour spontaneously arranges itself into a self-augmenting parasitic network of social relations which he calls 'capital'. A society of

---

faith healing (156)!

[23] Specifically, a universal class in a position to act on behalf of society as a whole.

individual humans thus becomes dominated by an interest alien to that of the individuals comprising it.

* * *

The conclusion of this section, therefore, is that while reductionist approaches may be safely dismissed, embracing holism is no guarantee of getting it right. There is holism and there is holism. Methodological holism combined with the *deus ex machina* of an invisible hand mechanism can still sustain the inappropriate and unwarranted reductionist policy prescription of *laissez-faire*. The bulk of this section has concerned the holism-reductionism dyad, at the cost of ignoring the distinction between what Bunge calls 'systemism' and what he calls 'holism'. Bunge drew attention to the important principle that 'there are no relations without relata'. We both reject the reductionist assumption that the properties of entities are just aggregates of the properties of substrate entities, but we also reject the assumption that the pattern of relationships constituting a system is substrate-neutral and in some sense logically prior to the properties of the substrate entities.

The implication is that writers in the invisible hand tradition, such as Smith and Hayek, adopt precisely this assumption. As far as Smith is concerned, the case is made in Chapter 4, where the invisible hand is interpreted as literally the hand of an omniscient, omnipotent and benign deity. As far as Hayek is concerned, the issue is dealt with in Chapter 5, where I argue that, for Hayek, macro level objects are understood as independent entities in their own right, owing nothing to their material bases in individual behaviour. In both cases the overall macro-level pattern is divorced from the need for a mechanism, such that micro level incentives are consistent with behaviour which sustains that pattern.

The consequence of this analysis is that the association between methodological standpoint and policy prescription asserted in Bunge (2000) can be challenged. There is no simple association between 'individualism' and *laissez-faire*, on the one hand, nor between 'holism' and intervention or centralisation, on the other. On the contrary, *both* 'individualism' (that is, in my terminology, reductionism) *and* 'holism' (a holism cut off from material foundations in self-seeking substrate activity) are associated with *laissez-*

faire, while 'systemism' (holism which is not so cut off) is associated with a range of policy prescriptions from mild to radical intervention.

In substance, Bunge's article is a significant contribution to our understanding of the methodology of economics – perhaps best embodied in his bold injunction to 'see agency through Weber's microscope, and structure through Marx's telescope' (154). The article, however, can be faulted for an over-simplification and misunderstanding of the relation between methodological premises and policy consequences. There do exist reductionist free marketeers of the kind Bunge describes: the pronouncements of Friedman, Lucas and Sir Keith Joseph fit this pattern. But to lump holists such as Smith and Hayek in with them, is both mistaken and allows us to ignore the fundamentally ideological role of invisible hand mechanisms in allowing economists to retain some approximation to efficiency as their default notion of how the capitalist economy actually works.

1.4 The structure of the thesis

From what has been said above, it is clear that a major part of the thesis must be concerned with the problem of the invisible hand. It is mandatory, then, to return to the *locus classicus* of the invisible hand in the writings of Adam Smith. This will occupy a pivotal chapter, Chapter 4, 'The Invisible Hand of God in Adam Smith'. This will be followed, in Chapter 5, by a close examination of the position of a leading twentieth century invisible hand theorist, Friedrich Hayek. The thesis concludes with an examination of the consequences of adopting a holistic perspective while rejecting invisible hand mechanisms: Keynes's proposals for planning and widening state engagement in the economy. Before turning explicitly to the invisible hand itself, however, I examine two episodes in twentieth century political economy where developments within game theory and the theory of social choice were perceived as challenges to 'the invisible hand theorem'.

Within modern neoclassical economics controversy has been aroused by the discovery of two apparent anomalies or paradoxes. Arrow's impossibility theorem, for example, appears to show that it is impossible to devise a procedure which can be relied upon to aggregate individual preferences into clear and acceptable collective preferences. And

the prisoners' dilemma seems to contradict the view that individual utility maximisation leads in general to collectively desirable outcomes – not just as a theoretical possibility but as a plausible description of a pervasive phenomenon. Considerable energy has been invested in the attempt to reconcile the prisoners' dilemma and Arrovian impossibility with the supposition of a ubiquitous and benevolent invisible hand.

One response to such difficulties, however, has been to jettison the idea of collective rationality altogether. Libertarian writers such as IMD Little, argue that only individuals can think or prefer one thing to another, and concepts of collective rationality and social preference are therefore devoid of meaning. The question as to whether a collection of individuals, a society, can have interests and preferences distinct from those of the individuals of which it consists, has parallels with two related questions in psychology, philosophy and computer science, namely, (a) how brains can be conscious when individual neurons are not, and (b) whether artificial intelligence is possible in principle. Both questions relate to the validity of the computational theory of mind.

Chapter 2 concentrates on the challenge to 'the invisible hand theorem' posed by the prisoners' dilemma. Attention will be drawn to a key feature of the problem facing agents when such dilemmas arise: that decision-making is *interdependent in content but independent in form.* The decision affects others but is taken as if it only affected the individual decision-maker. In Chapter 3 attention is turned to the problem of Arrovian impossibility, and, in particular, to the criticism of Arrow by libertarian thinkers. Taking IMD Little as a principal exponent of this view, a parallel is drawn between his rejection of the notion of a social welfare function and the rejection, by writers such as John Searle, of the possibility of artificial intelligence. Chapters 2 and 3 take the form of an extended review of Barry and Hardin (1982).

What, then, do we learn, with respect to the issues outlined above, from these two chapters? Firstly, they appraise the seriousness of the challenge posed by the prisoners' dilemma and Arrovian impossibility to the reductionist programme. The conclusion drawn is that both are, in fact, incompatible with the reductionist programme, but that the prisoners' dilemma is the more profound, drawing attention (a) to both common and conflicting interests of the agents involved, whereas the Arrow theorem is based only on conflicting interests, and (b) to the tension between social content and private form of

the decision making process. Secondly, the chapters draw attention to a particularly extreme liberal response to these challenges, and provide an explanation of why that response is wrong.

As just mentioned, in a holistic world the prisoners' dilemma and Arrovian impossibility pose considerable difficulties for the reductionist programme. In part this is an empirical matter. Do we really encounter prisoners' dilemmas[24] and non-single-peaked preferences in the world? One possible approach for the reductionists is to say that these do not in fact occur[25]. However, I am not aware of any major attempt to put this argument and therefore will ignore it here. A second possibility is to pick holes in the argument, trying to find an acceptable 'solution' to the two 'paradoxes'. Ken Binmore says well that 'Personally, I see no paradox at all in the fact that independent choice behaviour by rational agents should sometimes lead to Pareto-inefficient outcomes' (Binmore, 1994: 103). But why does he need to say this? Because 'The claim that [that both players defect] is the solution of the Prisoners' Dilemma seems paradoxical to many authors.' (ibid) 'Many' is in fact an understatement here, and very great efforts have, unsuccessfully, been made to circumvent the two propositions. The point is that there is no absolute or logical paradox involved, but both Arrovian impossibility and the prisoners' dilemma are paradoxical in the older, weaker sense of 'contrary to orthodox wisdom' – the sense in which the Good Samaritan is a paradox. Belief in the possibility of automatic mechanisms which aggregate (a) individual preferences to achieve a consistent notion of what is socially desirable, and (b) individual actions to achieve socially desirable outcomes, is incompatible with these two results.

A third approach, taken by the more extreme proponents of *laissez-faire*, is to deny collective rationality and social welfare, to deny, that is, that any meaning can be attached to these phrases. According to this view, only individuals can be rational or have preferences. The lesson of Chapters 2 and 3 is that this tactic, which illegitimately privileges a particular level, that of the individual agent, is fundamentally untenable.

---

[24] ie, one shot games and multi-player iterated games: just those versions where reciprocity cannot lead to the cooperative solution.

[25] Binmore (1994) and Axelrod (1984) are based on repeated two-player games, and much of what they say ceases to apply once *n*-player ($n > 2$) games are admitted.

Chapter 4 is devoted to an examination of the invisible hand in the work of the writer who invented the term. Adam Smith is revered as the father of modern economics. I shall argue that analysis of his writings, however, reveals him to be a representative of the 18th century tradition of 'natural theology', with some medieval elements. Smith is preoccupied with the need to preserve order in society. His scientific methodology emphasises reconciliation with the world we live in rather than investigation of it. He invokes a version of natural law in which the universe is a harmonious machine administered by a benign deity. Nobody is uncared for and, in real happiness, we are all substantially equal. No action is without its appropriate reward – in this life or the next. The social desirability of individual self-seeking activity is ensured by the 'invisible hand', that is, by the hand of a god who has moulded us so to behave, that the quantity of happiness in the world is always maximised.

In Chapter 5 attention turns to a prominent twentieth century exponent of the invisible hand. While Adam Smith proposed that individual self-seeking would lead to socially desirable consequences thanks to the intervention of the 'invisible hand' of a wise and kindly god, Friedrich Hayek seeks to re-establish the invisible hand in a secular age, replacing the agency of a deity with an evolutionary mechanism. A process akin to natural selection ensures that individual behaviours leading to undesirable social consequences are weeded out. Hayek's approach hypostatises the old. Hayek is inconsistent in his deployment of this evolutionary mechanism: only its – for him – benign effects are acknowledged, while its undesirable consequences are blamed on a culture of state intervention.

What is the specific contribution which these two chapters make to the overall thesis? I argued, earlier in this introductory chapter, that the notion of an invisible hand is fundamentally ambivalent: at once it sustains reductionism, and the *laissez-faire* policy prescription implied by it, while tacitly conceding the holist case: social order is emergent. What we want to know is, precisely how this automatic mechanism operates; how the invisible hand is to reconcile individual behaviours in the interest of the greater good of society.

In Smith and Hayek we see two allied but distinct answers to this question. In the writings of Adam Smith we see that the invisible hand is simply the hand of God. Smith's

God has a very simple structure, indeed he is reduced almost to a cipher. Repeatedly and unambiguously we are told that, at least for practical purposes, God has a utility function of one argument: total human happiness[26]. The machine grinds out the summation with perfect accuracy for ever. It is the totality which is always primary for Smith; individuals are assigned very subordinate roles. Deprived of any real freedom or autonomy, they are inveigled, deceived and frightened into doing God's will, that is, supposedly, maximising human welfare.

Is Smith then in the reductionist or holist camp? I argued above that the role of the invisible hand is to reconcile the reductionist programme with a holistic world. This is certainly the case for Smith. Smith admits a holistic world. His imagery is, in tune with his times, mechanical rather than organic: the world is a great machine rather than an organism. But that is not significant. What is key is that Smith's deployment of the invisible hand allows him to *assume* that the maximisation of human happiness takes place automatically – all we need to do is give each other enough space for the invisible hand process to work itself out. Macro level rationality is the spontaneous reflex of micro level rationality.

And what of Hayek? Hayek's standpoint is structurally very similar to Smith's – with the modification that a form of evolution is to replace the deity as the mechanism driving the invisible hand. Not only the allocation of resources between competing ends, but also the institutional environment within which that allocation process takes place, are subject to variation and selection. Only those allocations and institutions most conducive to human welfare survive this weeding and sifting process. Once again, the reductionist programme is reconciled with a holistic world. Social order, and other properties of the macro level, are emergent at that level, but they are born perfect and fully-formed. We have no need to intervene at the macro level. Social welfare is again maximised automatically by individual utility maximisation.

[26] Why 'for practical purposes'? Smith drew a distinction between God ('the Author of the universe') and the universe itself ('nature'): the latter, God's machine for implementing his will in the world, has this utilitarian character, while God himself has additional objectives, such as enjoying virtue 'for its own sake'. These additional characteristics of the deity have a purely cosmetic function in Smith's account.

What this, at first blush more plausible, account omits is how individual self-seeking behaviour leads to institutions and allocations which are just what society requires. Indeed, just as in Smith individual interests were pre-reconciled by God, so that aggregation was unproblematic, so, too, Hayek admits that this will only work for interests that are already reconciled. He then goes on to build his entire system on the wholly unjustified assumption that, indeed, they are thus reconciled.

The thesis concludes with a presentation of the policy stance of John Maynard Keynes and an examination of its links with his underlying social philosophy. The argument is that Keynes shares the holistic approach of Smith and Hayek, but without their reliance on invisible hand mechanisms. If spontaneous processes cannot be relied upon to generate desirable social outcomes then we have to take responsibility for achieving this ourselves. Individual self-seeking behaviour will lead to socially desirable outcomes only if the institutional framework is right. Setting up the institutional framework in which there are no prisoners' dilemmas – or, what comes to the same thing, abolishing the existing institutional framework which atomises individuals – implies large-scale collectivisation of the economy.

Chapter 2 The Prisoners' Dilemma[27]

2.1 Barry and Hardin: rationality at two different levels?

The purpose of this chapter and the next is to show that the issue of the articulation between the micro and macro levels does present problems for contemporary economics. An appropriate place to begin our investigation of the problems facing the neoclassical view of the micro-macro articulation of levels is a book of readings (Barry and Hardin, 1982) on precisely this subject, *Rational Man and Irrational Society?* edited by Brian Barry and Russell Hardin, political scientists at the University of Chicago.

Barry and Hardin's book consists of two parts: 'Individual actions and collective consequences' is a discussion of the prisoners' dilemma, while 'Individual preferences and collective decisions' is about the Arrow impossibility theorem. Both parts of the book are therefore concerned with the micro-macro dichotomy: the failure of rationality at the micro level, the level of the individual agent, to guarantee rationality at the macro level, the level of the whole society. The prisoners' dilemma illustrates circumstances in which rational individuals, attempting to do the best they can for themselves, do worse than if they had tried to do the best they could for all the agents taken together. Arrovian impossibility shows that there is no rule which can always aggregate individual ordinal preferences into social preferences without producing some perverse result, such as that the social ordering is just a restatement of the ordering of a particular individual, or that the society is indifferent between every alternative presented to it. Both results have been regarded as shocking and as deeply damaging to liberal notions of economic and political processes. They have been held to exemplify a contradiction between individual and collective rationality: 'A theme that unifies many of the pieces reprinted here is that

[27] An earlier version of this chapter appeared as Denis (1996a).

of rationality. Both the prisoner's dilemma and the Arrow general possibility theorem are thought to throw up problems for the concept of rationality' (367[28]).

Barry and Hardin are supporters of the libertarian view that, if there is a conflict between individual and collective rationality, we should keep the former and drop the latter. Their response to the supposed dichotomy between collective and individual rationality is to assert that it is based on an illicit extension of the concept of rationality from the individual to the society. Arrow's is a theoretical result based on sufficiently unrealistic assumptions that its relation to the real world is remote, while the prisoner's dilemma is a practical problem challenging us to construct institutions which avoid it.

> "The idea that the prisoner's dilemma or the Arrow theorem evince some sort of 'breakdown' or 'paradox' of rationality should be seen as a reaction to the discovery that the concept of rationality cannot be extended indefinitely to solve all problems of conduct and evaluation." (368) "[T]oo much is expected of the notion of rationality. There is no a priori reason to expect that, if people have very diverse ways of conceiving the bases for ranking some states of affairs above others, it will be possible to aggregate these preferences in a way that will satisfy the conditions laid down by Arrow." (385)

And a similar point can be made in the case of the prisoners' dilemma. Hence, 'There is, if we see things straight, no paradox in either case' (385). Arrovian 'social welfare functions' embody a 'kind of formal rationality' which, according to Barry and Hardin, is not very important. Barry and Hardin prefer a 'substantive rationality', meaning 'whether of not there are reasons which can be publicly supported for doing one thing rather than another' (385). And in the final words of the book they say that 'As far as the prisoner's dilemma is concerned, it should be regarded not as a paradox, but as a challenge. The problem is how to devise institutions that will get us out of them' (386). There is more than an echo of the Kantian procedure about this. Kant argued that unconstrained thought ('speculative reason') would lead to contradictions ('antinomies') and that the purpose of philosophy, therefore, was to seal off those areas where this was the case ('dialectic') and redirect thought into more productive and practical channels:

> "human reason, which naturally pursues a dialectical course, cannot do without this science [sc metaphysics], which checks its tendencies towards dialectic, and ... prevents the ravages which a lawless speculative reason would infallibly commit ... [T]he supreme office of censor which it [metaphysics] occupies, assures to it the highest authority and importance. This office it administers for the purpose of securing order, harmony, and

[28] Unqualified page numbers in this chapter and the next refer to Barry and Hardin, 1982.

> well-being to science, and of directing its noble and fruitful labours to the highest possible aim – the happiness of all mankind." (Kant, 1934: 480-81)

In the same way, Barry and Hardin note that thinking consistently about the aggregation of individual preferences and behaviours leads to paradoxes – contradictions between micro- and macro-level rationality. Instead of seeking an intellectually satisfying resolution of the paradox, we are advised to drop the issue: there is obviously something wrong with the concept of collective rationality so we had better stop using it and turn instead to the practical problems of constructing institutions. An intellectual problem is replaced by a practical one.

As we shall see, however, Barry and Hardin differ from Kant – and, as we shall see, from Adam Smith – in that Kant and Smith at least did not dismiss the idea of 'the happiness of all mankind' as an illusion. For Barry and Hardin, we cannot apply the concept of rationality to society as a whole, we cannot ask whether society is acting rationally to maximise its happiness and the idea of human happiness in the aggregate itself becomes vacuous.

2.2 The Prisoners' Dilemma

Table 1

Payoff matrix for a one-shot prisoners' dilemma game with ordinal payoffs

		X's move	
		Cooperate	Defect
Y's move	Cooperate	2, 2	1, 4
	Defect	4, 1	3, 3

In the payoff matrix shown in Table 1, above, the cells indicate the ordinal payoffs to the two players, *X* and *Y*. Each of the four outcomes is ranked between 1 and 4 for each player, with 1 meaning most preferred by that player. The payoff for *X* is presented first in each cell. Hence the top left cell shows the payoffs if both cooperate, and the top right

cell those if *X* defects and *Y* cooperates. In the latter case, for example, '1, 4' means that this outcome is *X*'s most desired outcome and *Y*'s least desired. In the case where *X* co-operates and *Y* defects, the outcome, represented in the bottom left cell, is reversed. If both players cooperate they both get their second best outcome and if both defect they get their third best outcome. These ordinal preferences can be illustrated by any cardinalisations which preserve the rankings. There is no need for any symmetry or equality of the cardinal utilities attached to each outcome for the two players.

A one-shot, two-player prisoners' dilemma is any game with the payoffs indicated in Table 1. We make the following two assumptions:

(1) The payoffs indicated accurately capture the preferences of the players. In other words, any fellow feeling, moral views, pleasure in cooperating *per se*, etc, are all already included in the ranking.

(2) Each player is rational: he actually behaves in the way which maximises his own utility.

Often a third assumption is specified: that moves are simultaneous – in the sense that each player makes his decision in ignorance of the other's decision. This is actually formally unnecessary: we shall see shortly that each player has a dominant strategy to defect – and this is entirely robust to relaxations of the simultaneity assumption.

Suppose *Y* cooperates. The best that *X* can do is to defect: he gets his first rather than his second best outcome. Suppose now that *Y* defects. The best that *X* can do is still to defect: by doing so he will now get his third rather than fourth best outcome. So *X* has a dominant strategy: whatever *Y* does, he does best by defecting. Since *Y* is in exactly the same position, he also has a dominant strategy to defect. Yet if we compare the outcomes on the main diagonal (ie those outcomes where the players make the same move) we can see that if both defect, both get a less desired outcome than if both had cooperated.

There are four possible outcomes: CC, CD, DC and DD, where DC, for example, means that *X* defects and *Y* cooperates. Three outcomes (CC, CD and DC) are Pareto-efficient,

as any change to another outcome would make at least one player worse off. Only DD is Pareto dominated: a change to CC would benefit both players. The one-shot prisoners' dilemma selects the single Pareto-inferior outcome. Unlike the other three outcomes, DD can also be seen as a Nash equilibrium: it is the only outcome in which each player is doing the best he can, given the actions of the other player.

In the prisoners' dilemma rational behaviour by atomistic agents cannot be relied upon to generate socially desirable outcomes. 'The point is that in this situation if each person acts rationally and pursues his own self-interest, the societal outcome is worse than if each individual does not. In short, the invisible hand of individual maximization seems to break down here' (Schotter, 1985: 49). Indeed, to the extent that the prisoners' dilemma game realistically portrays situations that arise in society – and inspection and introspection would indicate that it describes a situation all too frequently met with in real life – then socially inferior outcomes will be the norm.

The prisoners' dilemma game, however, is only the tip of the iceberg. Slightly varying the assumptions gives rise to a host of similar games generating sub-optimal outcomes. If moves are sequential instead of simultaneous, for example, appropriate adjustment of the payoffs available gives us the centipede game – a game which can never get started[29] even though both players would benefit from playing (Kemp and Philp, 1996). The extent of prisoners' dilemma-type macro-irrationality can be illustrated by reference to two points made by Schelling, the author of Reading 5. The first point simply draws attention to the fact that clashes between micro and macro conceptions of rationality extend well beyond the limits of the prisoners' dilemma narrowly defined. Schelling argues that the prisoners' dilemma is not the only situation where

> "there is one way that everybody can act so that everybody is doing what is in his own best interest given what everybody else is doing, yet *all* could be better off if the *all* made opposite choices ... [W]e should probably identify as the generic problem, not the inefficient equilibrium of prisoner's dilemma, but all the situations in which equilibria achieved by unconcerted or undisciplined action are inefficient – the situations in which everybody could be better off, or some collective total could be larger, by concerted or regulated or centralised decision." (108)

[29] A correspondent correctly points out that formally the game *does* get started in that the first player faces a decision whether to play or not, and if he decides not to play, that is his move (and the last one of the game), so by *not* playing, he *has* played and the game did get started (and terminated instantaneously). I submit that my use of words is more transparent.

As an example, he refers to the problem of what to call the time: whether we choose to call a given moment nine o'clock, or to put the clock back and call it eight o'clock. This, however, leads into the second point I want to make in connection with Schelling's discussion, namely, that Schelling draws attention to the problem intrinsic to the prisoners' dilemma and related games as a *coordination* problem. We may note, to anticipate a later discussion, that a coordination problem is exactly what arises when a decision is interdependent in content but independent in form. The decision has social consequences but is taken as if it had only individual consequences. If the individuals concerned could coordinate their decision-making – so that the payoff matrix could be changed to reflect the choices of just one (aggregate) player amongst two (or more) actions – then all would be better off. Gauthier makes the point in Reading 4, that 'the problem of coordination ... arises in those situations in which there is more than one set of possible actions in mutual equilibrium' (105). (If we use the strict notion of equilibrium rightly prevalent in economics, that of Nash equilibrium, then, as pointed out above, there is only one equilibrium: DD. Gauthier is using equilibrium in a looser sense in order to include the idea of the socially desirable outcome as an equilibrium. The whole problem, however, is that the socially desirable outcome in the prisoners' dilemma is not an equilibrium.)

The decision on what to call the time is a classic case of a nominal variable having real consequences. This decision is not a prisoners' dilemma game, yet it exhibits exactly the same coordination problems as the prisoners' dilemma proper. Barry and Hardin summarise Schelling's discussion of time thus:

> "The essence of the game involving time is that it is a coordination game in which it is more important that we all have the same time than what time it is. But that does not mean it is a pure coordination game, if we understand by that one in which all equilibria are equally good. Without organisation we will be locked in standard time all the year round, even if we would prefer daylight saving time in the summer." (109)

A simple payoff matrix for a coordination game, such as what to call the time, or which side of the road to drive on, might be set out as follows. The assumptions made here are the same as for the two player, one-shot game made above, with the addition of the simultaneity assumption. Whether a one-shot, 2-player game has any real significance will be addressed shortly.

Table 2

Payoff matrix for a one-shot coordination game with ordinal payoffs

		X's move	
		GMT	BST
Y's move	GMT	1, 1	3, 3
	BST	3, 3	1, 1

In the payoff matrix above, *X* and *Y* do not care whether to use GMT or BST so long as both use the same time system. The outcomes on the main diagonal where they both use the same time-naming convention are their equal first choices, and the off-diagonal outcomes their equal third choices. If players only communicate by means of their moves, and they face a one-shot game, then they face a probability of 0.5 that they will coordinate successfully.

The key point about these games – the one-shot, 2-player prisoners' dilemma and coordination games – is the failure of reciprocity: there is no way in a one-shot prisoners' dilemma game that either player can influence the behaviour of the other. Without reciprocity we cannot have coordination.

2.3 Iterated and *n*-player games

In the case of an iterated game – sometimes referred to as a 'supergame' – the decisions taken by players in any period can now be related to decisions made by the other player in the previous iteration. Each player now has to consider what it stands to lose in the future if it defects now *and if that defection causes the other player to defect in the future*. If either (a) the rate at which the player discounts future benefits is very high, or (b) the players' subjective probability of playing further rounds is too low, then, for practical purposes, we still have a one-shot game. Assuming this is not the case, cooperation becomes a viable strategy. The problem is that one does not know the likelihood of the antagonist to cooperate in the future or his responsiveness to one's own moves. It has been claimed on the basis of experimental evidence (Hofstadter, 1985 Ch

29 and Axelrod, 1984, 1990 Ch 2) that the best strategy is one, such as TIT-FOR-TAT, that is 'nice', that is, does not defect first, and which punishes defection by defection and thereafter forgives by returning to cooperation[30]. These last two are necessary – otherwise one's antagonist is unable to influence one's own behaviour (for better or worse) and therefore can treat the game as a one-shot. To resolve the prisoners' dilemma, the supergame must extend into the indefinite future to avoid the problem of backward induction. If the players know when it is going to end, they will defect on the final game. Since they cannot influence their antagonist's play in the final game, they will defect in the penultimate game, and in the one before that. The iteration unravels back to the present and defection becomes the dominant strategy once more. The technical name for this is 'backwards induction'.

With two-player iterated games with high probability of future rounds and shallow discounting of the future, however, there is a real basis for cooperation. Robert Axelrod makes this scenario the entire basis for a theory of *The Evolution of Co-operation* in the biological and social worlds. Many of 'the most important problems facing humanity ... take the form of an iterated Prisoner's Dilemma' (Axelrod, 1990: 190). Clearly, this is open to a very *laissez-faire* interpretation. If social problems which might otherwise legitimise state intervention are characterised by indefinitely iterated 2-player prisoners' dilemmas, then we can expect rational agents to learn to cooperate with each other, rendering state intervention unnecessary.

One way out of the quandary of the prisoners' dilemma, therefore, is to extend the number of periods. One way back into it, however, is to extend the number of players:

> "two players, who should rationally defect in a two-person single play game, commonly should cooperate in iterated play. In general, in an *n*-person single-play prisoner's dilemma, defection is narrowly rational. In iterated play it may not be. Whether it is depends on how many others' choices are likely to be affected by one's own. As *n* becomes very large, it is increasingly implausible that cooperation will be narrowly rational even in iterated play. Hence, in the dynamic analysis, narrow rationality can lead to cooperation in prisoner's dilemma (or collective action) .... Still, cooperation becomes less likely as *n* increases in iterated play." (33-34)

[30] These contentions by Axelrod, *et al*, have been severely criticised by Ken Binmore (1994: 194-203). Binmore's points have themselves been criticised by Matt Ridley (1996: 72-73).

Recall that the incentive to cooperate in a prisoners' dilemma game rests on the assumption that by doing so you will influence your antagonist to cooperate, thereby obtaining more desired outcomes for both players. Cooperation depends on reciprocity. But as *n* becomes large, it becomes impossible to influence the behaviour of other players. How large is 'very large'? Well, actually it is often the case that anything over two can be large enough to forestall cooperation, as the paper by Rapoport (Reading 3) indicates: 'In variants of the three-person Prisoner's Dilemma where the *single* defector gets the largest payoff and the single cooperator suffers the most severe punishment, it is all but impossible for all three to 'balance' themselves on the precarious cooperative solution.' (74)

The players face two problems: how to discriminate between players whom one wishes to reward for cooperating and those one wishes to punish for defecting, and how one's own behaviour can perceptibly reward or punish other players. A competitive market, for example, can be seen as a prisoners' dilemma supergame amongst *n* producers, where *n* is very large, and each producer would like all producers together to reduce output but is not prepared to reduce output unilaterally. In spite of it being an iterated game, cooperation cannot emerge because every producer would want to defect and there is no mechanism to punish just defectors.

One can therefore argue that Axelrod's conclusions are far too optimistic: two-player games certainly do exist, but there is no reason to believe that many of the important institutional dilemmas faced by humanity are not more realistically modelled as multi-player games. In the key issues addressed by Keynes (Chapter 6 of this thesis), for example, agents must decide whether to save or to spend and whether to hold their assets in the form of bonds or money. These decisions clearly impact on many other agents and constitute multi-player iterated games. Such situations cry out for state intervention – for example, to alter the payoff matrices by means of appropriate penalties for defection and incentives for cooperation. In the sequel, therefore, references to prisoners' dilemma without qualification are to be understood without distinction as one-shot, two-player games, finite iterated games, or *n*-player iterated games; that is, as just those situations where the dilemma is manifest.

The problem with the prisoners' dilemma can be stated simply. There is a common interest between the two players. But this common interest cannot be implemented because of its fragmentation into two atomistic players. The agents act irrationally from the collective point of view because they are compelled by the structure of the game to make their decisions *at an inappropriate lower level of aggregation*.

As an aid to perspective, before looking at the issues of interest, where the micro-macro dichotomy poses problems for human society, I will make some prefatory remarks on individual and collective rationality in a natural world context.

2.4 God's Utility Function

A recent book about Darwinism by Richard Dawkins contains a chapter entitled 'God's Utility Function' (Dawkins, 1995, Ch 4: 111-155). For the purposes of argument, Dawkins postulates some aggregate over-arching purpose in the world: 'God's utility function'. By observing nature we should be able to work out what that purpose is, what it is that is being maximised in the natural world. His conclusion to this exercise in reverse engineering is that there is no such over-arching purpose: 'The universe we observe has precisely the properties we should expect if there is, at bottom, no design, no purpose, no evil and no good, nothing but blind, pitiless indifference' (Dawkins, 1995: 155).

This does not mean that the world is senseless or chaotic: on the contrary, the structure and behaviour of the entire organic world can be explained by reference to a single principle, the survival of DNA. Each gene is a maximiser of a utility function with one argument: the probability of its own survival: 'Genes that cause individuals to maximise their descendants are the genes we expect to see in the world. The animals we are looking at inherit the genes of successful ancestors' (Dawkins, 1995: 127).

This individual, self-interested rationality leads to incoherent, purposeless behaviour as soon as higher levels of aggregation are considered:

> "Cheetahs give every indication of being superbly designed for something ... They appear to be well designed to kill antelopes. The teeth, claws, eyes, nose, leg muscles, backbone and brain of a cheetah are all precisely what we should expect if God's purpose in

> designing cheetahs was to maximize deaths among antelopes. Conversely [in the case of] an antelope we find equally impressive evidence of design for precisely the opposite end: the survival of antelopes and starvation among cheetahs. It is as though cheetahs had been designed by one deity and antelopes by a rival deity." (Dawkins, 1995: 122-123)

Even at the level of a single species, or groups within a species, the attempt by individuals to maximise the number of their descendants leads to inefficiency. Dawkins takes the example of the elephant seal. In the elephant seal the females are monopolised by 4 percent of the males in a harem system. The sex ratio is about 50:50 even though 96 percent of the males make no contribution, and, indeed, consume more than half the population's food resources. 'Any utility function that paid even a little attention to the economic efficiency of the community would dispense with the bachelors. Instead, there would be just enough males born to fertilize the females' (Dawkins, 1995: 124).

However, if males were in a minority, it would be to the advantage of an individual to have male offspring: the expected number of offspring of males is greater than that of females. Thus individuals who tended to have male offspring would be selected for until the sex ratio was equal again. An unequal sex ratio is not an evolutionarily stable strategy. But the consequence is massive inefficiency:

> "The wastefulness of the harem economy can be summarized as follows: Males, instead of devoting themselves to useful work, squander their energy and strength in futile struggles against one other. This is true, even if we define 'useful' in an apparently Darwinian way, as concerned with rearing children. If males diverted into useful channels the energy that they waste competing with each other, the species as a whole would rear more children for less effort and food consumed." (Dawkins, 1995: 138)

'God's Utility Function, Dawkins sums up, 'seldom turns out to be the greatest good for the greatest number. God's Utility Function betrays its origins in an uncoordinated scramble for selfish gain' (Dawkins, 1995: 142). The situation Dawkins is describing is precisely a multi-player prisoners' dilemma. The suggestion here is that the problem is *not* that genes make individual organisms or groups of organisms behave in ways – in some sense – contrary to their own interest, though that may be the case. The essence of the problem is that genes, individually maximising their future replication, do not collectively maximise their future replication. If – somehow – the elephant seal genes could cooperate instead of blindly competing with each other, the population could support a skewed sex ratio and make far more efficient use of its resources. The population would be more successful within the species and the species within its overall

environment. The probability of any randomly selected gene in the gene-pool surviving into subsequent generations would be increased. But they cannot do it. Individually rational behaviour by genes leads, in this as in many other cases, to collective behaviour which is irrational for those same genes.

To put it another way. Genes want to survive.[31] Individual behaviours are selected which maximise the transmission of the DNA responsible for those behaviours into the future. But the only evolutionarily stable strategy for the DNA is to do the best it can, *given what all the other DNA is doing*. However, at the collective level, the DNA is *not* doing the best it can, precisely because 'what all the other DNA is doing' is *not* 'given'. In particular, if 96 per cent of males in the species is redundant, then it would be better for the species to dispense with them. The ideal sex ratio is 2:50, or 200/(2+50) per cent = 3.85 per cent males[32]. **Individually** it is rational to invest in male offspring if the percentage of males falls below 50 per cent. **Collectively** it is rational to invest in males only if it falls below 3.85 per cent. Choosing female offspring with a probability of 0.9615 is cooperation and choosing female offspring with a probability of 0.5 is defection in this iterated multiplayer prisoners' dilemma game. If all elephant seal genes cooperated they would all enhance their replication over the Pareto-inefficient Nash equilibrium. In the case of partial cooperation, however, the cooperators would individually do worse than if there were no cooperation, and the defectors better. Hence cooperation cannot arise gradually, and if, somehow, it did arise, it would be vulnerable to invasion by defectors. Cooperation is not an evolutionarily stable strategy.

This section, therefore, illustrates a natural-world case where agents act rationally at an individual level, but irrationally at the aggregate level – because the decision making is independent in form but interdependent in content. It also gives substance to the view that the individual agent in the prisoners' dilemma game is a quite abstract category

[31] For a forceful rejection of the notion that genes can't 'want' anything, see Dennett (1995: 328-329).

[32] 4% of males are productive, and males constitute 50% of the population, therefore, for each hundred elephant seals 4% of 50%, or 2 males should be retained to mate with the 50 females. So the ideal sex ratio is 2:50. The 2 males as a percentage of the population of 52 (50 females plus 2 males) is approximately 3.85%. As a decimal fraction this is 0.03846; taking this away from unity to get the desired probability of female offspring gives us the figure of 0.9615.

which could be a gene or a species, a person or an institution, a nation or even a civilisation.

2.5 Collective and individual rationality

We are now in a position to look in more detail at the Barry and Hardin view of individual and collective rationality with respect to the prisoners' dilemma, which is essentially that collective rationality does not exist – the term is an unwarranted extension of the term 'rationality'. In the Epilogue, they concentrate their fire on Rapoport's use of the term 'collective rationality' in Reading 3. The Rapoport paper opens with these comments:

> "The dilemma in Prisoner's Dilemma arises from the circumstance that the question 'What is the rational choice?' is ambivalent unless 'rationality' is strictly defined. It turns out that in the context of non-constant-sum games like Prisoner's Dilemma actually two concepts of 'rationality' compete for attention, namely *individual* rationality, which prescribes to each player the course of action most advantageous to him under the circumstances, and *collective* rationality, which prescribes a course of action to both players simultaneously. It turns out that if both act in accordance with collective rationality, then *each* player is better off than he would have been had each acted in accordance with individual rationality." (72)

It is true that Rapoport immediately spoils this with references to ethical considerations, in particular, Kant's categorical imperative. This misses the point that *either* the rankings of the two players already impound all considerations of morality, *or* they don't and impounding them *changes* the rankings, releasing the players from any dilemma. But Rapoport is correct in asserting the contradictory claims of rationality operating at two different levels.

Barry and Hardin discuss Rapoport's contention at length.

> "If the end is taken as maximising one's utility," they argue, "then it looks as though there is an irrefutable case here for saying that means-end rationality requires one to play the noncooperative strategy in a one-shot prisoner's dilemma. What can be said against this conclusion?" (379)

They then refer to the passage in Rapoport's article just cited and seize on the references to morality which follow it. After some discussion of ethical considerations and the Kantian categorical imperative, they sum up as follows:

> "We have isolated a plausible moral principle ["that ... you should not take advantage of the forbearance of others ... that if you know the other prisoner has not confessed you ought not to confess yourself"]; but contrary to Rapoport's suggestion ["that rationality need not be identified with the pursuit of self-interest"], morality and self interest do conflict here. An agent who follows the moral prescription in a prisoner's dilemma will do less well than one who does not." (381)

It is fairly easy to show that 'moral' behaviour is irrational from the individual point of view in the one-shot the prisoners' dilemma – if the prisoners' dilemma payoff matrix accurately represents utility payoffs it is irrational to cooperate because of some external moral principle. But that is not the point. The point is that Rapoport was correct to distinguish between individual and collective rationality and to point out the conflict between the two. It is a simple fact that the individually rational outcome, DD, is Pareto inefficient, that is, it is socially irrational.

Barry and Hardin attempt to puncture the claims of 'collective rationality' by pointing out the consequences to the individual players of acting on its strictures:

> "the best thing for a self-interested individual is for the other player to follow 'collective rationality' while he follows 'individual rationality' himself. Now if somehow both could be locked into 'collective rationality', it is true that both would have good reason for choosing that option rather than being left free. But the essence of the prisoner's dilemma is that there is no mechanism for simultaneously determining the choices in that way." (384)

In this last statement Barry and Hardin effectively concede the case they have been contesting, namely, that there *is* indeed a conflict between individual and collective rationality. And, more precisely, that there is no spontaneous mechanism or institution to coordinate individual actions in order to achieve that collective rationality. 'It is, then,' they continue, 'simply not true that the prescriptions of what Rapoport calls 'collective rationality' have any relevance to a self-interested actor, except inasmuch as he will hope that the other player will be attracted by them, to his own disadvantage' (384).

Again Barry and Hardin miss the point. The point is that collective rationality is thwarted in the prisoners' dilemma situation, and there is just no way round it. The individually rational choice is DD, the collectively rational choice is CC, and since there can be no mechanism for the collective interest to be realised other than through the actions of individuals, the latter is not achieved.

2.6 Hobbes and Rousseau

Further light is shed on Barry and Hardin's approach by their treatment of Gauthier, the author of Reading 4 'Reason and Maximisation'. Noting that the existence of prisoners' dilemma-type situations poses problems for the identification of rationality with individual utility-maximisation, Gauthier seeks to draw out a distinction between two kinds of rationality parallel to Rapoport's individual and collective rationality. Gauthier distinguishes between *interdependent* and *independent* action, which he identifies with activity in *civil society* and in *the state of nature* respectively. The rational modes of behaviour appropriate to these contexts are individual utility-maximisation in the state of nature and 'agreed optimisation' in civil society; he also refers to these as 'straightforward' and 'constrained maximisation', respectively. In the latter case, 'rational persons with full knowledge will perform actions leading to an optimal outcome' (105). But the question remains, how we are to achieve this more civilised state of 'civil society'.

Gauthier refers to Hobbes' view[33] that agreement alone is insufficient to move from the state of nature to civil society, and that force is required to underpin such agreement: contract must be backed by legal sanction. Without such backing, in the prisoners' dilemma, for example, 'any mutually beneficial agreement would require each person to act irrationally, and so no one has reason to make such an agreement' (98), on this view. But Gauthier disagrees with Hobbes. Hobbesian man, socialised only by force, despite a change in appearances, actually remains in the state of nature, Gauthier believes. 'The real difference between the state of nature and civil society must be a difference in man, and not merely in the external relations of men' (103). He turns, therefore, to Rousseau's view that the 'passage from the state of nature to civil society produces in man a very remarkable change, in substituting justice for instinct in his conduct, and giving his actions the morality which previously they lacked'.[34] Hence, for Gauthier, civil society, and hence also 'agreed optimisation', is to consist of agreement backed by morality.

[33] The reference is to Hobbes (1651) *Leviathan* Ch 14, 15, 17

[34] Jean-Jacques Rousseau *Du contrat social* (1762) I, viii; trans Gauthier.

The problem with this is that Rousseau does not say, in the passage cited, that the transition from state of nature to civil society is accomplished by a change in morality, but, on the contrary, that the change in morality he refers to is achieved by the transition from the state of nature to civil society. The change in morality is a result, not a cause. To return to the problem of prisoners' dilemma games, the whole problem would be avoided in a society with the moral values which Gauthier endorses (and I with him), because the payoffs attached to the different outcomes would not have the form of a prisoners' dilemma. But that is to wish the problem away. If the payoff matrix correctly represents the players' preferences, then we have not yet achieved Gauthier's civil society and there is no way that players can escape from the dilemma by incorporating a further moral dimension into their decision making. Thus Barry and Hardin are able to use Gauthier's inconsistency to attack the very notion of collective rationality. In the Introduction to Reading 4, they criticise Gauthier's use of

> "'independent' and 'interdependent' as if they were equivalent to 'straightforward maximising' and 'constrained maximising'. [For] the issue is of course whether or not what Gauthier calls constrained maximising (eg playing the cooperative move in a prisoner's dilemma) *is* rational maximising of any sort in the absence of genuine interdependence between the choices of the parties, that is to say in the absence of a real possibility of one party's choice affecting another's." (89)

It is this issue of 'genuine' interdependence which lies at the heart of the problem. In what sense is behaviour in the prisoners' dilemma setting interdependent? Clearly, Barry and Hardin's criterion for 'genuine interdependence', the possibility of influencing other players' moves, is violated. They are talking about *reciprocity*. When that sort of interdependence is present, when there is reciprocity, then there *is no* prisoners' dilemma. Equally clearly, however, there is *some* element of interdependence at work here. When one player makes a decision, that decision has consequences for the other player. There is an externality at work: the private costs and benefits of each possible move, those affecting the individual player, do not fully reflect the social costs and benefits, those affecting both (or all) players. Again, without this condition, there would be no dilemma: each player's moves would be his business and his alone; there would be no game. We can summarise the problem by saying that decision-making in the prisoners' dilemma game is *interdependent in content but not in form.* The decision affects others but is taken as if it only affected the individual decision-maker.

This interpretation of the prisoners' dilemma is very significant. Virtually every economic transaction that agents make can be said to have some element of this substantial interdependence combined with formal independence. While it is possible to imagine a barter transaction in which no-one other than the two agents involved had any interest, every transaction involving money must necessarily impact on other agents throughout the economy – at the very least influencing the demand for money. The scope for macroeconomic externalities is thus considerable. We have already seen that this scope goes far beyond the prisoners' dilemma proper to encompass a host of other situations where there is a coordination problem and consequent sub-optimal outcome.

But it is also the case that if the problem lies with decision making that is interdependent in content but independent in form, then the solution implies bringing form and content back into mutual accord. It is difficult to see how this can be done by removing the substantial interdependence. That would seem to involve a shift to a Robinson Crusoe economy – without even the company of a Man Friday. Removing the formal independence on the other hand – given the ubiquity of partially overlapping and partially conflicting interests which lie at the heart of the human condition – must surely imply very widespread intervention in and detailed supervision of the workings of the economy by central authority.

2.7 Conclusion

In this chapter I have sketched out a number of propositions. In particular, that coordination problems are endemic in both the natural and social worlds and that the collective irrationality such problems lead to is epitomised by the prisoners dilemma. The idea that prisoners' dilemmas are anomalous or paradoxical does not arise from a mistaken application of the concept of rationality to a sphere where it has no meaning, but from the correct insight that behaviour in prisoners dilemma-type contexts is genuinely pathological. I have argued that the source of the dilemma in prisoners' dilemmas and, more broadly, in coordination problems in general, is the combination of independence in form and interdependence in content of the decision making process. A group of agents having a common interest may be unable to implement that interest if it is fragmented into a number of parts, each of which is compelled to make decisions at an inappropriate lower level of aggregation. Finally, if correct, this argument implies a

search for solutions in the direction of extensive supervision of the economy by central authority. The next chapter, on Arrovian impossibility, continues the discussion begun here.

Chapter 3 Arrow's Impossibility Theorem[35]

3.1 Introduction

I wrote in the previous chapter that the prisoners' dilemma had been perceived as a challenge to the Smithian 'invisible hand theorem' – the view that individual rationality of economic agents will, in general, lead to collective rationality at the level of the economy as a whole. I also argued that the attempt to explain away the problem, by attributing it to a mistaken extension of the concept of rationality from the individual to the collective sphere, was ultimately untenable. In this chapter attention is turned to the problem of Arrovian impossibility, and, in particular, to the criticism of Arrow by libertarian thinkers. Taking IMD Little as a principal exponent of this view, a parallel is drawn between his rejection of the notion of a social welfare function and the rejection, by writers such as John Searle, of the possibility of artificial intelligence. As in Chapter 2, a principal reference point is the collection of writings on *Rational Man and Irrational Society* edited by Barry and Hardin (1982).

The background to the debate over the Arrow impossibility theorem is as follows. Consumers have preferences over the alternative consumption bundles with which they are faced. Modern microeconomics is based on the idea that everything we need to know to discuss consumer behaviour can be obtained from ordinal rankings by the consumers of these alternative consumption bundles without any need to know how much they enjoy a bundle, or how much more they enjoy this bundle than that. An older, cardinal, theory required that, in principle, these levels and differences in utility were measurable. The liberal tradition had for long held that alongside all the particular interests in society, there was a general interest, the interest of society itself. With the

[35] An earlier version of this chapter appeared as Denis (1996b).

cardinal utility theory this conception of a general interest received an obvious interpretation: Benthamite utilitarianism said just add up the utility levels of all the individuals in society and that total represents the welfare of the whole of society. Then any policy can be judged according to how it increased or decreased that total of social welfare. Significant difficulties with the measurement and interpersonal comparison of utility levels led to the abandonment – for the most part – of cardinal utilities and the adoption of the weaker, and hence more robust, ordinal theory. Although it was now no longer obvious how to construct a social welfare function (hereafter SWF), it was assumed that this was merely a technical problem and that the task could still be accomplished in principle. Arrow, however, in a number of publications, in particular the two editions of *Social Choice and Individual Values* (1951, 1963), showed that it was in principle impossible to derive any SWF purely from individual rankings of social alternatives which satisfied certain elementary criteria, such as consistency and non-dictatorship. This was a great shock and, indeed, it would not be an exaggeration to say that some theorists despaired at this result. Plott sets the tone here in Reading 12 (of Barry and Hardin, 1982):

> "The subject began with what seemed to be a minor problem with majority rule. 'It is just a mathematical curiosity', said some ... But intrigued and curious about this little hole, researchers ... began digging in the ground nearby ... What they now appear to have been uncovering is a gigantic cavern into which fall almost all of our ideas about social actions. Almost everything we say and/or anyone has ever said about what society wants or should get is threatened with internal inconsistency. It is as though people have been talking for years about a thing that cannot, in *principle*, exist ... " (231-32)

As the editors say, in their Introduction to the reading, 'Plott maintains that the Arrow results, and others like it in the theory of social choice, undermine the whole tradition of liberal or individualistic political theory that has been developed in the last few centuries' (230).

A major purpose of this chapter is to evaluate this response.

3.2 The conditions

Arrow's theorem says that no social choice procedure can exist which satisfies the five conditions **O**, **U**, **P**, **D** and **I**. The argument proceeds via a *reductio ad absurdam*: we assume that all the conditions hold and then demonstrate that they lead to a

contradiction. Relaxing any of the conditions removes the impossibility of deriving a SWF, but, in general, such SWFs will exhibit perverse features. The conditions are expressed in many different ways in the literature and what follows is based on Barry and Hardin's Introductory essay to Part II. It should, however, be noted that the conditions specified in this form are not genuinely primitive: condition **O**, for example, contains criteria of both uniqueness and transitivity. Transitivity by itself could be expanded into two (or more) conditions. The uniqueness criterion is replicated in Condition **I**, rendering the latter partly redundant.

Condition **O**

The SWF is a unique **O**rdering of the alternatives facing society based only on individual orderings. An ordering is a consistent ranking. In particular this implies transitivity for both preference and indifference. If society prefers *x* to *y* and *y* to *z*, then it prefers *x* to *z*, and so on.

Condition **U**

The social choice rule must have **U**nrestricted domain: it must work for *every logically possible combination* of individual orderings.

Condition **P**

The social choice rule must be **P**areto-efficient: if one individual prefers *x* to *y* and all other individuals either prefer *x* to *y* or are indifferent between *x* and *y*, then the SWF must prefer *x* to *y*.

Condition **D**

There must not be a **D**ictator, that is, a person whose preference of *x* for *y* is always (in every logically conceivable constellation of preferences) the social preference, for any *x* and *y*, regardless of the preferences of others.

Condition **I**

The social ordering of any pair of alternatives *x* and *y* is a function solely of the individual orderings of *x* and *y*: it is **I**ndependent of irrelevant alternatives – individual orderings of *x* and *z*, for example.

3.3 Proof of the theorem

One further concept is needed: that of *decisiveness*. If a group or individual is decisive over *x* and *y*, and prefers *x* to *y*, then society prefers *x* to *y*, whatever anyone else's preference may be[36]. We also assume that society consists of a finite number of individuals. The proof then proceeds by showing that (a) if there is an SWF which satisfies conditions **O**, **U**, **P**, and **I** (that is, all except non-dictatorship), then for some constellation of preferences there must be a decisive individual, and (b) if there is a decisive individual then he is a dictator.

Consider any pair of alternatives, *x* and *y*, where society prefers *x* to *y*. It cannot be the case that everyone in society prefers *y* to *x*, by condition **P** (Pareto). There must be a set of decisive individuals. If only one person prefers *x* then the set only contains one person; if everyone prefers *x* then the set of all the individuals in society is decisive. Normally, it will be a set of intermediate size, but that is irrelevant. There will thus be a (non-empty) decisive set for each pair of alternatives where society is not indifferent between the two. Consider the set of all of these decisive sets. From the assumption that society consisted of a finite number of individuals, this set of decisive sets must have a smallest member, or a subset of equally large smallest members, in which case we pick any member of this subset. We can show that there must be a possible pattern of preferences for which this smallest set of decisive individuals has only one member.

We will suppose initially that this smallest decisive set, *V*, has more than one member, and show that this leads to a contradiction. Suppose *V* is decisive over *x* and *y*, and that

[36] Decisiveness is not the same as dictatorship. Decisiveness holds if there is *one pair* of alternatives and one individual or group such that the individual or group can make the choice between the two alternatives without reference to anyone else's preferences. Dictatorship implies the same, but for *every pair* of alternatives.

it (and hence society – because V is decisive) prefers x to y. This must lead to a contradiction. Since it consists of more than one member we can divide it into two parts, one, V_1, consisting of one member and the other, V_2, consisting of all the other members of V. We also give the name V_3 to the set of all the members of society not in V. Condition **U**, unlimited domain, tells us that the SWF must work for any logically possible pattern of preferences. So we can pick any pattern of preferences we like. Suppose the pattern of preferences is that for V_1, $x > y > z$[37], for V_2, $z > x > y$, and for V_3, $y > z > x$. For convenience of reference this information is set out in Table 3.

Table 3

Preferences of the three sets of agents, V_1, V_2 and V_3

Rank	V_1	V_2	V_3	S	S'
1	x	z	y	x	x
2	y	x	z	y	} y, z
3	z	y	x	z	

Now, we know that V is decisive over x and y, so for society $x > y$. this is shown in the column headed S. But where does society rank z? Suppose society preferred z to y. Only V_2 prefers z to y – both V_1 and V_3 prefer y to z – so that would make V_2 decisive. But V_2 is one person less than V, the smallest decisive set, so that is not possible. So society either prefers y to z (Column S) or is indifferent between y and z (Column S'). Hence society must prefer x to z, given $x > y$ and $y \geq z$, by transitivity (condition **O**). These two alternatives ($x > y > z$, and $x > y = z$, are shown in columns S and S', respectively, of Table 3. But now V_1 is decisive since both V_2 and V_3 prefer z to x. However, V_1 consists of only one person, so the assumption that the smallest decisive set, V, consisted of more than one person turns out to be self-contradictory. We have shown, therefore, what was required, that there is a pattern of possible preferences such that there is a decisive individual. That completes the first part of the proof.

[37] Where, in an obvious notation, the greater than sign, >, means 'is preferred to'.

The second part of the proof shows that a decisive individual is a dictator. Suppose *A* is decisive for *x* against *y* and that he also prefers *x* to *z*. Also suppose, invoking condition **U**, that everyone prefers *y* to *z*. Condition **P** says that society prefers *y* to *z*. If every individual prefers *y* to *z*, it would certainly be Pareto-inefficient for the SWF to prefer *z* to *y*, or even to be indifferent between them. Hence, $x > y > z$ for society and so, invoking transitivity, $x > z$. But condition **I**, independence of irrelevant alternatives, says that the social choice between *x* and *z* is independent of individual preferences over *y*. Hence, regardless of others' preferences, *x* is preferred to *z* by society if and only if *A* prefers *x* to *z*. Hence, if *A* is decisive for *x* against *y*, we can replace *y* by any other alternative which *A* finds less desirable than *x*. Similarly we can replace *x* by any alternative which *A* finds more desirable than *y*. Hence for any possible pair of alternatives, *A*'s preference is decisive: he is a dictator. But the first part of the proof showed that there was a possible set of preferences such that there was decisive individual. Hence, for this pattern of preferences there is a dictator. But this violates condition **D**, of non-dictatorship. The concept of an SWF which simultaneously satisfies all five conditions is therefore inconsistent.

What this shows is that there is a possible pattern of individual preference orderings such that a social ordering derived from them which satisfies the Pareto and independence conditions must be dictatorial if it is to be consistent.

3.4 Scope for relaxing the assumptions

All of the assumptions mentioned are invoked in the proof, so relaxing any will make it possible, in principle, for us to construct a SWF. However, the conditions are generally regarded as quite minimal so the resulting SWFs are unlikely to be attractive. If we are prepared to accept a dictator, for example, there is no problem in constructing an SWF – but now the SWF has nothing social about it. The SWF abdicates before the task of *aggregating* individual preferences. Again, recall that we commenced the proof by considering any pair of alternatives, *x* and *y*, where society prefers *x* to *y*. If we are unable to do this, because society is indifferent between every pair of alternatives, then we can derive a SWF, which, however, is completely vacuous: $x = y$ for all *x* and *y*. This is going to the opposite, but equally useless, extreme.

We assumed that society consisted of a finite number of individuals, and if we relax this assumption then the proof fails, for there need not be a smallest decisive set. Or, to put it another way, if the smallest decisive set were infinitely large, removing one member would not leave a residue smaller than the original set. Again the proof would fail. Since actually existing societies consist of a finite number of individuals, this is scarcely helpful. To consider infinite societies we would have either (a) to consider 'individuals' not to be individual at all but rather infinitely subdivisible, or (b) to regard the individuals composing society to be not just those currently living but also the unborn and/or the dead – and even that would be problematic. In neither case would it be possible to establish individual preferences, let alone aggregate them.

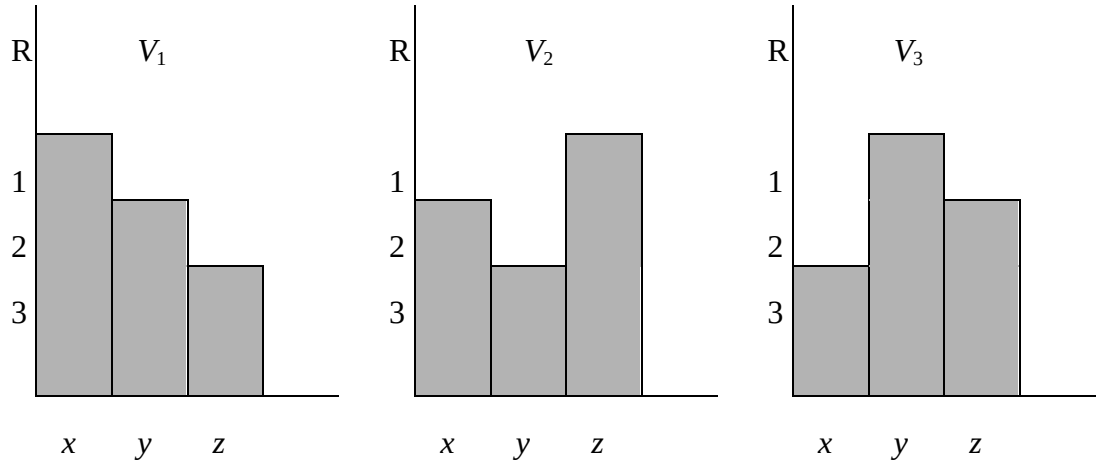
SWFs violating condition **P** or transitivity are no more attractive than those already considered. An SWF capable of choosing Pareto-inferior outcomes would clearly not be maximising society's welfare, which would contradict the notion of an SWF. Violations of transitivity would also lead to irrationality at the macro level. Consider the set of preferences presented in Table 3, which was used in the proof of the impossibility theorem. This set will play a significant role in the sequel, when we turn to the 'Paradox of Voting'. Condition **U** says that our SWF must apply to this set. Suppose that we examine the three pairs of alternatives, *x* and *y*, *x* and *z*, and *y* and *z*, in turn, and adopt a simple majority voting rule. We will obtain an SWF in which $x > y > z > x$, in each case by a two-to-one majority. These are referred to as cyclical majorities. Note carefully that this does *not* say that society is indifferent between the three outcomes – which might be unhelpful but would not be inconsistent. What it says is that *every* outcome is preferred to both alternatives. Permitting intransitivity means that such instances cannot be excluded. In a formal system, once you can prove a contradiction you can prove any statement. So the contamination of irrationality immediately spreads beyond these three alternatives: every possible alternative will be preferred to every other. This is worse than the violation of condition **P**.

Can we make more progress by relaxing condition **U**? Considerable research effort has gone into attempts to obviate the impossibility result in this direction. For example, we could require the SWF only to work where preferences are single-peaked. This would eliminate the possibility of patterns of preferences such as that in Table 3, since, as can be seen from Figure 1, below, which plots rank in the individual orderings against outcome

for the three sets of actors, V_1, V_2 and V_3, the set of individuals composing V_2 were assumed to have dual-peaked preferences. Since the proof depended on this pattern of preferences, it will not hold if such preferences are not in the domain to be considered. However, the resulting SWF will have nothing to say about the cases where preferences are not single-peaked. Now it is well-known that the latter, far from being an exotic theoretical possibility, is a practical problem of pandemic proportions. No SWF of any interest can simply remain silent on these cases. One way out would be to say that society is indifferent between all the outcomes in a preference cycle. This is referred to as 'taking the transitive closure' of the preference relation.

Figure 1

'Peakedness' of preferences for the three sets of agents, V_1, V_2 and V_3



In the proof of the impossibility theorem, the three groups of agents were assumed to have preferences as shown above. The vertical axis shows the ranking the individual ascribes to each of the three outcomes, *x*, *y* and *z*, with 1 indicating the most preferred outcome.

This procedure immediately runs into two major problems: (a) it violates yet another assumption, condition **I**, and (b) by doing this we have in any case restored condition **U** and, with it, impossibility. The point here is that it is not possible to tell whether there is a preference cycle without checking preferences between other alternative pairs than the one under consideration. We cannot decide society's preference between *x* and *y* without also knowing that between *x* and *z*, and *y* and *z*. The social ordering of *x* and *y* is no longer independent of 'irrelevant' alternatives. Turning now to the second problem, condition **U** could be satisfied by a rule for one set of circumstances together with a rule for the remaining circumstances plus a rule for deciding when to apply which rule. That

is exactly the situation which now obtains. Let us rehearse the relevant part of the proof of Arrow's impossibility theorem, substituting the new rule.

Suppose that the smallest decisive set, *V*, has more than one member, and that it is decisive for *x* against *y*. This must lead to a contradiction. Since it consists of more than one member we can divide it into two parts, one, V_1, consisting of one member and the other, V_2, consisting of all the other members of *V*. We also give the name V_3 to the set of all the members of society not in *V*. Condition **U**, unlimited domain, tells us that the SWF must work for any logically possible pattern of preferences. We first of all have to decide whether preferences are cyclical or not and then either treat as normal or invoke the new transitive closure rule, as appropriate. Suppose, as before, that the pattern of preferences is that for V_1, $x > y > z$, for V_2, $z > x > y$, and for V_3, $y > z > x$ (see Table 3). Clearly we do here have cyclical preferences so we invoke the revised rule and say that society is indifferent between *x*, *y* and *z*. But this contradicts our assumption that *V* is decisive for *x* against *y*. Hence the initial assumption, that the smallest decisive set contained more than one member, turns out to be contradictory. So there is still a pattern of preferences for which there is a decisive individual and, therefore, also, a dictator. The proof still goes through.

Condition **I** (together with the uniqueness criterion associated with it – and with condition **O**) has, of all the five conditions, provoked the most controversy, not to say confusion, in the literature. The reason is, perhaps, that, at first blush, is seems utterly counter-intuitive to exclude the non-uniqueness which goes with cardinal individual preferences, and the dependence on 'irrelevant' alternatives which goes with Borda-type systems. One's instinct is to feel that if a SWF can work with relatively information-poor ordinal preferences, then, surely, it should work that much better with relatively information-rich cardinal preferences. Again, Borda systems, which (*pace* Barry and Hardin p219) are neither purely ordinal nor cardinal, are more information rich than purely ordinal systems. It seems perverse to rule out such procedures *ab initio*.

To take this view, however, is to lose sight of what Arrow is trying to do. The point is to demonstrate the possibility or impossibility of building SWFs on the basis of purely ordinal individual preference orderings. It is no good showing that such an SWF is possible (or at least not proven impossible) if that possibility was due entirely to leaving

the door open for cardinal preferences. That would defeat the object of the exercise. The SWF is impounding into the social ordering some extra information – either an arbitrary element that has nothing to do with social welfare or information about individuals' cardinal preferences. Either way, for the ordinalist project to succeed, such influences must be excluded. Again, the Borda system allows agents to reveal not just their ordering of two alternatives but also some information about the intensity with which that preference is held. Preferences over third, fourth, etc, alternatives are not 'irrelevant' and cannot be ignored. We have *some* grounds – albeit inconclusive – for believing that the individual who ranks *x* and *y* tenth and twentieth respectively, holds his preference for *x* over *y* with greater intensity than another individual, who ranks them 16th and 15th, respectively, prefers *y* over *x*. Borda impounds this partial evidence on cardinal preferences. Again, if the SWF is only possible because Borda has not been excluded, the ordinalist project fails. It is therefore essential for the conditions **O** and **I** to be retained. Dilution here spoils the whole point of the exercise.

3.5 The libertarian response

We have already seen the seriousness with which many took the Arrow results. Plott, in Reading 12, 'Axiomatic social choice theory: an overview and interpretation', spells out the breadth and depth of its consequences, on his interpretation. To understand the extent of the concern Arrovian impossibility has stimulated, we need to get a feel for the scope of the alternatives and processes that Arrow's theorem is about.

> "An option or social alternative," Barry and Hardin explain, "could be a complete description of the amount of each type of commodity, the amounts of various types of work done by each individual, the production level of each firm, the type of government agencies and the services provided by each, etc ... The set of feasible options could be a consumption possibilities set ..." (244)

As for the processes which might exist to make choices between the alternatives society is faced with,

> "the process could be a competitive process, a capitalistic process, a socialistic process, or any other kind of process. There is no need, for example, for the process to be directed in that some judge, administrator, or planner uses the defined social ranking to determine the best option and then directs its implementation. The process could be any type of game, voting process, market process, political process ..." (244)

The very comprehensiveness of the Arrow results has forced some writers into a fundamental re-evaluation. Plott states the case for abandoning the very notion of social preferences:

> "Some, like myself (Plott 1972)[38] would claim that the concept of social preference itself must go. Buchanan (1954a, b)[39] was right in his original criticism of Arrow, that the concept of social preference involves an illegitimate transfer of the properties of an individual to the properties of a collection of individuals. For me, the Arrow theorem demonstrates that the concept of social preference involves the classic fallacy of composition, and it is shocking only because the thoughts of social philosophers from which we have developed our intuitions about such matters are subject to the same fallacy." (242)

'In order to see how extreme this position is', Plott continues, we should investigate the ways in which this fallacious concept of social preference continually dogs our thought about our social environment. The idea that 'No one really accepts or uses the idea of a social preference' is completely wrong, he says:

> "many commonly used concepts are equivalent to the concept of a social ranking ... concepts like social needs, group wants, etc. These are simply expressions of priorities and are thus rankings of options ... Take for example the concept of economic welfare. To different options one attaches a number ... indicating the level of welfare. Certain forms of cost-benefit analysis are attempts to operationalise such a formula. But indicators of social welfare clearly imply a ranking of social options according to the numbers which indicate the levels of welfare. The ranking satisfies all of our principles of social preference, and thus the [Arrow] theorem stands as a criticism of any such formula. The only admissible definitions of welfare are those which are dictatorial." (244-45)

This is incorrect. It is not 'the concept of economic welfare' *per se* that the impossibility theorem rules out, but comparisons of alternative states of economic welfare based solely on *ordinal* information about individual agents' preferences.

Lest the significance of Plott's final remark about dictatorship should escape the reader, we will see how the theme is developed in Reading 14, 'Social choice and individual values' by IMD Little. Little argues that there cannot be any such thing as a social preference (in the sense of a preference on the part of society rather than a preference of an individual about society), a 'scientific ethic', or an 'objective moral code'. 'We can

[38] The reference is to Charles R Plott (1972) 'Ethics, Social Choice Theory and the Theory of Economic Policy' *Journal of Mathematical Sociology* 2: 181-208.

have an "objective" ethic only when there is universal agreement ... when all moral judgements will [in any case] have become futile and redundant' (279).

So Arrow is chasing a chimera. 'The very use of the phrase 'social welfare function' suggests ... if such were possible ... an objective moral code' (278). Not only is the idea false; it is much worse: it is an ideology containing the seeds of dictatorship.

Society, in Little's view, is riven by the clash of antagonistic interests. Philosophers have interpreted this antagonism to be merely a superficial appearance masking a more essential harmony of interest. This general interest is then identified with the state and used to delude individuals into conspiring against their own freedom to pursue their individual interests. On the one hand we have an identification between the idea of the general interest and Arrow's conception of an SWF, and on the other between the idea of the general interest and dictatorship. Hence Little is able to locate Arrow in an intellectual tradition leading from Rousseau via Hegel to modern totalitarianism:

> "Political philosophers used to worry over the question 'Why ought I to obey the state (or society, or the general will)?' The pseudo-puzzle of how one can both be free and be subject to law is a variant of the pseudo-puzzle of how duty and self-interest can be reconciled. The philosophers tried to show that it was always really in one's own interest to abide by whatever the 'objective' code might be. Since, plainly, conflict must arise, at least in the absence of complete initial consensus, and since such consensus was obviously absent, they invented the doctrine of a metaphysical consensus. When people actually (ie really) disagreed about some matter affecting the common interest, they were really (ie metaphysically) agreeing. Rousseau ... was first responsible for this nonsense ... 'Each of us puts his person and all his power under the supreme direction of the general will.' (1950: 15)[40] 'Whoever refuses to obey the general will shall be compelled to do so by the whole body ... he will be forced to be free.' (ibid: 18) ... It is but a very short step from here (a step which Hegel took) to maintain that acceptance of the social order (or obedience to the state) is really only self-obedience. Arrow's problem clearly has so much in common with Rousseau's that it seems worthwhile to point out again the insidious danger of this approach. Modern totalitarian philosophy may be not altogether unjustly fathered on Rousseau." (278-79)

Little is far from being isolated in these views. We have already seen how close they are to Plott's. And Barry and Hardin, in the Chicago tradition of political economy, share the view that 'collective rationality' is a meaningless term – an illicit extension of the

[39] The reference is to James M Buchanan (1954a) 'Individual Choice in Voting and the Market' *Journal of Political Economy* 62: 334-343, and James M Buchanan (1954b) 'Social Choice, Democracy and Free Markets' *Journal of Political Economy* 62: 114-123.

concept of rationality beyond its proper domain. While Little tars Arrow with the brush of totalitarianism, Barry and Hardin subject him to criticism of a severity bordering on the intemperate in their Introduction to Reading 13, by Arrow. Arrow's views are, they say, 'extraordinary', 'grotesque', 'bogus', 'clearly ... once again, absurd' and 'an illusion' (249-51).

It is Little, however, who can be taken as a representative of a trend, and it will therefore be worth examining his views in more detail. Later in this thesis, in the chapter on Adam Smith, we shall see that the stick Little uses to beat Arrow, the claim, as Pope puts it, that 'all discord, [is] harmony not understood', is far more aptly turned on Smith himself and his followers.

3.6 Little: the argument against the existence of an SWF

Perhaps the first point to make is that Barry and Hardin misunderstand the context of the debate. With cardinal utility theory it was straightforward – in principle – to construct an SWF: just add up all the individual utilities. With ordinal utility this was no longer obvious, and the next step was simply to show whether it was or was not possible to construct an SWF within this paradigm. Arrow showed that it was not. Barry and Hardin, Little and Plott ignore this. Thus, for Barry and Hardin,

> "[t]he single most penetrating and fundamental statement in the article by Little comes ... when he says 'that it is foolish to accept or reject a set of ethical axioms one at a time. One must know the consequences before one can say whether one finds the set acceptable ...' ... Nobody has any immediate views about the desirability of, say, the independence of irrelevant alternatives, and we should refuse to be bullied by a priori arguments to the effect that would be 'irrational' not to accept it." (265-66)

But this is to ignore the purpose of Arrow's work. The acceptance of condition **I**, as I have argued above, is essential if we are to carry out Arrow's research agenda. If we leave it out and we find that an SWF is not in principle impossible, but that possibility is in fact a consequence of the possibility of obtaining more than bare ordinal information about individual preferences, then we simply have not answered the question we set out to address. It has nothing whatsoever to do with the 'ethical desirability' of condition **I**.

---

[40] The reference is to Jean-Jacques Rousseau (1950) *The Social Contract* trans GDH Cole, New York: EP Dutton.

A similar argument applies to the other four axioms. They do not need to be accepted 'one at a time' – whatever that might mean – they have to be taken all together or Arrow simply hasn't answered the – perfectly legitimate – question he has set himself.

A primary objective of Little's paper is to distinguish between two interpretations of an SWF: (a) as a social decision procedure (legitimate), and (b) as a judgement about social welfare (illegitimate). He criticises Arrow on two grounds (a) to the extent that Arrow's theorem is an objection to social decision procedures, Little complains that it is too extreme – viable social decision procedures can and do exist, and (b) to the extent that an SWF is considered as a social preference, which according to Little cannot in principle exist, it is, he thinks, misleading and dangerous to talk about it as though it could.

Little's major point, therefore, is to argue that a SWF, if it means anything at all about preferences, merely reflects someone's *opinion* about what is good for society, not what the society wants:

> "the so-called 'social welfare function', postulated by welfare economists, should on my view be regarded as a social ordering *only* in the sense that it orders states of society ... Instead of writing, with Bergson, $W = W(U_1, ..., U_n)$, we can write $W_i = W_i(U_1, ..., U_n)$ (i = 1, ..., n). There is no need ... to introduce a further (social) welfare function of the form $W = W(W_1, ..., W_n)$. We can deduce the whole effective corpus of welfare economics from, say, $W_{10} = W_{10}(U_1, ..., U_n)$ – remembering only to put 'in the opinion of individual No. 10' after 'welfare' whenever we use the term." (272)

Because each *individual* ordering of the social alternatives is now to have the status of a candidate *social* ordering – a possible ordering *for* society – Little must now address the issue of whether, and to what extent, the welfare of others is impounded into one's own ordering:

> "In the preceding section it was supposed that each individual ordered states solely according to what he himself would get ... But of course there is no need to suppose this ... In fact, quite generally, we may suppose that they arrange all states in order of what they regard as ultimate desirability, taking everything they know and feel into account ... The problem, then, is to form a 'master'-ordering; in Rousseau's language the problem is the well-known one of discovering the general will." (275)

Actually, it is completely irrelevant, in the aggregation procedure assumed by Arrow, whether individuals take other individuals' welfare into account or not. The SWF is a purely formal procedure aggregating individual preferences without regard to their

construction or content. Indeed, if we look at the outcome, there is a potential perversity here. Suppose there are two individuals, *A* and *B*, with their own sets of preferences and that the possible outcomes are points in *n*-dimensional space (where *n* is the number of issues), designated by *a*, *b* and *s*, representing the two individual interests and the general interest respectively. The Arrow impossibility theorem says that no ordinalist procedure can guarantee to find s. Suppose that we have a constellation of preferences such that an SWF is, in fact, able to find *s* – the theorem by no means prohibits this. If *A* and *B* represent their preferences accurately the outcome is *s*. If, however, one party, say *B*, instead of representing just his own preferences, puts forward a compromise between his own and *A*'s preferences, the outcome will be somewhere between *s* and *a*. *B* is penalised for being 'reasonable'. By taking others' interests into consideration, *B* is left worse off and, possibly, there is a loss to society as well. On the other hand, if *B* were to misrepresent his preferences as being further away from *A*'s than in fact they are, he would be rewarded for this by drawing the social outcome further towards *b*.

Little makes use of his assumption that individual orderings impound opinions about what is good for society, as we shall see. He points out, quite fairly, that an Arrovian SWF must be a formal procedure and hence capable of being mechanised. He asks us to imagine a machine into which we feed all the individual preferences which the machine aggregates according to the formula in order to output a social preference printed on a card. This result could not, he says, consistently be accepted by anyone whose own ordering differed from it. To the potential objection that the individual orderings are just that, *individual* orderings, and therefore that anyone may accept the SWF as such although it diverge from their own preferred ranking, Little responds '[t]his objection does not, however, apply in the case under consideration, because in this most general version it is presumed that the individual orders take the welfare of others into consideration' (276), which, as we have just seen, is simply incorrect.

> "Now if I think x is better than y," continues Little, "and the machine announces that y is better than x ... I can say 'to hell with the machine; x is better than y whatever it says.' If I always did this (and how could I do anything else, since, remember, we are supposing that all other people's values are already known; indeed I have taken into account everything I think significant!) then I should naturally have refused to accept the condition of non-dictatorship. It is in the nature of value judgements that the only order which I can fully accept is one that coincides with my own, regardless of the orders of other people. In other words, no one can consistently accept the condition of 'non-dictatorship'." (277)

There are two remarks to be made about this passage. Firstly, there is an ambiguity here about the meaning of 'accepting' an SWF – between (a) adopting it as one's own ordering, and (b) understanding that it is society's ordering. The only order one can 'fully accept', that is, adopt as one's own, is one that already coincides with one's own ordering. But one is not asked to 'fully accept' the SWF, simply to understand that it *is* the social ordering. This Little cannot understand, since he is unable to conceive of a social ordering in the first place. The second point is that the last sentence, and, indeed the whole drift of the article, is that, if they are to be consistent, everyone must wish to be a dictator. This merely shows that the logical consequence of consistent individualism is solipsism. It is also somewhat ironic, given his remarks linking Arrow and Hegel with totalitarianism.

Let us return to the argument about mechanising the SWF. Little begins by citing Arrow on the benefits of the axiomatic procedure: 'one of the greatest advantages of abstract postulational methods is the fact that the same system may be given different interpretations' (Arrow, cited in Little, 275). But Little says that by interpreting his decision-making process as a SWF he has given it a 'nonsensical interpretation':

> "Imagine the system as a machine which produces a card on which is written 'x is better than y' or vice versa, when all individual answers to the question 'is x better than y?' have been fed into it. What significance can we attach to the sentence on the card, ie, to the resulting 'master'-order? First, it is clear that the sentence, although it is a sentence employing ethical terms, is not a value *judgement*. Every value judgement must be *someone's* judgement of values. If there are n people filling in cards to be fed into the machine, then we have n value judgements, not n + 1. The sentence which the machine produces expresses a ruling, or decision, which is different in kind from what is expressed by the sentences fed into it. The latter express value judgement; the former express a ruling between these judgements. Thus we can legitimately call the machine, or function, a decision-making process.
>
> "But what would it mean to call the machine a social welfare function? One would be asserting, in effect, that if the machine decided in favour of x rather than in favour of y, then x would produce more social welfare than y or simply be more desirable than y. This is clearly a value judgement, but it is, of course, a value judgement made by the person who calls the machine a SWF. Thus, in general, to call the machine a SWF is to assert that x is better than y whenever the machine writes the sentence 'x is better than y'. Now we may suppose that the individual who calls the machine a SWF is one of those who has fed his own value order into it. It is clear that this person must be contradicting himself unless the 'master'-order coincides with his own ordering ... In other words it is inconsistent both to call the machine a social welfare function and to accept the condition of non-dictatorship." (275)

I have already commented on the solipsism Little expresses in the final sentence of this passage. He is simply incapable of raising his eyes above the limited horizon circumscribing 'the individual'. We have also seen how vacuous is Little's assertion that anyone accepting an SWF which was not identical with his individual ordering would be acting inconsistently. He is not asked to accept it *instead* of his individual preferences, merely to understand that it is what *society* prefers. But what is interesting about the passage is the assertion, as something so obvious, once stated, that it needs no supporting argument, that society cannot form and hold a preference. Little says that there are n preferences not $n + 1$. In fact, there *are* $n + 1$ preferences, considered formally: n personal preferences and one social one. The latter cannot really be considered as extra, however: it has more the relationship of a whole to its individual parts.

What is interesting here is the privileging of the level of the individual person and the denial that the society or the machine or, in general, the *system*, can have preferences or intentions. This is very reminiscent of John Searle's denial – particularly in his notorious 'Chinese room' thought experiment – of the possibility of machine 'intentionality'. It is to Searle's standpoint we now turn, in the next section, in order to shed additional light on Little's argument.

3.7 Searle and Little

In setting up his by now well-known Chinese Room thought experiment, Searle's objective is to attempt to discredit artificial intelligence (AI) theory and the computational theory of mind. An AI program, for example one for reading stories in Chinese and answering questions about them, is a formal procedure and can be written down in English. Searle, who knows no Chinese, asks us to imagine him in a sealed room with such a program (he doesn't know what it is) written down in the form of a book of rules. Chinese texts – a story and then questions about it, are passed under the door, although Searle has no knowledge of what they are. He manipulates the symbols on the paper according to the instructions and passes the results back to the Chinese people on the other side of the door. Unknown to him they are answers to the questions about the stories. They are good enough to pass the Turing Test, that is, to convince his Chinese audience at least 50 per cent of the time that the room contains a native-

language Chinese speaker. But Searle still doesn't understand Chinese. Hence – and this is the punch-line – even if we admit AI programs powerful enough to pass the Turing Test, they still won't be *intelligent* because they won't *understand* what they are doing. However good AI programs and machines get at mimicking intelligent life, it still won't be intelligent because when *we*, humans, do something it *means* something to us – but when machines do exactly the same thing it means nothing to them. The modern *élan vital* which is to distinguish, not the living, but the thinking being, is *intentionality* (Searle, 1980, 1984).

One of Searle's keenest critics has been computer scientist and psychologist Douglas Hofstadter, who, together with the philosopher Daniel Dennett, has gone to considerable lengths to defend the computational theory of mind. Hofstadter's response

> "is basically the 'Systems Reply': that it is a mistake to try to impute the understanding to the (incidentally) animate simulator [Searle]; rather it belongs to the system as a whole ... The weakness of Searle's position is that ... he merely insists that some systems [humans] have intentionality by virtue of their 'causal powers' and that some [machines] don't." (Hofstadter and Dennett, 1982: 374-5)

It would be out of place to go into more detail here on the original Chinese Room. The interested reader should consult: Searle (1980), Searle (1984) Ch 2 'Can computers think?', Hofstadter and Dennett (1982) Ch 22, Poundstone (1988) Ch 11 'Mind: Searle's Chinese Room', and Dennett (1993).

Little's thought experiment can be re-cast in the Chinese Room format as follows. It is Little, now, who is imprisoned in the room with a book of instructions and who has to process the characters on slips of paper which are pushed under the door. Unknown to him the slips of paper are individual orderings, the book constitutes the SWF and the output which he pushes back under the door is the social ordering. Now, just as Searle thought that no understanding was taking place as he personally did not understand any Chinese, Little can claim that no preferring is going on in the SWF Room, as he cannot perceive it. Just as the answers Searle provides to the questions were not *his* answers (he doesn't even know that they *are* answers, let alone what they say) so the social preferences provided by Little are not *his* preferences (he, likewise, doesn't even know they *are* preferences, let alone what they say). All Little experiences is (a) his own preference which has been submitted for processing along with everyone else's, and (b)

the utter tedium of carrying out the mechanical processes dictated by the book of rules. The answer is the same: it is not to be expected that Little will experience himself preferring anything since the preference formulated in the SWF Room is not *his* preference. The preference attaches to the whole system: all the individuals who are asked in some way to code their preferences, the formula for aggregating them, the room and its furnishings, Little (or the computer we would normally expect to do his job), and so on. What is the special ingredient – corresponding to Searle's notion of intentionality – whose absence prevents the output of the SWF Room from being a genuine preference? Little does not tell us.

In conclusion, therefore, we have seen that Little adopts an illicit dualism which privileges the level of the individual person. His approach implies that there are two fundamentally diverse kinds of thing in the world: individual humans, which can prefer, and everything else, which cannot. Systemists such as Hofstadter and Dennett have argued strongly that all sorts of systems – including genes and memes, and complexes of genes and memes, individual organisms and collectives of individual organisms, species and populations – can be sensibly thought of as having interests and hence as preferring one state of affairs to another. For Little, however, the privileged status of the individual is simply an assumption without explanation.

3.8 How serious a problem is Arrow's impossibility theorem?

The above account has shown how the libertarian right has responded to Arrow's impossibility theorem by rejecting the concepts of social rationality and of a social welfare function. I have also suggested that this rejection is ultimately untenable, as it illegitimately privileges the individual and adopts a dualistic standpoint. We need now to say something about the seriousness of the challenge the impossibility theorem poses for the invisible hand theorem.

In my opinion the challenge posed by Arrow's theorem is not so serious as that posed by the prisoners' dilemma. The latter shows something profound: namely, how agents can be locked into a situation where their individually rational behaviour leads to socially sub-optimal outcomes. Further it shows that this occurs when there is interdependence but no reciprocity, when decision making is social in content but privatised in form.

What does Arrow's theorem show? It shows that there is no aggregating procedure which is *guaranteed* to produce an acceptable SWF, based only on *ordinal* individual preferences, for *every* conceivable constellation of preferences. Is this so serious? That depends on (a) whether the constellations of preferences encountered in reality are anything like the constellations that cause the Arrow problem, and (b) whether the restriction to ordinal preferences is a legitimate one.

On (a), I have already explained, above, that the problem lies in non-single-peaked preferences, and that such preferences do occur widely in reality. Non-single-peaked (or bimodal) preferences occur in the presence of indivisibilities: when one wants 'all or nothing', for example. Half a loaf is better than no bread, but half a baby is worse than none. Babies are indivisible, a fact that Solomon was able to use in judging a difficult case. Indivisibilities are ubiquitous in a heterogeneous world. Non-single-peaked preferences form the basis for the 'paradox of voting'. Suppose three voters having the preferences illustrated in Table 3 and Figure 1, above or in the example below. Think of them as parties or candidates standing for election. Voting would produce a two-to-one majority for *x* over *y*, for *y* over *z* and for *z* over *x*. Majority voting cannot produce a consistent result. Now, although this is known as the 'paradox' of voting, there is not much paradoxical about it. All it says is that if people have different preferences, they may not be able to agree. Indeed, one can imagine a simpler case: two agents, *A* and *B* have preferences over two alternatives regarding a £20 note, (i) that *A* has the £20, and (ii) that *B* has it. Again, they will not agree. Nothing very profound seems to be going on here. The paradox of voting, and hence the Arrow theorem, which incorporates it, only illustrates the problems that arise when agents have conflicting interests; the bite in the prisoners' dilemma arises from its combination of conflicting *and* converging interests.

On (b), the seriousness of the challenge posed depends on how strictly neoclassical one is. The neoclassical paradigm depends upon the assumption of ordinal and interpersonally non-comparable preferences. Whether it is right to do so is debatable and has inspired a huge literature, and it would be inappropriate to take up this very controversial issue here. I will simply note that we do actually make interpersonal comparisons so frequently and unselfconsciously as to invite the speculation that the brain contains special organs for that very purpose. If I bang my finger with a hammer,

and you lose your leg in a road traffic accident, few would hesitate to say who was likely to be worse off, or who had suffered the larger decrement in happiness. But for those neoclassicals also committed to the ordinalist paradigm, as in fact most are, the Arrow result is indeed a problem.

There is a further point to be made about the significance of Arrow's result. This starts by recognising that it is a special case of Gödel's incompleteness theorem. The latter says that in any consistent formal system (of at least the level of complexity of arithmetic) there must be true statements whose truth cannot be proven within the system, that is, it cannot be both complete and consistent. The SWF is a formal system and must at the very least impound arithmetic – how else is it to aggregate preferences? – and Arrow's general possibility theorem shows that it cannot be complete (ie satisfy condition **U**) and consistent (condition **O**) while at the same time satisfying the other three conditions (**P**, **I**, **D**).

Gödel proceeds by formulating a statement in a formal language, say, PC (for predicate calculus), which says 'this statement cannot be proved in PC'. It would not be possible to prove the statement without proving a contradiction – for if one proved that it was not provable, by proving it one would have proved that it *was* provable. Hence a statement and its negation would both be true and PC would be inconsistent. Hence it is the case that it is unprovable; but that is just what it asserts, so it is also true.

So Gödel works by setting up a formal system and then importing a paradox, namely, a version of the Liar Paradox, 'this statement is false'. The version actually used says, not 'this statement is false', but 'this statement is unprovable in PC'. Arrow also works by setting up a formal system and importing a paradox: the paradox of voting. The paradox of voting, as we have seen, says that if three voters hold the preference orderings

$$A: \quad x > y > z$$
$$B: \quad y > z > x$$
$$C: \quad z > x > y$$

then majority voting on each pair yields the binary social preference rankings

S: $x > y$, $y > z$, and $z > x$, in each case by a majority of 2 to 1.

Hence, for society, every option is preferred to every other option: the social ranking is intransitive (and hence not an ordering).

In the Arrow proof, *A*, *B* and *C* are, respectively, one member of the smallest decisive set, the complement of the smallest decisive set, and the smallest decisive set minus one person, *A*. The decisive set referred to is decisive for *x* against *y*. The pattern of choices set out above is then deployed to show that wherever *z* is placed on the social preference ranking, this must imply that either *A* or *C* is decisive, contradicting the assumption that they were both smaller than the smallest decisive set. This result is then used to show that if all five conditions are satisfied simultaneously, the system is inconsistent.

Completeness in the Gödel context means that all and only the true statements are provable. Completeness in the Arrow context means that for every possible pattern of preferences there is a unique social ordering. Incompleteness is shown in Gödel by importing a 'paradoxical' or self-defeating statement: if it can be proved (ie, if the system is complete) then the system must be inconsistent. Incompleteness is shown in Arrow by importing a 'paradoxical' or self-defeating pattern of preferences: if an SWF is derived from it the system is inconsistent.

It has been suggested that the Arrow impossibility theorem presents as much a problem for proponents of planning as for those of *laissez-faire*[41]. This is false, even apart from the cardinal-ordinal issue mentioned above[42]. In the case of Gödel's theorem, a human intellect standing outside the formal system in question can detect the truth status of the statement 'this statement is unprovable in PC'. We can see that the statement is true and that the formal system is incomplete. Further, we can see how to make it more complete by incorporating true but unprovable statements into the system as axioms, thereby expanding the system. Similarly, as we encounter paradoxical preference constellations, we can expand the preference aggregating procedure to encompass them by identifying

[41] In correspondence in response to an earlier version of this chapter.

[42] Planners who attempted to implement the programme 'from each according to their abilities, to each according to their need' would be required to measure utilities cardinally and to make inter-personal comparisons, if the second half of the slogan is to mean anything.

cyclical preferences, taking their transitive closure and impounding the result as a new axiom of the aggregating procedure. This, to be sure, will still leave the aggregation procedure vulnerable to further paradoxical preference constellations, but as soon as one is encountered it can be treated in the same way. If the number of policy options is finite then eventually all comparisons will have been made and we will have a complete social welfare ordering. Otherwise, the result is an infinite regress with each obstacle being overcome, only to give way – potentially, at least – to a new one. It is, perhaps, in the nature of things that we can get there, not all in one go, but only as an unending series of approximations.

3.9 Conclusion

In this chapter I have argued that Arrovian impossibility, like the prisoners' dilemma, has been perceived as presenting a fundamental challenge to the Smithian 'invisible hand theorem'. This challenge has prompted a number of conservative thinkers in the social sciences to argue that we should abandon the concept of collective rationality itself. Micro-level rationality is all that we can ask for. The policy prescription, therefore consists of maximum freedom for individuals to maximise their individual utilities, coupled with passive acceptance of whatever emerges at the macro level. As Barry and Hardin say, sympathetically summarising Little's view, 'The outcome is the outcome, and it may have nothing apart from that to be said for it' (268). Pursuing a libertarian critique of the idea of a SWF, these writers accuse Arrow, and hence, implicitly, mainstream neoclassical economics, of standing in a tradition leading from Rousseau and Hegel to modern totalitarianism.

In this chapter I have also argued that Little's argument against the possibility in principle of a social welfare function closely parallels Searle's Chinese Room thought experiment designed to show the impossibility in principle of artificial intelligence. Both arguments fail for the same reason: they involve a dualistic vision of the world in which individual humans are set apart from the rest of nature by some innate quality – such as 'intentionality' – the absence of which is supposed to prevent systems other than individual humans from being conscious or forming purposes.

Finally, I noted the links between Arrow's and Gödel's results, and argued that, although still an important challenge for the neoclassical paradigm, the Arrovian theorem is of less theoretical significance than the prisoners' dilemma. While the essence of the Arrow theorem was a conflict of interest between agents, that of the prisoners' dilemma was a combination of conflict and convergence of agent interests.

Chapter 4 The Invisible Hand of God in Adam Smith[43]

4.1 Introduction

The two previous chapters have looked at the mid- and late-twentieth century response of political economy to two perceived anomalies which have been seen as challenges to the invisible hand hypothesis: the prisoners' dilemma and Arrow's impossibility theorem. In the present chapter attention returns to the roots of this tradition in the writings of the eighteenth century father of nineteenth and twentieth century economics: Adam Smith.

Adam Smith has been lauded by the economists – both orthodox and heterodox – of the last two centuries as the founder of their discipline. As Heilbroner says 'An aura surrounds Smith, endowing his name with an authority not enjoyed by any other worldly philosopher except Marx' (Heilbroner, 1986: 1). On the orthodox side, for just two examples of Smith's continuing influence at all levels of the discipline, see the references to the 'invisible hand' in the 'industry standard' introductory economics textbook, Begg, Fischer and Dornbusch (1991: 9, 50, 260), and in the 1996 Nobel Prize Lecture (Mirrlees, 1997: 1311). On the heterodox side, it is well known that Marx, for example, divided political economists into 'scientific' and 'apologetic' classes, with Smith and Ricardo representing the pinnacle of the scientific or 'classical' group and, roughly, everyone after Ricardo being consigned to the apologetic, or 'vulgar' trend (Marx, 1972: 501). I want to argue here, however, that there is a very significant apologetic[44] aspect

[43] An earlier version of this chapter, and material derived from it have appeared as Denis (1997, 1999a, and 2000). The chapter as a whole is at point of writing under consideration at *Research in the History of Economic Thought and Methodology*.

[44] The question of what it is that is to be defended, or apologised for, is a subsidiary theme of the chapter. Essentially, Smith is trying to defend two potentially incompatible things: the existing system of ranks and orders of society in politics, and the 'simple system of natural liberty' in economics.

to Smith, which has as yet received little attention, and, further, that this apologetic aspect is intimately concerned with Smith's conception of the articulation between micro and macro levels, between individual actions and social consequences.

It is of course easy to point to specific passages and throw up one's hands at the ease with which Smith satisfies himself that we are living in the best of all possible worlds – and just as easy to dismiss such passages as *obiter dicta* unrelated to his basic theme. Here, for example, is a famous passage, the second, in fact, of the three occasions on which Smith makes use of the notion of an 'invisible hand', but for all that not always given in full:

> "The rich ... are led by an invisible hand to make nearly the same distribution of the necessaries of life, which would have been made, had the earth been divided into equal portions among all its inhabitants, and thus without intending it, without knowing it, advance the interest of the society, and afford means to the multiplication of the species. When providence divided the earth among a few lordly masters, it neither forgot nor abandoned those who seemed to have been left out in the partition. These last too enjoy their share of all that it produces. In what constitutes the real happiness of human life, they are in no respect inferior to those who would seem so much above them. In ease of body and peace of mind, all the different ranks of life are nearly upon a level, and the beggar, who suns himself by the side of the highway, possesses that security which kings are fighting for." (**TMS** IV.1.10)

So the poor should be content with their lot – they are just as well off as the rich in the things that really matter. Perhaps the typical reaction on reading this is to dismiss it as a vulgar aside, a mere personal prejudice, having no bearing on Smith's scientific researches. This, however, would be profoundly mistaken. The thesis of this chapter is that Smith's whole system of thought can be understood as aiming, not so much at discovery of the world, but at reconciliation with it – indeed, he plainly says as much – and the notion of the 'invisible hand' lies at the heart of this reconciliation.

This chapter examines the evidence in Smith's works – in particular *The History of Astronomy*, *The Theory of Moral Sentiments*, and, to a lesser extent, *The Wealth of Nations* – that this was indeed his approach. The next section constitutes a bibliographical preamble discussing the relationship of each of these three works to Smith's overall *Weltanschauung*. Then a section on *The History of Astronomy* argues that in his major methodological work Smith presents a view of science as an activity aimed, in the first instance, at reconciling us with the world, rather than at theoretically apprehending it. Section 4 presents Smith's stoic conception of the world as a

harmonious machine operated by a utilitarian deity. This conception first arises and is presented with great clarity in *The Theory of Moral Sentiments*; it is then applied to, or rather, simply imposed upon, the social world in *The Wealth of Nations*. A subsequent section establishes the links between Smith and his contemporaries, showing how profoundly in tune he was with the *Zeitgeist* of the second half of the eighteenth century. The section also discusses his failure to deal with some critical contradictions in his system. The conclusion notes two possible responses to Smith: that an evolutionary mechanism can replace a utilitarian deity as a mechanism ensuring that macro optimality corresponds to micro rationality; and, alternatively, the recognition that there is no such automatic mechanism behoves us to construct one ourselves.

A preliminary caveat is in order. My purpose is not to deny the enormous contribution which Smith has made to the development of economics. That would be absurd. A full account of Smith would present those contributions alongside what (as I argue in this chapter) are major shortcomings in Smith's work. My purpose here is far more circumscribed; namely, to trace one particular feature, albeit a key feature, of Smith's thought. The object is to show how Smith believed that the hand of God would invisibly, 'by that eternal art which educes good from ill' (**TMS** I.ii.3.4), ensure that uncoordinated individual actions would always lead to 'the greatest possible quantity of happiness' (**TMS** VI.ii.3.1), and to show how this belief is related to his philosophy as a whole.

4.2 Bibliographical note

I shall be referring to the two works which Adam Smith published in his lifetime, the *Wealth of Nations* and the *Theories of Moral Sentiments* – hereafter referred to as **WN** and **TMS**, and one posthumous work, known, misleadingly, as the *History of Astronomy*, and referred to here as **Astronomy**. The edition of these works that I will refer to is that contained in Adam Smith (1976-1980) *The Glasgow Edition of the Works and Correspondence of Adam Smith* Oxford: Clarendon Press/OUP (reprinted (1981-82) Indianapolis: Liberty Fund). The relevant volumes of the *Works* are:

– Volume I (1976) *The Theory of Moral Sentiments*, ed AL Macfie and DD Raphael[45] (hereafter **TMS**);

– Volume II (itself in two volumes) (1976) *An Inquiry into the Nature and Causes of the Wealth of Nations*, ed RH Campbell, AS Skinner and WB Todd (**WN**);

– Volume III (1980) *Essays on Philosophical Subjects*, ed WPD Wightman and JC Bryce (**EPS**). This volume includes the **Astronomy**, the *Letter to the 'Edinburgh Review'* (1756), and a number of other miscellaneous items by, and about, Adam Smith.

This *Works* is regarded as the definitive edition: enormous scholarly efforts have been made, by comparison with all published and unpublished sources, to present Smith's work in its most mature and finished form, with variations reported in footnotes and appendices throughout. References to Smith's writings, by convention, will be, as appropriate, to the Part, Section, Chapter, subsection and paragraph to preserve consistency with other editions. Editorial introductions are listed separately in the bibliography to this thesis and referred to as such for the sake of clarity.

Adam Smith (1723-1790) published two books in his lifetime. His first, **TMS**, first published in 1759, was also his last: the 6th edition in 1790 contained extensive revisions and additions worked up by Smith in the last year of his life. **WN** was first published in 1776; the 5th edition appeared in 1789 and a 6th, posthumous, edition in 1791. The *History of Astronomy* is a fragment of an uncompleted larger work entitled *The Principles which Lead and Direct Philosophical Enquiries; Illustrated by the History of Astronomy, by the History of the Ancient Physics, and by the History of the Ancient Logics and Metaphysics*. The fragments come to just under 100 pages, most of which consists of the part on the *History of Astronomy*, giving the whole its conventional name. **Astronomy** was first published in 1795 in a posthumous volume entitled *Essays on*

[45] The editors themselves sometimes seem a little hazy as to who exactly the editors of the individual volumes are, and what their order of priority should be. According to **TMS**: ii, it is 'edited by A.L. Macfie and D.D. Raphael'; everywhere else, including the title page, the order is reversed. The title page of **EPS** says 'edited by W.P.D. Wightman and J.C. Bryce' while everywhere else, including on the cover, Bryce's name is omitted.

Philosophical Subjects edited by Joseph Black and James Hutton. It was written at various times, but textual analysis reveals that even the later parts were written before 1758, and most likely the main part was drafted in the late 1740s (Wightman, 1980: 7-8). This makes it nearly contemporaneous with the composition of **TMS** (1st edition, 1759). Shortly before he died, Smith had the bulk of his papers burnt (**EPS**: 327n), only passing a few to his literary executors for them to make up their own minds on the question of possible publication. In an earlier (1773) letter to Hume, Smith specifically singled out **Astronomy** as being possibly worth publishing (Wightman, 1980: 27; **EPS**: 328n).

The point here is to establish that the three works referred to can be viewed as the products of a unified system not as disparate milestones on an intellectual career culminating in **WN** – in which **TMS** and **Astronomy** have interest only as stages in Smith's thought and as the prehistory of the later work. Anikin is a clear example of the stance I am arguing against here: **TMS**, he says, 'is important today mainly as a stage in the formation of Smith's philosophical and economic ideas' (Anikin, 1975: 187). On the contrary: **Astronomy** and **TMS** are products of the same period and the same system of thought; the changes in **TMS** between the first and sixth editions are such only as to clarify and give more substance to this basic system; and the system remained unchanged in its basic outlines after five editions of the **WN**.

> "The so-called 'Adam Smith problem' was a pseudo-problem based on ignorance and misunderstanding. Anybody who reads **TMS**, first in one of the earlier editions and then in edition 6, will not have the slightest inclination to be puzzled that the same man wrote this book and **WN**, or to suppose that he underwent any radical change of view about human conduct. Smith's account of ethics and of human behaviour is basically the same in edition 6 of 1790 as in edition 1 of 1759 .... It is also perfectly obvious that **TMS** is not isolated from **WN** (1776)." (Raphael and Macfie, 1976: 20)

Indeed, Viner (1958: 215), contradicting his main thesis of irreconcilability between **WN** and **TMS**, notes that Smith was already including his major economic principles, 'the essence of his fully developed doctrine, as expounded in the *Wealth of Nations*' in a lecture of 1749, at the same time that he was writing the **Astronomy**, and long before publication of **TMS**. Contrary to the view which sees major discontinuities between Smith's views in **TMS** and **WN**, not to mention earlier works, we can treat Smith's system of thought as a unity, rather than a process. It is that unity which we are to investigate here, and the question to be addressed is: How are micro and macro levels articulated in Smith, how does the invisible hand actually work?

4.3 Smith's methodological stance

> "Starting with Adam Smith's history of astronomy, the main theorists of classical economics sought to capture the essence of the scientific method in order to employ it in the sphere of economic research .... For Smith, the essence of science was the evocation of order, wonder[46] and intellectual delight; it was primarily an esthetic response." (Mirowski, 1989: 198)

As noted in the previous section, what is commonly known as Smith's 'History of Astronomy' is more properly called, in full, *The Principles which Lead and Direct Philosophical Enquiries; Illustrated by the History of Astronomy; by the History of the Ancient Physics; and by the History of the Ancient Logics and Metaphysics*. The full title makes clear that Smith's intention is to set out his conception of scientific method[47]. As far as Smith is concerned in his discussion of successive schools of thought in these *Histories*, the purpose of a system of thought is not so much to disclose the truth of how the world is, but, principally, to soothe the imagination which had previously been agitated by wonder at the marvels of the world[48].

> "Smith's doctrine ... measures the value of philosophical systems solely in relation to their satisfaction of the human craving for order ... it leaves all science essentially hypothetical." (Richard Olson (1975) *Scottish Philosophy and British Physics, 1750-1880* p 123, cited in Raphael and Skinner, 1980: 12)

---

[46] Mirowski is wrong to imply, as he does here, that 'wonder' is something desirable. On the contrary, for Smith, wonder was a dis-ease of the imagination, caused by the incoherent appearances of nature, the purpose of science being, precisely, to return one's imagination to its ease.

[47] It should be noted that no clear distinction between 'science' and 'philosophy' existed at this time, '[i]n fact the terms philosophy, physics, arts, sciences, and natural philosophy are used almost indiscriminately' in Smith (Wightman, 1980, 1982: 12).

[48] As we shall see in the next section, Smith held a *harmonious* view of nature. He would have denied that that there was any inconsistency between investigating the world and accepting it: investigation would simply reveal harmony. The confidence that reason would confirm the prior wisdom of religion and sentiment permitted Smith, like many philosophers and scientists of the period, from Newton to Hegel, to make genuine discoveries. This ambivalence as to the status of reason leads to a fundamental ambiguity in Smith (as also, for example, in Hegel): 'much in Smith ... is Janus-faced' (Mirowski, 1989: 168). I submit, nevertheless, that reading works such as the **Astronomy** allows us to see that it was always reconciliation with the natural and social order that was uppermost in Smith's mind.

At the level of appearances, Smith says, the world throws up phenomena which appear incoherent and therefore inflame the imagination. This inflammation is to be regarded as a disagreeable sensation. 'When we first encounter anything that is nor familiar or expected, Smith argues, we are struck by the feelings we call Surprise and Wonder. These are not welcome feelings.' (Heilbroner, 1986: 15) The job of a science is to soothe the imagination by suggesting connections between things, and by tracing the unknown back to the familiar, so that the observer may regain his tranquillity:

> "Philosophy is the science of the connecting principles of nature. Nature ... seems to abound with events which appear solitary and incoherent ... which therefore disturb the easy movement of the imagination .... Philosophy, by representing the invisible chains which bind together all these disjointed objects, endeavours to introduce order into this chaos of jarring and discordant appearances, to allay this tumult of the imagination, and to restore it ... to [its former] tone of tranquillity and composure ... Philosophy, therefore, may be regarded as one of those arts which address themselves to the imagination". (**Astronomy** II.12)

Or, more pithily, 'it is the end of Philosophy, to allay that wonder, which either the unusual or seemingly disjointed appearances of nature excite' (**Astronomy** IV.34). We do not understand what we seek to explain by science, 'but by categorizing things we come to be at peace with them … We draw the venom of Wonder by applying the poultice of familiarity' (Heilbroner, 1986: 16).

For Smith, therefore, it is just irrelevant to talk about the truth or otherwise of the findings of a science – what matters is its success or otherwise in 'smoothing the passage of the imagination betwixt … seemingly disjointed objects' (ibid), it is this criterion alone which we should bear in mind when considering the sequence of schools of thought in a science such as astronomy:

> "Let us examine, therefore, all the different systems of nature, which ... have successively been adopted by the learned and ingenious; and, *without regarding their absurdity or probability, their agreement or inconsistency with truth and reality*, let us consider them only in that particular point of view which belongs to our subject; and content ourselves with inquiring how far each of them was fitted to sooth the imagination, and to render the theatre of nature a more coherent ... spectacle". (ibid, emphasis added)

Indeed, Smith concludes his discussion of Newton's system of astronomy by confessing that it is so compelling that he had, himself, been seduced into speaking of the latter's system as if it embodied real knowledge of the world:

> "even we, while we have been endeavouring to represent all philosophical systems as mere inventions of the imagination, to connect together the otherwise disjointed and discordant phaenomena of nature, have insensibly been drawn in, to make use of language expressing the connecting principles of this one [sc Newton's philosophical system], as if they were the real chains which Nature makes use of to bind together her several operations." (**Astronomy** IV.76)

And this is a measure of the success of Newton's system. The implication is, as Raphael and Skinner (1980: 19-21) point out, that it would be mistaken, or at best off the point, to regard Newton's connecting principles as 'the real chains' of Nature.

There is a further point to be made about Smith's methodology here. For him, science starts off, as indeed all science must, with the level of appearances: but then, instead of penetrating those appearances to reality, the truth, to the essence of the thing, science *remains* at the level of appearances, merely contrasting one set of appearances with another. In place of a congeries of apparently incoherent, isolated phenomena, Smithian science gives us a coherent and interconnected vision of the world[49]. But, for Smith, that vision is no more real, no less apparent than either the raw appearances or the connecting principles proposed by rival explanations. The criterion for choosing between these appearances is not their greater or lesser degree of truth, but a purely *aesthetic* consideration: which is the more pleasing? Thus a scientific explanation of a phenomenon is to be preferred to none, and a later system is preferred to an earlier one, because and to the extent to which they are able to provoke greater admiration (**Astronomy** II.12). Though much to be preferred to the earlier systems, there is no suggestion – the idea is without interest to Smith – that the Newtonian system is more *profound*, indeed, it may well be replaced when an even more pleasing system is proposed. 'Philosophy' is to be traced, he says, 'from its origin, up to that summit of perfection to which it is at present supposed to have arrived [with Newton], and to which, indeed, it has equally been supposed to have arrived in almost all former times' (ibid). In every period, Smith says, science is believed to have reached 'the summit of perfection' as the science of that period is just the scientific explanation the period finds most pleasing. Whether there is any *progress* in this is left moot.

[49] Doing so is already a step towards penetrating appearances to the reality hidden behind them, but we are here concerned with Smith's objectives, not with the discoveries he made despite them.

It has been suggested[50] that Smith denies that there is any such thing as the truth, an objective reality to the world apart from the models and images of it which we construct. This in turn, it is argued, is a very modern view of the world, popular, for example among some twentieth century physicists. His disdain for the truth has also been linked to Hume's alleged scepticism. This is a misinterpretation both of Smith and Hume[51]. For Smith, there is indeed objective truth, but human, finite minds cannot grasp, or even approach it: only the infinite mind of God can grasp all the ultimate 'connexions and dependencies of things'. There is thus an unbridgeable gulf between the finite and the infinite, between the human and the divine. This was a very common medieval view of the nature of infinity; see Rucker (1995: 4), for example, for a discussion of this point in Thomas Aquinas. This contrast will be touched on in the discussion of **TMS** below; indeed, it forms the basis for the very restricted role of reason and philosophy (the sphere of finitude), relative to that of sentiment and religion (the sphere of infinity), in Smith's system.

Unsurprisingly, perhaps, Smith's editors find some difficulty in dealing with this aspect of his thought:

> "When Smith writes that scientists have imagined inventions he does not say that they have invented science fiction – or any other sort of fiction. But he does contrast an invention by the imagination with a discovery of the truth, and so he implies that scientific theory cannot be true." (Raphael and Skinner, 1980: 21)

Raphael and Skinner make two claims here: (a) that Smith implies that scientific theory cannot be true, and (b) that Smith does not say that scientists are writing science fiction. In my opinion these two claims are incompatible – if what scientists say 'cannot be true', then it must, surely, be fiction. But aside from that, the second claim is actually false. Smith did not, of course, use the actual term 'science fiction', but he did call Descartes' work, for example, 'an entertaining romance'[52] – and he meant no sneer by this, for he admired Descartes greatly[53]:

---

[50] By various correspondents in response to earlier versions of this chapter.

[51] This is not the place for a detailed consideration of Hume. In his discussion of Smith's 'theistic' invisible hand theory, Macfie (1967: 108) says 'Smith … is stating a form of faith which was general in his school', and notes that 'Hume's scepticism is mainly as to possible proof' of that faith, not scepticism as to the validity of the doctrines of that faith.

[52] In a similar vein, David Hume referred to history as an 'agreeable entertainment' (cited in Marwick,

> "We need not be surprised ... that the Cartesian philosophy ... though it does not perhaps contain a word of truth ... should nevertheless have been so universally received by all the learned in Europe at that time. The great superiority of [Descartes'] method ... made them greedily receive a work which we justly esteem one of the most entertaining romances that have ever been wrote" (Adam Smith *Lectures in Rhetoric and Belles Lettres* (1748-1750), cited in editorial footnote 3, **EPS**: 244)

Again, later, in his *Letter to the 'Edinburgh Review'* (1756: §5), he says of the Cartesian philosophy, that 'in the simplicity, precision and perspicuity of its principles *and conclusions*, it had the same superiority over the Peripatetic system, [as] the Newtonian philosophy [had]' (**EPS** p244, emphasis added).

Perhaps anticipating some 'postmodern' writers by more than two centuries, therefore, Smith regarded the *narrative* as the appropriate focus for the attention of an investigating philosopher. Although completely untrue, a romance, the principles and conclusions of Descartes' narrative are to be regarded as as much an improvement over previous approaches as Newton's is, because it provides simple, precise and perspicuous ... entertainment. Descartes' vortices reconcile us to our world, even though 'these pretended causes of those wonderful effects, not only do not actually exist, but are utterly impossible, and if they did exist, could produce no such effects as are ascribed to them' (**TMS** VII.ii.4.14).

Contra Raphael and Skinner, 'science fiction' is exactly what Smith thinks scientists are producing, but science fiction with a particular slant, science fiction with the purpose of soothing the imagination and reconciling us to the world about us. As Mirowski points out, the primary function of science for Smith is the evocation of order.

This section has set out the main lines of Smith's methodological stance and suggested links between his methodology and his underlying intellectual goals. The purpose of doing so is to illuminate his political economy and, in particular, his notion of 'the invisible hand'. It is standard amongst economists today to believe that philosophy is redundant, and, in particular, that Smith's non-economic writings can tell us nothing

---

1970: 31)

[53] See Raphael and Skinner (1980: 1 n 2); also Mirowski (1989: 163 ff): 'Smith ... has not been given adequate credit as the prime suspect in the smuggling of Cartesian economics into the backyard of

about his economics. The purpose of this chapter is to make the contrary case. Smith's writings on methodology set out a research programme which Smith then followed in his psychological (**TMS**) and economic (**WN)** investigations: 'the philosopher who began the essay on the history of astronomy with a theory of scientific systems is himself applying that theory in his construction of an economic system' (Raphael, 1985: 77).

4.4 Smith's *Weltanschauung*

4.4.1 *All is for the best in this world and we should accept our lot with joy*

Smith's Weltanschauung is adopted, with minor modifications, from the Stoics[54]; the points where Smith does, and does not, agree with the Stoics, are not, however, germane to the theme presented here[55]. Smith believes that the universe, or Nature, is an enormous, sophisticated and subtle machine. This machine is supervised by an omnipotent, omniscient and beneficent, indeed, a utilitarian[56], deity. The sole aim of the machine (and, probably, of the deity himself, see **TMS** VII.ii.3.18), is the maximisation of happiness:

> "all the inhabitants of the universe, the meanest as well as the greatest, are under the immediate care and protection of that great, benevolent, and all-wise being, who directs all the movements of nature; and who is determined, by his own unalterable perfections, to maintain in it, at all times, the greatest possible quantity of happiness." (**TMS** VI.ii.3.1)

Newton' (ibid: 164).

[54] This assertion has been questioned by some correspondents; nevertheless, I think it clear that Stoicism constituted a critically important element in Smith's thinking: 'Stoic philosophy is the primary influence on Smith's ethical thought. It also fundamentally affects his economic theory .... Stoicism never lost its hold over Smith's mind' (Raphael and Macfie, 1976: 5-6). The argument is also set out in Clarke (1996, 1998). In any case, acceptance of this point is not a precondition for understanding or accepting the argument presented in the remainder of the present chapter.

[55] See Macfie (1959: 225) for some of the ways Smith modifies Stoic doctrine.

[56] Smith himself is certainly not a utilitarian. That would require him to hold a consequentialist view of morality rather than the deontological view he actually does hold. See Sen and Williams (1982: 3-4) for the argument that utilitarianism lies at the intersection of welfarism and consequentialism. It would also require him to believe, what he does not believe, that it is possible for human actions and institutions to increase the total quantity of happiness in the world. The deity, however, is another matter. For more on the relation between Smith and utilitarianism, see Macfie's review of Lionel Robbins's *Theory of Economic Policy* (Macfie, 1967: 152-161).

> "[T]hat divine Being['s] ... benevolence and wisdom have, from all eternity, contrived and conducted the immense machine of the universe, so as at all times to produce the greatest possible quantity of happiness". (**TMS** VI.ii.3.5)

So the world is perfect: we *do* live in the 'best of all possible worlds' – Smith is a true Panglossian. Since the world is really perfect, our apparent troubles stem from our finite, partial view of the world. The purpose of philosophy, therefore, is to cultivate a fine indifference to whatever occurs:

> "The wise and virtuous man is at all times willing that his own private interest should be sacrificed to the public interest of his own particular order or society. He is at all times willing, too, that the interest of this order or society should be sacrificed to the greater interest of the state or sovereignty, of which it is only a subordinate part. He should, therefore, be equally willing that all those inferior interests should be sacrificed to the greater interest of the universe, to the interest of that great society of all sensible and intelligent beings, of which God himself is the immediate administrator and director .... [Since the] benevolent and all-wise Being can admit into the system of his government, no partial evil which is not necessary for the universal good, he [sc the wise and virtuous man] must consider all the misfortunes which may befal himself, his friends, his society, or his country, as necessary for the prosperity of the universe, and therefore as what he ought, not only to submit to with resignation, but as what he himself, if he had known all the connexions and dependencies of things, ought sincerely and devoutly to have wished for." (**TMS** VI.ii.3.3)

Smith sustains this theme by making use of the analogy of soldiers marching cheerfully off to be slaughtered in defence of a 'forlorn station', that is, an indefensible position:

> "They cheerfully sacrifice their own little systems to the prosperity of a greater system .... No conductor of an army can deserve more unlimited trust, more ardent and zealous affection, than the great Conductor of the universe. In the greatest public as well as private disasters, a wise man ought to consider that he himself, his friends and countrymen, have only been ordered upon the forlorn station of the universe; that had it not been necessary for the good of the whole, they would not have been so ordered; and that it is their duty, not only with humble resignation to submit to this allotment, but to endeavour to embrace it with alacrity and joy." (ibid VI.ii.3.4)

The message is clear: what is good is good and what is bad is good as well; everything is for the best, so – whatever happens – rejoice, and accept. Lest the reader should be tempted to wonder whether these passages represent a stage in Smith's thought long passed by the time he came to write **WN**, I should point out that, though similar ideas can be found in the earlier editions, these passages themselves are taken from Part VI, a new section written by Smith, in the last year of his life, for the 1790 edition.

Smith has another tactic for convincing us that all is for the best. His first move is to say that what appears bad is actually good, but we don't see it because we are only finite

minds. His second tactic, tacitly admitting, perhaps, that the first won't wash, is to introduce an afterlife to balance the books. All our virtue and vice will be appropriately rewarded, if not here, then hereafter:

> "Our happiness in this life is ... upon many occasions, dependent on the humble hope and expectation of a life to come: a hope and expectation deeply rooted in human nature .... [T]here is a world to come, where exact justice will be done to every man ... " (**TMS** III.2.33) "Nature teaches us to hope, and religion, we suppose, authorises us to expect, that it [sc injustice] will be punished, even in a life to come. Our sense of its ill desert pursues it ... even beyond the grave .... The justice of God, however, we think, still requires, that he should hereafter avenge the injuries of the widow and the fatherless, who are here so often insulted with impunity." (**TMS** II.ii.3.12)

Smith combines the idea of justice in the hereafter with that of the limits to reason and the scope for religion and sentiment, which we will examine in more detail below. To those such as the wrongly condemned man, Smith says,

> "humble philosophy which confines its views to this life, can afford, perhaps, but little consolation .... Religion can alone afford them any effectual comfort. She alone can tell them, that it is of little importance what man may think of their conduct, while the all-seeing Judge of the world approves of it. She alone can present to them the view of another world ... where their innocence is in due time to be declared, and their virtue to be finally rewarded" (**TMS** III.2.12).

Indeed, we are not only led to a belief in a life after death by our religious sentiments, but by an intellectual consideration of the idea of justice, itself:

> "When we thus despair of finding any force upon earth which can check the triumph of injustice, we naturally appeal to heaven, and hope, that the great Author of our nature will himself execute hereafter, what all the principles which he has given us for the direction of our conduct, prompt us to attempt even here; that he will complete the plan which he himself has thus taught us to begin; and will, in a life to come, render to everyone according to the works which he has performed in this world. And thus we are led to the belief of a future state, not only by the weakness, by the hopes and fears of human nature, but by the noblest and best principles which belong to it, by the love of virtue, and by the abhorrence of vice and injustice." (**TMS** III.5.10)

4.4.2 *Why, then, bother with considerations of morality?*

The idea that things seem good or bad to us only because of our limited perspective, and that, if we knew 'all the connexions and dependencies of things', we would realise that everything is good, is extremely important in Smith. However, it does raise the question of why we should then be concerned as to the moral qualities of our behaviour. The argument proceeds in three steps. Firstly, in answer to the question, Why do we approve

of moral actions and disapprove of immoral actions?, Smith says that our moral response to an action is a *sentimental reaction*, that is, it is produced by our instinctive feelings, in particular, the emotion of *sympathy*. By sympathy we enter, to a limited degree, into the feelings of those affected by the action in question, the victims or beneficiaries:

> "our sense of the merit of good actions is founded upon a sympathy with the gratitude of the persons who receive the benefit of them ... Gratitude and resentment ... are ... counterparts to one another; and if our sense of merit arises from a sympathy with the one, our sense of demerit [must] ... proceed from a fellow-feeling with the other." (**TMS** II.i.5.7)

These emotions are placed within us by the deity as part of the grand design. Our instinctive response to murder, for example, is directly implanted in us by a Nature seen as an active and conscious principle in the world:

> "with regard ... to this most dreadful of all crimes, Nature ... has ... stamped upon the human heart ... an immediate and instinctive approbation of the sacred and necessary law of retaliation." (**TMS** II.i.2.5) "As every man doth, so shall it be done to him ... retaliation seems to be the great law which is dictated to us by Nature." (**TMS** II.ii.1.10)

Secondly, however, Smith tells us that the wise man will recognise that whatever happens to him is for the best, and that however unpleasant it appears, that is only because we as limited beings do not see the distant, but only the proximate, consequences and ramifications of actions. Hence, morality is based only on a consideration of the proximate consequences of the action whose morality we are to appraise. If the first round effects are unjustly detrimental to some person or group, the subsequent ramifications will prove beneficial to themselves or others to a degree that more than counterbalances the evil done at first. Nevertheless, our appraisal of the morality of the action will focus on its first round effects alone and condemn it: 'The ancient stoics were of the opinion,' he reports, with approval,

> "that as the world was governed by the all-ruling providence of a wise, powerful and good God, every single event ought to be regarded, as making a necessary part of the plan of the universe, and as tending to promote the general order and happiness of the whole: that the vices and follies of mankind, therefore, made as necessary a part of this plan as their wisdom or their virtue; and by that eternal art which educes good from ill, were made to tend equally to the prosperity and perfection of the great system of nature. No speculation of this kind, however, how deeply soever it might be rooted in the mind, could diminish our natural abhorrence for vice, whose immediate effects are so destructive, and whose remote ones are too distant to be traced by the imagination." (**TMS** I.ii.3.4)

Thus Smith argues that our abhorrence of vice is due to our failure to follow through all the ramifications of an immoral act. If we were to do so, he implies, we would accept vice with equanimity as generating remote positive effects which at least outweigh the proximate negative ones. Reason, or 'speculation', however, cannot change our basic emotional instincts.

The third step is thus to find a way to endorse the morality of a moral action. Why should we bother to make the distinction between moral and immoral actions if their effects are the same? Smith has no clear answer to this; he does, however, have two unclear answers. The first approach is to duck the issue and say that this is a positive, not a normative science, that is moral which is considered moral:

> "the present inquiry is not concerning a matter of right ... but a matter of fact. We are not at present examining upon what principles a perfect being would approve of the punishment of bad actions; but upon what principles so weak and imperfect a creature as man actually and in fact approves of it." (**TMS** II.i.5.10)

Essentially, what distinguishes the moral from the immoral here is an aesthetic matter: it is a question of what *feels* better, even though reason can make no distinction. Perhaps Smith says more than he intends to here for there is a clear logical implication in the contrast he employs: although we, imperfect creatures, may regard this action as moral and that as immoral, a perfect being, conscious of all their most distant consequences, would see matters in quite another light and, presumably, would not make this distinction. That Smith could not accept this implication is shown by his alternative response. For the second approach is, precisely, to argue that by acting morally we place ourselves on the same side, as it were, as the deity:

> "by acting according to the dictates of our moral faculties, we necessarily pursue the most effectual means for promoting the happiness of mankind, and may therefore be said, in some sense, to co-operate with the Deity, and to advance, as far as in our power, the plan of Providence. By acting otherways, on the contrary, we seem to obstruct, in some measure, the scheme which the Author of nature has established for the happiness and perfection of the world, and to declare ourselves ... in some measure the enemies of God." (**TMS** III.5.7)

True, this is not very logical, for it evades the question, how we could conceivably displease a god by our choice of action, given that, according to Smith, he is in a position to dictate exactly the mixture of feelings, the strengths and weaknesses and so on, making up each personality, and hence the behaviour to which each person will be led.

Again, it seems inconsistent to speak of more and less effectual means for promoting the happiness of humanity when God has already determined to maximise the quantity of happiness in the world; the means chosen are presumably those chosen by him, and hence are 'necessarily perfect'. Essentially, therefore, Smith's endorsement of morality is *sentimental*, based on its feeling of rightness, and revealed divine sanction, without rational justification. Morality and virtue are essentially *ornaments*, having no rational function: 'those sentiments and qualities which are the great ornaments of humanity, and which seem to raise it to a resemblance of divine perfection, [are] the love of virtue and beneficence, and the abhorrence of vice and injustice' (**TMS** III.5.4).

In spite of his harmonious view of nature, Smith senses that there is a tension between reason and sentiment, between the logic of his position and his religious feelings. Does God love virtue and hate vice in view of their consequences, or, like us, 'for their own sakes'? Characteristically, he comes down on the side of sentiment:

> "That the Deity loves virtue and hates vice ... not for their own sakes, but for the effects which they tend to produce; that he loves the one, only because it promotes the happiness of society, which his benevolence prompts him to desire; and that he hates the other, only because it occasions the misery of mankind, which the same divine quality renders the object of his aversion; is not the doctrine of nature, but of an artificial ... refinement of philosophy. All our natural sentiments prompt us to believe, that as perfect virtue is supposed necessarily to appear to the Deity, as it does to us, for its own sake, and without further view, the natural and proper object of love and reward, so must vice, of hatred and punishment." (**TMS** II.ii.3.13, editions 1-5, omitted in 6.)

Smith cannot logically say that virtue is preferable to vice *because* it leads to the happiness of society[57], as this would be inconsistent with his claim that God has in any case arranged everything to maximise happiness in the world at every instant. But it is equally quite illogical for him to say that virtue appears to God as it does to us – for the only reason it appears thus to us, according to Smith, is that that is how God has made us, in order that we may play our predestined part in the great plan. But there is no higher spirit to place such feelings in God's nature, nor any higher plan which he is supposed to play a subordinate part in.

* * *

[57] Although virtue does have precisely this 'utility', particularly in relation to justice: see the next

To conclude, the message of the present subsection is that, for Smith, morality has an aesthetic basis. For example, Smith lists as one of the four sources of approbation, that

> "when we consider [ethical] actions as making a part of a system of behaviour which tends to promote the happiness either of the individual or of the society, they appear to derive a beauty from this utility, not unlike that which we ascribe to any well-contrived machine." (**TMS** VII.iii.3.16)

Macfie is at pains to point out that beauty is more important to Smith than usefulness, indeed, utility seems only to be considered to the extent that it, itself, entails beauty. In Smith, Macfie says,

> "a sense of propriety is something different from good consequences in some sense pleasant. Here, indeed, Smith is consistent and specific. It is to the 'beauty' of 'fitness' that he constantly returns – the beauty of the 'well-contrived machine.' Such a machine, says Smith, provides 'a thousand agreeable effects,' while a rusty jarring machine 'would displease' and be 'necessarily offensive.' So what is valuable from the point of view of virtue is the fitness of the machine, the 'fine polish' it gives 'to the wheels of society,' not the results for which the machine is the necessary means." (Macfie, 1961: 15. Macfie's references to Smith are from **TMS** VII.iii.1.2)

To put it in modern terms, Macfie says, 'there is no doubt that for him [sc Smith] it is beauty that has final value, utility instrumental value (apart from its own inherent beauty or fitness)' (Macfie, 1961: 16 n 13).

4.4.3 *Every cloud has a silver lining*

I said in the previous subsection that, according to Smith, God was in a position to choose the mental composition of individual persons, and hence to lead them to desirable behaviours. We now need further to investigate this theme, epitomised in Smith's view that 'That [sc God's] wisdom ... contrived the system of human affections, as well as that of every other part of nature' (**TMS** VI.ii.2.4).

This is important for a number of reasons. Firstly, I need to justify this claim of predestination in Smith. Secondly, Smith's argument here is a further illustration of his Panglossian view that everything is predetermined by the deity, predestined to turn out for the best. And, thirdly, because, again, these arguments further illustrate Smith's view,

---

subsection of this Chapter, and **TMS** II.ii.3.1ff.

mentioned before, that if we are misled by appearances, then this deception, too, is part of the plan and hence a Good Thing.

A major instance concerns the predisposition to benevolence and the very much stronger one, not just to obey, but to enforce, the 'sacred laws of justice' (**TMS** II.ii.2.3), which God has placed in our personal make-up, what Smith calls 'this constitution of Nature' (**TMS** II.ii.3 title). Man, he says, 'who can only subsist in society, was fitted by nature to that situation for which he was made' (**TMS** II.ii.3.1). While it would be nice if everyone could cooperate from sheer love of one's fellows (ibid), we can still live without society-wide benevolence; but not without justice: 'Society may subsist, among different men, as among different merchants, from a sense of its utility, without any mutual love or affection ... but the prevalence of injustice must utterly destroy it' (**TMS** II.ii.3.2-3).

Nature has therefore endowed men with consciences in order that they may behave justly:

> "Though Nature, therefore, exhorts mankind to acts of beneficence, by the pleasing consciousness of deserved reward, she has not thought it necessary to guard and enforce the practice of it by the terrors of merited punishment in case it should be neglected. It is the ornament which embellishes, not the foundation which supports the building, and which it was, therefore, sufficient to recommend, but by no means necessary to impose. Justice, on the contrary, is the main pillar that upholds the whole edifice. If it is removed, the great, the immense fabric of human society, that fabric which to raise and support seems in this world ... to have been the peculiar and darling care of Nature, must in a moment crumble into atoms. In order to enforce the observation of justice, therefore, Nature has implanted in the human breast that consciousness of ill-desert, those terrors of merited punishment which attend upon its violation, as the great safe-guards of the association of mankind, to protect the weak, to curb the violent, and to chastise the guilty." (**TMS** II.ii.3.4)

It is clear that Smith is saying here that Nature, in order to preserve society, has placed in our personalities a desire for justice, even if it is unclear whether this is based on a love of justice for its own sake, or a fear of retribution. A sense of justice is an endowment of nature, but nature seen as an active force in the world, conscious and intentional. Speaking of **TMS**, Heilbroner says

> "But whence come these higher principles of nature? ... What is necessary is that we assume human nature to contain such a saving element. The imperatives of duty and the voice of conscience must be there from the start, available to us in critical situations. They

must be part of the human makeup, placed there by the Deity that has arranged for our collective well-being." (Heilbroner, 1986: 60)

Despite Smith's claim that justice is fundamental for society, *order* is in reality of more basic importance to him. Speaking of the tendency for members of the different 'orders and societies' in the state to resist any diminution in their 'powers, privileges and immunities', he argues that

> "This partiality, though it may sometimes be unjust, may not, upon that account be useless. It checks the spirit of innovation. It tends to preserve whatever is the established balance among the different orders and societies into which the state is divided; and while it sometimes appears to obstruct some alterations in government which may be fashionable and popular at the time, it contributes in reality to the stability and permanency of the whole system." (**TMS** VI.ii.2.10)

The assumption is that what is, is likely to be best, and should in general be preserved, even at the expense of justice. Having said that, however, we should note that, for Smith, just as there can be no profound antagonism between investigation and reconciliation, there cannot be any serious conflict between order and justice. Indeed, everyone always gets their just deserts in the end:

> "notwithstanding the disorder in which all things appear to be in this world, yet even here every virtue naturally meets with its proper reward, with the recompense which is most fit to encourage and promote it; and this too so surely, that it requires a very extraordinary concurrence of circumstances entirely to disappoint it." (**TMS** III.5.8)

And should such extraordinary circumstances occur, everything can be set to rights, and the books balanced, as we have already seen, in the hereafter. It is precisely this concept of heaven as a mechanism for balancing the books that allows Smith to defend principles, such as the partiality of the orders of society in defence of their own interests, and the contempt 'unjustly' bestowed upon poverty and weakness instead of on vice and folly (**TMS** II.ii.3.4), when they conflict with the claims of justice.

When Smith speaks of justice he is thinking of order, when he talks of order he is thinking of property:

> "The poor man must neither defraud nor steal from the rich, though the acquisition might be much more beneficial to the one than the loss could be hurtful to the other .... by [doing so] he renders himself the proper object of the contempt and indignation of mankind; as well as of the punishment which that contempt and indignation must naturally dispose them to inflict, for having thus violated one of those sacred rules, upon the tolerable

> observation of which depend the whole security and peace of human society. There is no commonly honest man who does not more dread the inward disgrace of such an action, the indelible stain which it would for ever stamp upon his own mind, than the greatest external calamity which, without any fault of his own, could possibly befal him; and who does not inwardly feel [that such an action] is more contrary to nature, than death, than poverty, than pain, than all the misfortunes which can affect him" (**TMS** III.3.6).

Thus theft by the poor from the rich – even when, as he concedes, it would augment social welfare – calls down more Smithian abuse upon their heads than any other crime. In one passage a murderer or parricide, by contrast, is dismissed as merely 'ungrateful' (**TMS** II.ii.3.11), while in another, murder, though stigmatised as 'this most dreadful of all crimes' (**TMS** II.i.2.5), is dealt with matter-of-factly without any of the excitement shown in his discussion of theft from the rich. Again, it is well known that Smith regarded the state as an institution guarding the rich from the poor:

> "Till there be property there can be no government, the very end of which is to secure wealth, and to defend the rich from the poor." (Smith *Lectures on Jurisprudence*, cited in **WN** V.i.b.12 n21) "Civil government, so far as it is instituted for the security of property, is in reality instituted for the defence of the rich against the poor, or of those who have some property against those who have none at all." (**WN** V.i.b.12)

This fact, however, has been subjected to the almost comical misinterpretation that somehow this represented a *complaint*, a plea on behalf of the underdog. Viner (1958: 233), for example, cites these passages as evidence for Smith's desire to limit government activity, and Raphael (1985: 8) says that the **WN** passage strikes a 'radical note'. Nothing could be further from the truth. The context of these passages shows unambiguously that Smith was simply, and, in his view, uncontroversially, setting out how things were and how they should be:

> "The affluence of the rich excites the indignation of the poor, who are … prompted by envy to invade his possessions … which [are] acquired by the labour of many years, or perhaps of many successive generations …. He is at all times surrounded by unknown enemies, whom, though he never provoked, he can never appease, and from whose injustice he can be protected only by the powerful arm of the civil magistrate continually held up to chastise it [sc the injustice of those enemies]. The acquisition of valuable and extensive property, therefore, necessarily requires the establishment of civil government."

Robert Heilbroner gets closer to the real meaning of these passages:

> "We come to Smith correctly expecting to find a great social thinker in the conservative tradition, but we are not likely to anticipate finding in him that 'Laws and government

> may be considered … as a combination of the rich to oppress the poor'[58] … Smith could speak in these seemingly radical … terms because neither he nor any of his contemporaries imagined a society in which exploitation and oppression would not be present" (Heilbroner, 1986: 3).

* * *

> "In the *Theory of Moral Sentiments*, Smith develops the doctrine of a beneficent order in nature, manifesting itself through the operation of the forces of external nature and the innate propensities implanted in man by nature …. [T]he essence of Smith's doctrine is that Providence has so fashioned the constitution of external nature as to make its processes favourable to man, and has implanted *ab initio* in human nature such sentiments as would bring about, through their ordinary working, the happiness and welfare of mankind." (Viner, 1958: 216-17)

Our strengths are thus implanted in us by divine providence. Not only our strengths but our weaknesses, too, however, are endowed by nature. A particularly striking example concerns the tendency of a fickle public to admire people merely for being lucky:

> "Fortune has ... great influence over the moral sentiments of mankind, and, according as she is either favourable or adverse, can render the same character the object, either of general love and admiration, or of universal hatred and contempt. This great disorder in our moral sentiments is by no means, however, without its utility; and we may on this as well as on many other occasions, admire the wisdom of God even in the weakness and folly of man. Our admiration of success is founded upon the same principle with our respect for wealth and greatness, and is equally necessary for establishing the distinction of ranks and the order of society. By this admiration of success we are taught to submit more easily to those superiors, whom the course of human affairs may assign to us; to regard with reverence, and sometimes even with a sort of respectful affection, that fortunate violence which we are no longer capable of resisting; not only the violence of such splendid characters as those of a Caesar or an Alexander, but often that of the most brutal and savage barbarians, of an Attila, a Gengis, or a Tamerlane." (**TMS** VI.iii.30)

This is a remarkable passage. Admiration for the merely lucky is, admittedly, a 'great disorder' in our morals. But even our folly reflects God's wisdom, and this particular folly, like everything else, has been given us by God for a reason. The good thing about this weakness is that it reconciles us with our rulers, even those who only achieved this status by means of 'fortunate violence', inspiring us even to a kind of affection for brutal tyrants such as Tamerlane (or Timur Lenk), who reputedly made mountains of his enemies' skulls.

As Smith reminds us, this view of the role of fortune in moral sentiments parallels that of public admiration of the great in preference to the good:

[58] The reference is to Smith (1976) *Lectures on Jurisprudence* Oxford: Clarendon Press, p208.

> "This disposition to admire, and almost to worship, the rich and the powerful, and to despise, or, at least, to neglect persons of poor and mean condition ... is ... the great and most universal cause of the corruption of our moral sentiments. That wealth and greatness are often regarded with the respect and admiration which are due only to wisdom and virtue; and that the contempt, of which vice and folly are the only proper objects, is often most unjustly bestowed upon poverty and weakness, has been the complaint of moralists in all ages." (**TMS** I.iii.3.1)

And they were wrong – in Smith's view – since, as we have seen, even injustice can be part of a higher Good. Smith, himself, incidentally, was happy to contribute to this contempt for the poor (though the case of the rich who become poor was another matter altogether):

> "The mere want of fortune, mere poverty, excites little compassion. Its complaints are too apt to be the objects rather of contempt than of fellow-feeling. We despise a beggar; and ... he is scarce ever the object of any serious commiseration. The fall from riches to poverty, as it commonly occasions the most real distress to the sufferer, so it seldom fails to excite the most sincere commiseration in the spectator." (**TMS** III.3.18)

Even this 'universal cause of moral corruption', however, is god-given and has its purpose: it is 'necessary both to establish and maintain the distinction of ranks and the order of society' (**TMS** I.iii.3.1).

> "The distinction of ranks, the peace and order of society, are, in a great measure, founded upon the respect which we naturally conceive for [the greatly fortunate ... the rich and powerful] .... The peace and order of society is of more importance than even the relief of the miserable .... Moralists ... warn us against the fascination of greatness. This fascination, indeed, is so powerful, that the rich and the great are too often preferred to the wise and the virtuous. Nature has wisely judged that the distinction of ranks, the peace and order of society would rest more securely upon the plain and palpable difference of birth and fortune, than upon the invisible and often uncertain difference of wisdom and virtue. The undistinguishing eyes of the great mob of mankind can well enough perceive the former: it is with difficulty that the nice discernment of the wise and the virtuous can sometimes distinguish the latter. In the order of all those recommendations, the benevolent wisdom of nature is equally evident." (**TMS** VI.ii.1.20)

So even this particular weakness, which Smith has earlier damned in the most severe terms, is evidence of the 'benevolent wisdom of nature', and it is so because there has to be a ruling stratum, and Nature has judged it best to have an obvious one to which the masses can easily be led to give their loyalty.

There is a further point concerning the admiration of wealth which illustrates Smith's view that deceptive appearances can still be desirable. For Smith, the outward

appearance of great disparity in wealth between the rich and the poor conceals a very large measure of real equality in welfare. In the passage from **TMS** (IV.1.10) cited at the beginning of this chapter, Smith argued that 'In what constitutes the real happiness of human life, they [sc the poor] are in no respect inferior to those who would seem so much above them.' So the sources of real happiness were divided by divine providence – or by the rich who are, in turn, led by divine providence – so that we all get an equal share. This theme is repeated throughout Smith's works, often combined with the notion that great happiness and grief are occasioned not by a state or condition but by a change in condition.

> "The never-failing certainty with which all men, sooner or later, accommodate themselves to whatever becomes their permanent situation, may, perhaps, induce us to think that the Stoics were, at least, thus far very nearly in the right; that, between one permanent situation and another, there was, with regard to real happiness, no essential difference .... Happiness consists in tranquillity and enjoyment. Without tranquillity there can be no enjoyment; and where there is perfect tranquillity there is scarce any thing which is not capable of amusing. But in every permanent situation, where there is no expectation of change, the mind of every man, in a longer or shorter time, returns to its natural and usual state of tranquillity." (**TMS** III.3.30)

He illustrates the point with an anecdote about an imprisoned count who amused himself 'with feeding a spider' (ibid)! The view that all permanent conditions are alike, and that it is only changes which matter, is reflected in his statement, reported earlier, that the poor are, at best, ignored, while the impoverished rich are pitied. Smith draws the conclusion that much of the evil in life can be attributed to failure to understand this point:

> "The great source of both the misery and disorders of human life, seems to arise from over-rating the difference between one permanent situation and another. Avarice over-rates the difference between poverty and riches .... The person under the influence of those extravagant passions [sc avarice], is not only miserable in his actual situation, but is often disposed to disturb the peace of society, in order to arrive at that which he so foolishly admires ... [although] in all the ordinary situations of human life, a well-disposed mind may be equally ... contented .... In all the most glittering and exalted situation that our idle fancy can hold out to us, the pleasures from which we derive our happiness, are almost the same with those which, in our actual, though humble station, we have at all times at hand, and in our power." (**TMS** III.3.31)

But in even *this* cloud there is a silver lining! It is in the extremity, or extravagance, of the emotion that the problem lies. Merely to be deceived by appearances, on the contrary, is often desirable:

> "The poor man's son, whom heaven in its anger has visited with ambition ... admires the condition of the rich .... He is enchanted with the distant idea of this felicity .... and, in order to arrive at it, he devotes himself for ever to the pursuit of wealth and greatness .... Through the whole of his life he pursues the idea of a certain artificial and elegant repose which he may never arrive at, for which he sacrifices a real tranquillity, that is at all times in his power, and which, if in the extremity of old age he should at last attain to it, he will find to be in no respect preferable to that humble security and contentment which he had abandoned for it. It is then ... that he begins at last to find that wealth and greatness are mere trinkets of frivolous utility .... And it is well that nature imposes upon us in this manner. It is this deception which rouses and keeps in motion the industry of mankind." (**TMS** IV.1.8-10)

This 'deception by nature' (Raphael and Macfie, 1976: 8), which leads people to fulfil what they think are their own purposes, only to find they were fulfilling the purposes of a superior force or interest, is the counterpart in Smith of the 'cunning of reason' in Hegel[59], and the 'divine tactic' of history in Burke (Sabine, 1951: 519), both whom are known to have read and admired Smith[60].

The deception of nature is not ancillary but fundamental to Smith's principal doctrine. This becomes clear in the first few pages of **TMS** (I.i.1.1-13), where we find that, according to Smith, the whole structure of moral sentiments is built on illusion. The basis for morality is *sympathy*, that is, our ability to a limited extent to enter into the emotions of other people. But this participation in the pains and pleasures of others is achieved solely by an act of the imagination, divorced from the material causes of those pains and pleasures in the person we sympathise with. This sympathy even extends to fictional characters, people in the past and the dead – people, that is, who are incapable of feeling pain and pleasure in the first place, as well as to the insane, who are incapable of comprehending the degradation their illness has brought them to, and persons experiencing what we can never experience, such, if we are male, as a woman in labour. This shows sympathy to be a 'very illusion of the imagination' (**TMS** I.i.1.13), the imagination of 'what perhaps is impossible' (**TMS** I.i.1.11). We place ourselves, in the imagination, in the position of the other person, without in fact being in that position, and often without it being possible that we ever could be in such a position. We cannot help

[59] See Hegel (1952: §§344, 348) for the best expression of the 'cunning of reason' in Hegel, even though the term itself is not employed there.

[60] For Hegel, see the favourable comments on the political economy of Smith, Say and Ricardo in *The Philosophy of Right* (Knox, 1952: §189 and Addition); for Burke, see the long extracts from his review of **TMS** and letter to Smith of 1759 in Raphael and Macfie (1976: 27-28).

it: it is a god-given compulsion from which even the most hardened criminal is not completely immune (**TMS** I.i.1.1).

Smith's God treats individual humans in an extremely cavalier manner, subjecting them to all sorts of illusions and deceptions, and other weaknesses and indignities, and in general treating them like puppets, often with quite deleterious consequences to the individual in question, supposedly in the interest of maximising human welfare. A classic case of this occurs at the end of the first chapter of **TMS**, where he applauds even the fear of death as bad at the individual but good at the social level: 'one of the most important principles in human nature [is] the dread of death, the great poison to the happiness, but the great restraint upon the injustice of mankind, which, while it afflicts and mortifies the individual, guards and protects the society' (**TMS** I.i.1.13).

In connection with this we should perhaps recall the value which Smith really placed on the individual in the context of the overall system of which he is part. Before his God, says Smith, man appears as a 'vile insect' (**TMS** II.ii.3.12[61]). Again, in *The History of the Ancient Physics* he describes 'a God of all … who governs the whole by general laws, directed to the conservation and prosperity of the whole, without regard to that [sc the conservation and prosperity] of any private individual' (**Astronomy**: Physics 9).

So Smith's God teaches us that it is permissible to 'poison the happiness', to 'afflict and mortify the individual', to disregard 'the conservation and prosperity … of any private individual' – in the interest of society, of 'the whole'; and if we are to consider the individual a 'vile insect' relative to the totality of which he is part, we will certainly be unrestrained by respect for individual lives and individual suffering in pursuit of what we take to be the interest of that totality. Smith's love of 'the ennobling hardships and hazards of war' (**TMS** III.2.35) is germane here: 'War is the great school for acquiring and exercising … magnanimity.' It teaches a 'habitual contempt of danger and death' which 'ennobles the profession of a soldier, and bestows upon it ... a rank and dignity superior to that of any other profession' (**TMS** VI.iii.7). Indeed, a 'great warlike exploit' attracts a measure of 'esteem' just because it is military, even 'though undertaken contrary to every principle of justice' and by 'very worthless characters' (**TMS** VI.iii.8) Passages showing a quite militaristic outlook on society (**TMS** VI.ii.3-4), passages

introduced in the 6th edition of **TMS** at the end of Smith's life, have already been cited above. Twentieth century individualists and 'libertarians' claim intellectual descent from Smith: one wonders whether they have read him. As Alec Macfie says,

> "[T]he belief, so popularly accepted in the economic world, that Smith was primarily an individualist, is the very reverse of the truth. For him as for Hume, the interests of society were the end. By all means let the individual be encouraged to chase 'trinkets', so long as this conduced to that end. Smith's 'invisible hand' had a good smack of cynicism in its composition." (Macfie, 1961: 23)

* * *

The message of this section is thus that, according to Smith, people do things for apparent reasons – the real reasons being often hidden from them, and it is desirable that they should do so. They act justly from a sense of justice, but the reason why justice has been given us in this way is so that society may subsist; we admire the rich, the fortunate and the powerful, instead of the wise and virtuous, because it is in our nature to do so, but those feelings have been implanted in us to reconcile us to our lot; we mistake wealth for happiness, and are led to do so, so that trade and industry may flourish; we investigate the world thinking to discover its truth, so that by means of ever more pleasing stories about the world we may be reconciled to it.

4.4.4 *Review*

It is worth pausing here to summarise the points made so far about Smith's *Weltanschauung*:

1 The universe is a machine administered by a deity.
2 The sole purpose of the machine is to maximise happiness.
3 All parts of that machine, including individual people, play their allotted roles.
4 We do what we do because it is what we are led to do by the feelings implanted in our nature by the deity. All is part of the plan.
5 Even human folly and weakness are part of God's plan.
6 Everyone has nearly the same level of happiness.
7 We should therefore be content with our lot.

[61] eds 1-5 only.

8 The failure to realise this, mistaking wealth for happiness, leads people to be industrious: the economy depends on their being so deceived.

9 People mistake wealth and good fortune for wisdom and virtue.

10 This allows them to be reconciled to class distinctions and oppressive rulers.

11 We like morality and dislike immorality because we only see their proximate effects on human welfare.

12 This weakness is also a good thing as (a) it allows us to be moral and hence on the same side as God, and (b) morality, particularly justice, is a prerequisite for society.

13 Appearances are part of the divine plan.

In the next subsection we will see how these ideas relate to Smith's notion of an 'invisible hand'.

4.4.5 *The invisible hand*

Smith uses the term 'the invisible hand' on three occasions. The first occasion is in referring to 'the invisible hand of Jupiter' in **Astronomy**. There is a contrast between the role of the invisible hand here, on the one hand, and in **TMS** and **WN**, on the other: the action of the former is seen only in 'the irregular events of nature' rather than the 'ordinary course of things' (**Astronomy** III.2). In polytheism and 'early heathen antiquity', Smith says,

> "it is the irregular events of nature only that are ascribed to the agency and power of their gods. Fire burns, and water refreshes ... by the necessity of their own nature; nor was the invisible hand of Jupiter ever apprehended to be employed in those matters. But ... irregular events were ascribed to his favour or his anger .... Those ... intelligent beings, whom they imagined, but knew not, were naturally supposed ... not to employ themselves in supporting the ordinary course of things, which went on of its own accord, but to stop, to thwart, and to disturb it." (ibid)

Smith says that this was because humans acted in this way to change the course of events which would have occurred without human intervention and so primitive peoples supposed that their gods acted likewise (ibid). This, says Smith, is 'the lowest and most pusillanimous superstition' (ibid). Smith contrasts this view of gods, like men, as responsible for only the exceptional, with his own view of the whole world, including

societies and individuals within it, as a great machine designed and managed for the best interest of all by a divine administrator:

> "In the first ages of the world, the seeming incoherence of the appearances of nature, so confounded mankind, that they despaired of discovering in her operations any regular system. Their ignorance, and confusion of thought, necessarily gave birth to that pusillanimous superstition, which ascribes almost every unexpected event, to the arbitrary will of some designing, though invisible beings, who produced it for some private and particular purpose. The idea of an universal mind, of a God of all, who originally formed the whole, and who governs the whole by general laws, directed to the conservation and prosperity of the whole, without regard to that of any private individual, was a notion to which they were utterly strangers." (**Astronomy**: Physics 9)

So, firstly, not only the irregular, but, and much more importantly, the most regular occurrences are the work of the deity; and, secondly, human actions, too, far from being contrary to nature, are profoundly in harmony with it. Natural events and human actions alike and without exception[62] are part of the divine plan: 'Instead of acting capriciously, it [sc the invisible hand] becomes [the hand of] the 'all-wise Architect and Conductor', the 'author of nature', who governs and animates 'the whole machine of the world" (Macfie, 1971: 598).

In contrast to that in the **Astronomy**, Smith's use of the expression in **TMS** and **WN** is in a context where Smith is presenting his own views, not criticising someone else's. The second instance of Smith's use of the term 'invisible hand', in **TMS**, has already been given at the beginning of this chapter. In **WN** he says:

> "By preferring the support of domestick to that of foreign industry, he [sc 'every individual'[63]] intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse for society that it was no part of it. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it." (**WN** IV.ii.9)

---

[62] We shall see later that Smith *does* admit exceptions – and in this he is logically inconsistent.

[63] ie, every capitalist. Smith naively adopts the standpoint of the individual capitalist and momentarily forgets that there exist other agents, who have *no* role in 'directing ... industry'. It seems very ironic that the first of the two arguments for individual liberty which Smith gives here, is essentially a *mercantilist* argument: we do not need government intervention in foreign trade to give preference to domestic industry, because individual capitalists will be led by the invisible hand to prefer domestic industry without intervention.

In both cases he claims that the invisible hand will ensure that the unintended outcome of self-seeking behaviour will be socially desirable. Without it, in the **TMS** case, individuals would be subject to large differences in welfare; and in the **WN** case, the total wealth available to society would be smaller than it actually is (and more of it will fall into the hands of foreigners[64]).

It should by now be clear that the use of the phrase 'an invisible hand' is just another expression of Smith's particular interpretation of Stoic philosophy. The machine of the universe is managed by a deity determined on the utilitarian objective of the maximisation of happiness, and our emotions and motives are predestined by that deity to lead us to behave in a manner consonant with the divine plan. The administration of the plan is carried out by God – but, of course, we cannot see anything: his hands are invisible[65]: '[T]he invisible Hand [is] the name that Smith gives to the covert intervention of the Deity into the affairs of humankind' (Heilbroner, 1986: 57). Hence the concept of the invisible hand requires no separate treatment. We have already seen how agents are 'deceived by nature' to act in socially desirable ways, how the unintended consequences of our desire for justice, or riches, make society possible. The notion of an invisible hand is of a piece with this philosophy.

This is essentially the view of the invisible hand, and of the continuity of the invisible hand between **TMS** and **WN**, taken by Peter Gay:

> "Adam Smith … was something of a cosmic optimist who trusted unintended consequences. The "benevolence and wisdom" of the "divine Being" have "contrived and conducted the immense machine of the universe" in such a way that man may follow his private inclination and obey his most powerful passions, and yet benefit the social order. By taking care of his own happiness, man is led to promote the happiness of others – this is the notorious "invisible hand which leads men to "advance the interest of society" without intending it, without even knowing it. All is for the best in the only possible world that God could have made.
>
> "In *The Wealth of Nations*, Adam Smith keeps these philosophical concerns alive, but with greater subtlety than before, with far greater respect for harsh truths and for the exceptions that modify all rules, and with an impressive command of social realities. Like Diderot, Adam Smith learned much in the 1760s; like Diderot, he did not discard his essential philosophy, but complicated it." (Gay, 1969: 361)

[64] Once again: to report Smith's view of what is and what is not socially desirable is not to endorse it.

[65] Smith even furnishes us with an account of why God is invisible (**TMS** III.2.31, eds 3-5 only). If we could see him, Smith says, we would be so dazzled that we would be unable to go about our normal business.

Heilbroner reads Smith in much the same way. The theme of the invisible hand, he writes,

> "runs through all of the *Moral Sentiments*. The idea behind it is … [that] [m]an is by his human nature incapable of foreseeing the consequences of his actions beyond a very narrow range. How then does he know what course to follow, when he cannot use his faculties to anticipate the outcome of his own actions, much less those of his fellow actors? The question is answered in much the same way as the provision of a sense of duty and conscience. The Deity, when he created the world, gave to humankind a surer guide than reason. This was the call of its passions … the Invisible Hand refers to the means by which 'the Author of nature' has assured that humankind will achieve His purposes despite the frailty of its reasoning powers. The means are a number of powerful instincts and promptings that the Deity has instilled within us, which we obey because we have to, quite unconscious of their long-term social purpose. In this way, 'without intending it, without knowing it' the pursuit of our immediate desires brings us to follow courses of action that would otherwise require a Godlike intelligence to pursue." (Heilbroner, 1986: 60)

> "[In *The Wealth of Nations*] the remarkable capacities of the market system … are again evidence of the Invisible Hand. No participant in the market has in mind – or … has the power to effect – the orderly provisioning of society … [G]rowth starts as a consequence of the Invisible Hand, which has implanted within us that all-important confusion of wealth with betterment." (Heilbroner, 1986: 152-153)

In a similar vein, Alexander Gray criticised Smith's invocation of

> "the invisible hand, the divinity which shapes our selfish ends to public purposes. Frankly, we do not believe it; rather we have learned that the interests and the prosperity of the individual may be in conflict with the well-being of the community, that no such simple process of mathematical integration as Smith suggests is permissible" (Gray, 1931: 147).

Latter day Smithians, however, wishing to propagate a very different interpretation of Smith's theory, have had much trouble explaining the meaning of Smith's use of the phrase 'an invisible hand'.

> "Smith's 'invisible hand' had nothing to do with divine guidance. The phrase makes only one unfortunate appearance in *The Wealth of Nations* – unfortunate because it has been so totally misrepresented. There is no question, either in the specific context where he used the phrase or in the larger context of the argument of the entire book, but that the invisible hand is the hand of competition, which places immense pressure on individuals to behave in ways that simultaneously promote the public interest as well as the private interest." (Rosenberg, 1990: 21)

The fallacious view that the 'invisible hand' is not to be taken literally, but was a metaphor (or even simile), for competition, is extremely widespread. The invisible hand had, as we have seen, *everything* to do with divine guidance, and Rosenberg makes no

attempt to make the contrary case. His argument that Smith's invisible hand is 'the hand of competition' is not to the point. Of course this is true: Smith's whole argument is that God's wisdom works itself out through competition, through the 'simple system of natural liberty', as well as in other ways, such as our desire for the approbation of the 'impartial spectator'. But the notion of competition by no means exhausts the notion of the invisible hand, to which it is wholly subordinate.

> "We are used to thinking of the Invisible Hand as a term that describes the manner in which a free market economy is kept on an even course despite the absence of any steersman. But ... the Invisible Hand plays a far more important role than that of a ghostly economic planner. Without it, neither morality nor social order would be possible." (Heilbroner, 1986: 57)

A much earlier version of this chapter followed conventional usage in referring to a *metaphor* of 'the invisible hand' (see, for example, Hahn, 1982: 1; Raphael, 1985: 65). Also, Barry (1988: 19): 'since Adam Smith{ XE "Smith, Adam" } the equilibrium state has been said to be created by a metaphorical 'invisible hand' ...' and 'Adam Smith wrote … of a metaphorical 'Invisible Hand'' (1988: 27). Roche (1977: 10) speaks of 'Adam Smith and his analogy of the invisible hand'. I now think this mistaken. Smith was very consistent in flagging any such comparison by the use of simile instead of metaphor. Smith intended us to read his statements in **WN** and **TMS** of agents being 'led by an invisible hand' quite literally: the invisible hand leading them is just the hand of God. Had he desired another interpretation he would have written 'led *as*' or '*as if*', or '*as though* by an invisible hand'. Interestingly, the word(s) 'as' or 'as if' *are* frequently inserted into the passage in **WN** in question in a – presumably unconscious – misrepresentation as simile of what Smith saw only as literal truth. See, for example, Begg, et al (1991: 9): 'Smith argued that individuals pursuing their self-interest would be led 'as by an invisible hand' to do things that are in the interests of society as a whole.' Again, Schotter (1985: 11) informs us that 'Probably the greatest contribution of Adam Smith{ XE "Smith, Adam" } was his insistence that the freedom of individuals to maximise their own interests leads 'as if by an invisible hand to promote an end which was no part of his intention',' which falsifies what Smith wrote (as well as being ungrammatical). Generations of economists have thus been taught that Smith said '*as if* by an invisible hand', and I have found extraordinary resistance to the idea that he did not say that.

Raphael, one of the editors of the *Glasgow Edition* **TMS**, is, like Rosenberg, uncomfortable with the 'invisible hand' imagery:

> "Adam Smith's image of the invisible hand is not a piece of theology .... He uses the phrase for vivid effect, to give us a picture of an imaginary controlling device, but he knows very well that the effect comes about automatically through the interplay of individual interest and the system of exchange .... He did not believe that the God of theism controlled the working of the economy .... He drew on the familiar heritage of religious language simply in order to make his readers appreciate the remarkable character of the phenomenon. I do not mean that he deliberately placed a false halo around it. He was led by an invisible hand to choose evocative words." (Raphael, 1985: 66-67)

As we have seen, the invisible hand concept certainly was theological, and the 'controlling device' quite real. It is not a matter here of making a case for a new and radical reading of Smith's meaning: the theological interpretation is the first and most obvious meaning to strike the reader of what Smith actually wrote. It is the non-theological interpretation, the interpretation which says that, in spite of what Smith wrote, he actually meant something different, which requires demonstration. What is remarkable is the regularity with which those writers who wish to separate the invisible hand from the invisible mind which guides it simply resort to assertion without setting out the case for the their alternative interpretation. Raphael says that the working of the economy is secured 'automatically through the interplay of individual interest and the system of exchange'. But, again, the point is, why should this interplay lend itself to 'automatic' coordination? Raphael, like Rosenberg, falsely contrasts the intervention of God with the workings of competition. This failure to see divine intervention in the ordinary, automatic, day-to-day workings of the world, including the economy, is just what Smith deplored in the 'pusillanimous superstition' of primitive societies.

In his attempt to rescue Smith from the grasp of the invisible hand, Raphael writes 'No doubt Smith would say that the beneficial results [of the invisible hand] are ultimately due to nature or the divine author of nature, but he does not mean that God pulls the strings all the time' (ibid: 66). This raises two important points. Firstly, the idea that the reconciliation of individual plans and interests is 'ultimately' due to God. This is correct. God, in Smith, does not intervene *directly*, unmediatedly, in human affairs. We do not know what Smith did or did not privately believe – very likely he shared his friend, David Hume's well known scepticism regarding miracles. Certainly the public Smith of the **Astronomy**, **TMS** and **WN** shows no evidence whatsoever of belief in such miraculous

direct intervention. What he does very clearly show is a belief that human happiness is the distal, not proximal, consequence of God's will, mediated by the totality of natural and social phenomena. The latter, including the 'simple system of natural liberty' are the indirect manifestations of God's will. Secondly, Raphael raises the issue of whether Smith's God is pulling the strings 'all the time'. This is a profound ambiguity in Smith. Either God *is* pulling the strings all the time, and we are literally puppets with no personal autonomy whatever. In this case the correct philosophical response would be utter fatalism and apathy. Or we need some guidance on when to treat outcomes as representing God's will and when not to. We will revisit this problem in Smith later in this section and again later in the chapter.

The final point to note in connection with Raphael's presentation of the issue concerns his facetious defence of Smith on the grounds that it was the 'invisible hand' what dunnit: 'He was led by an invisible hand to choose evocative words'. For the charge against Smith, which Raphael is attempting to refute, is that he sanctified the working of the economy by 'plac[ing] a false halo around it'. As we shall see, later in this chapter, that is exactly what Smith was trying to do, and in this he was in step with his generation. Raphael's defence that he did not do so 'deliberately', but allowed himself to be led to do it by an invisible hand, effectively concedes the case.

To return to the point at issue. Hayek, too, makes it clear that he regards the phrase as unfortunate:

> "Adam Smith and the other great Scottish individualists of the eighteenth century – even though they spoke of the "invisible hand" – provided ... an explanation [of how the interaction of the efforts of individuals can create something greater than they know]" (**CRS**: 392-393).

For details of this explanation, supposedly given by Smith, Hayek refers us, in a footnote, to the first essay, 'Individualism: True and False', in his earlier (1949) book *Individualism and Economic Order* (**IEO**). Sadly, his promise is not redeemed. What Hayek does – and this point will be amplified in the chapter on Hayek – is to argue for the *spontaneity* of the order which emerges from the market in Smith's notion of the invisible hand, abstracting from the *optimality* with which Smith endowed it, and then tacitly to assume that optimality has been established.

Blaug, after perfectly reasonably criticising the notion of the invisible hand as involving a fallacy of composition – what is good for the individual is necessarily good for society – complains that

> "a legend has grown up that the whole of the *Wealth of Nations* rests on this kind of naive reasoning, the so-called doctrine of 'the spontaneous harmony of interests'. But 'the obvious and simple system of natural liberty', which is said to reconcile private interests and economic efficiency, turns out upon examination to be identical with the concept of perfect competition; the 'invisible hand' is nothing more than the automatic equilibrating mechanism of the competitive market." (Blaug, 1978: 59)

But the whole point is, that competition was able to act as an 'automatic equilibrating mechanism' in Adam Smith, solely because the individual interests which it had to balance had already been pre-reconciled by a kindly Great Administrator of the system of the universe. While it is perfectly true that the whole of **WN** does not rest on this 'naive reasoning', it is the case that part of it does – and that part is precisely the doctrine of the invisible hand, which lies at the heart of Smith's notion of competition.

Macfie, too, deprecates the role of the invisible hand in Smith: 'It is ... not in the invisible hand that his most valuable and original contribution was made.' (Macfie, 1967: 125) 'If we were asked today to state in one phrase the essence of Smith's message, we should most truly reply by calling him the apostle of natural liberty' (ibid: 130).

At this point some reference should be made to Viner's (1927, 1958) version of the 'Adam Smith Problem', presented in a lecture in 1926. Viner sustains the view that **TMS** and **WN** are irreconcilable by scouring **WN** for exceptions to Smith's general endorsements of 'natural liberty' and its corollary of free competition and *laissez-faire*. On the basis of an apparently impressive mass of such exceptions Viner argues that there is a change of methodology between **TMS** and **WN**: in the former we have the theistic invisible hand type of argument (which I ascribe to *both* **WN** and **TMS)**, where a natural harmony of interests is *deduced* from the assumed attributes of the deity; in the latter we have an *inductive* invisible hand theory, only valid where the facts justify it.

Now, it is true that there are many exceptions to the desirability of *laissez faire* policies in **WN** – exceptions, indeed, to which twentieth century Smithians and invisible hand theorists would do well to give more attention – and Viner performs a valuable service

by collating them. Nevertheless, they remain *exceptions*. To argue that Smith's scientific methodology varied so abruptly between the psychological and economic aspects of his scientific work, and without a single hint anywhere that he was simultaneously adopting such contrary standpoints in the two disciplines, does unacceptable violence to the unity of Smith's thought. I have argued in this chapter that Smith started out with a clear conception of the nature of scientific thought; it is also the case that he carried out the programme implicit within it with relentless consistency throughout his life. If there had been only one edition each of **TMS** and **WN**, then, while still incorrect, the case would appear more plausible, that Smith changed his position between 1759 and 1776. However, this is not the case, and at the end of his life, in 1790, Smith was still saying substantially the same things in his revisions to both books as he had in their first editions. The weakness of Viner's argument can be seen when he argues that we can discount the sixth edition of **TMS** as Smith by now was 'elderly and unwell' (Viner, 1958: 231). In the context of the corpus of Smith's work as a whole, Viner's hypothesis can be seen to be implausible in the extreme. As Macfie argues on the critical question of the role of the invisible hand in **TMS** and **WN**: 'the almost theological view of the invisible hand ... [is] exactly carried over from the *Moral Sentiments* into the *Wealth of Nations*' (Macfie, 1959: 211[66]; see also ibid: 223-4.)

In the 1920s, Viner could call it 'a commonplace among the authorities on Adam Smith that it is impossible fully to understand the *Wealth of Nations* without recourse to the *Theory of Moral Sentiments*' (Viner, 1958: 216). Today this is no longer true, and, it seems to me, Viner himself must share some of the blame for the partial occlusion of that fundamental insight into the meaning of the invisible hand in **WN**.

The discussion raises two serious issues, however. Firstly, it is the case that the expression 'an invisible hand' only appears once in **WN**, and the theistic explanation of it does not appear at all. A great deal of empirical material, however, does appear, and, while the overwhelming bulk of that material is directed towards showing the superiority of the *laissez faire* system, he does indicate exceptions to its desirability. It is easy to see how modern, nineteenth and twentieth century readers of **WN** in isolation from Smith's

[66] Even here the use of 'almost' shows Macfie's hesitation to take literally what Smith explicitly says about the role of God in human affairs. Elsewhere he describes "the purpose of the 'invisible hand'" as 'theological' without qualification (Macfie, 1961: 19).

other works and from those of his contemporaries, should assume that this was a predominantly empirical study drawing the conclusion that in general, free competition was a good thing. It is easy to overlook the fact that the empirical material only plays the role of illustrating a preconceived order. Smith does not in fact anywhere make the inductive judgement that, as a generalisation, individual self-seeking behaviour leads automatically to socially desirable outcomes – on the contrary, this is assumed beforehand and illustrated by details of many empirical circumstances where it is *asserted*, over and over again, that this has occurred, or would occur if only enterprise were free. It is only by exploring the totality of Smith's thought, not only in the **WN** but in **TMS** and **Astronomy** as well, that we can clearly see the *a priori* and deductive nature of Smith's procedure, the assumption that the spontaneous system of free enterprise will lead to desirable outcomes because, in general, agents' interests are pre-reconciled by the invisible hand of a benevolent deity[67]. Whatever the stylistic and presentational differences between **TMS** and **WN**, this *faith* remains the starting point of Smith's account of the invisible hand throughout: 'the invisible hand here [sc in **WN**] remains to control the individual conflicts and excesses of competition, and to safeguard the public good through healthy competition. Such is his faith' (Macfie, 1959: 212).

The second point is that the supposition that there is a divine plan, in which all agent interests are fundamentally in harmony, does not seem to allow of 'exceptions'. We will return to this point in Section 5c, below. In brief, the point is that there is a difference between two kinds of inconsistency. Viner alleges an *arbitrary* inconsistency in which Smith switches, without comment, between two fundamentally contrasting standpoints. Why should Smith have done this? In my interpretation, however, Smith is inconsistent because his standpoint compels him to be: the inconsistency is implicit in his world view. He has adopted that standpoint and has to live with the consequences. The exceptions he notes were, in general and in modern language, those associated with externalities, public goods and market power. Smith was faced with the choice of being dogmatic – even in these circumstances the invisible hand will sort things out, which is what his theory actually implies, or moderate – dropping the theory without explanation when its consequences strain credulity. Wisely, he chose the latter: a more rigorous and

[67] In **TMS** 'are to be found the fundamental doctrines of the *Wealth of Nations* ... and the famous work cannot be properly understood without some knowledge of the *Theory*' (Macfie, 1961: 12).

intellectually consistent approach would have made **WN** far less plausible, palatable and effective for its purpose.

Hegel, deducing the State, in the *Philosophy of Right*, from the theological category of the Idea elaborated in the *Science of Logic*, should, in faith to his system, have reproduced the Prussian state exactly as it was. As is well known, however, he could not resist idealising the existing state and ended up with an improved, more consistent version. By breaking the link between the ideal and actual in this way, he only showed that his mode of procedure was capable of being used to deduce and justify whatever state system one desired. Smith is in the same position. Within his system, natural liberty, which gives the invisible hand its operational scope, must be absolute. If the invisible hand is God's hand, surely it must be sacrilegious to attempt to restrain or constrain its movements, however mysterious and obscure they may be[68]. The macro level is God's sphere of competence, only the micro level is ours: 'the care of the universal happiness of all rational and sensible beings, is the business of God and not of man' (**TMS** VI.ii.3.6).

In **WN**, however, Smith says

> "To restrain private people, it may be said, from receiving in payment the promissory notes of a banker ... is a manifest violation of that natural liberty which it is the proper business of law, not to infringe, but to support. Such regulations may ... be considered as in some respect a violation of natural liberty. But those exertions of the natural liberty of a few individuals, which might endanger the security of the whole society, are, and ought to be, restrained by the laws of all governments; of the most free, as well as of the most despotical. The obligation of building party walls, in order to prevent the communication of fire, is a violation of natural liberty, exactly of the same kind with the regulations of the banking trade which are here proposed." (**WN** II.ii.94)

But to say this undermines Smith's whole edifice: what are to be the rules governing government intervention? If God gets it wrong here, where else does he get it wrong? Is there really a god governing our lives after all? The idea of an invisible hand evaporates leaving us with something much more mundane and imperfect: an accidental and unreliable coincidence between individual and social interests, the spontaneity

[68] In another context, Flew says that the variety and contrariety of conclusions drawn from a single premiss 'must constitute a strong reason for challenging the legitimacy of the sort of derivation proposed' (Flew, 1967: 5).

proposition without the optimality proposition. This issue will be revisited at the end of the chapter.

* * *

The phrase 'an invisible hand' occurs throughout nineteenth century literature – in Mary Shelley (1818) *Frankenstein* Ch XII, Thomas Hardy (1874) *Far From the Madding Crowd* Ch 42, and in HG Wells (1898) *The War of the Worlds* Ch 6, to give just three examples – in each case in utterly pedestrian contexts. Raphael (1985: 67) gives an instance of its use in the early eighteenth century, when a captain wrote in his log that the ship had been saved from sinking by 'the invisible hand of Providence'. That the idea of the guiding hand of an unseen god ensuring the desirable social consequences of self-seeking behaviour – without the phrase of the 'invisible hand' itself, however – was a commonplace of late eighteenth century social commentary is shown by Hayek by reference to Smith, Tucker, Ferguson and Edmund Burke (**IEO**: 7). Taking the last as example: 'The benign and wise disposer of all things ... obliges men, whether they will it or not, in pursuing their own selfish interests, to connect the general good with their own individual success' (Burke (1795) *Thoughts and Details on Scarcity* cited in **IEO**: 7).

The next section looks in more detail at the relationship between Smith and his contemporaries.

4.5 Smith's intellectual environment[69]

4.5.1 *The 'Heavenly City' of the 18th Century* Philosophes

> All are but parts of one stupendous whole,
> Whose body nature is, and God the soul;

[69] Much of this section relies on Becker (1932). Becker has been heavily criticised, notably in Peter Gay ('Carl Becker's Heavenly City' (1957) reprinted in Gay, 1964: 188-210). The points made in this section remain substantially untouched by Gay's criticism, which boils down to little more than the complaint that Becker exaggerates. The same point could be made about Gay. Unfortunately, this is not the place for a thorough analysis of the problems raised by Gay's very interesting discussion of Becker, of Smith (Gay, 1969: passim), or of the *Philosophes*' 'Revolt Against Rationalism' (Gay, 1969: 187-207).

....
All discord, harmony not understood;
All partial evil, universal good:
And, in spite of pride, in erring reason's spite,
One truth is clear, *Whatever is, is right*.
(Alexander Pope, cited in Becker, 1932: 66[70])

Adam Smith was very much a man of his time. Smith's 'modified Stoicism typical of Cicero' was 'almost conventional in the Enlightenment' (Macfie, 1959: 210). Robert Heilbroner stresses the need to locate Smith in his times in order to understand his works (Heilbroner, 1986: 1). He also stresses that what links Smith and his contemporaries is the conservatism of both:

> "We come to Smith expecting to find a great monument of conservative economic thought, and we will not be disappointed – Smith is indeed the greatest of all conservative economists ... his conclusions about mankind are profoundly conservative ... Adam Smith, like all his contemporaries, believed firmly in the need for a well-defined social hierarchy and a firm adherence to the principle of property. ... In this essentially conserving vision of social continuity and order, the Enlightenment thinkers found the basis for their distinctive brand of philosophical and historical conservatism. [They were] [c]onvinced of the need for – indeed, the inescapable necessity of – a stratified, property-based social system ... Neither he [sc Smith] nor any of his contemporaries imagined a society in which exploitation and oppression would not be present ... [A]ll the *Philosophes*, including Smith, share one limit to their social imaginations. This is an inability to imagine that the lower orders might some day exercise sovereignty over society. Democracy, with all its implicit threats to property and hierarchy, was not yet on the political agenda and would not be put there until the French Revolution" (Heilbroner, 1986: 1-3).

This is a theme which is taken up at length in Carl Becker's *The Heavenly City of the Eighteenth-Century Philosophers* (Becker, 1932), in which, especially in Ch II 'The Laws of Nature and of Nature's God' (Becker, 1932: 33-70), he makes a powerful case that the intellectuals of this period[71] were not in any meaningful sense 'modern', but that, on the contrary, they were living in a medieval world and 'demolished the Heavenly City of St Augustine only to rebuild it with more up-to-date materials' (Becker, 1932: 31):

---

[70] The italicised concluding statement is the exact counterpart of Hegel's assertion that 'the real is the rational' (Hegel, in Knox, 1952: 10), and has exactly the same purpose, namely, to 'reconcile us to the actual' (ibid: 12). See also Wallace (1975: §6).

[71] He includes in the term *philosophes*, amongst others, from France: Montesquieu, Voltaire, Volney, Diderot, Savigny and Rousseau; from Germany: Leibniz, Lessing, Herder and Goethe; from Britain: Locke, Hume, Ferguson and Adam Smith{ XE "Smith, Adam" }; and from America: Jefferson and Franklin (Becker, 1932: 33).

> "We are accustomed to think of the eighteenth century as essentially modern in its temper .... And yet I think the *Philosophes* were nearer the Middle Ages, less emancipated from the preconceptions of medieval Christian thought, than they quite realized or we have commonly supposed .... [T]hey speak a familiar language .... But I think our appreciation is of the surface more than of the fundamentals .... [I]f we examine the foundations of their faith, we find that at every turn the *Philosophes* betray their debt to medieval thought without being aware of it .... They had put off the fear of God, but maintained a respectful attitude towards the Deity. They ridiculed the idea that the universe had been created in six days, but still believed it to be a beautifully articulated machine designed by the Supreme Being according to a rational plan as an abiding place for mankind .... they renounced the authority of church and Bible, but exhibited a naïve faith in the authority of nature and reason .... [T]he underlying preconceptions of eighteenth century thought were still ... essentially the same as those of the thirteenth century." (Becker, 1932: 29-31)

On the overall aim of the philosophers, he cites Hume – with whom Smith shared a mutual admiration and close friendship – as an example, 'Hume is representative of his century' (Becker, 1932: 39). Like Smith, Hume was sufficiently concerned with preservation of the social order to be willing to lay down his pen in its service. In his own words:

> "I am at present castrating my work ... that is, endeavouring it shall give as little offence as possible." (cited in Becker, 1932: 38) "A man has but a bad grace who delivers a theory, however true, which leads to a practice dangerous and pernicious. Why rake into those corners of nature, which spread a nuisance all around? ... Truths which are *pernicious* to society ... will yield to errors, which are salutary and *advantageous*.." (ibid)

Here, as in Smith, we find the idea that error can be 'advantageous'. Following through the programme just mentioned,

> "in mid career Hume abandoned philosophical speculations for other subjects, such as history and ethics, which could be treated honestly without giving 'offense'" (Becker, 1932: 38-39). "These are, no doubt, the reasons why Hume locked his *Dialogues* away in his desk[72] ... his contemporaries, could they have looked into that locked desk, would have found ... the brilliant argument that demolished the foundations of natural religion .... Hume ... refused to publish his Dialogues, and never, in public at least, failed to exhibit a punctiliously correct attitude toward the Author of the Universe." (Becker, 1932: 78[73])

[72] Hume's *Dialogues Concerning Natural Religion* remained unpublished until after his death. It was Adam Smith**{** XE "Smith, Adam" **}** who persuaded him to suppress it. The manuscript was originally left to Smith, Hume's literary executor, but at the last moment Hume gave it to a nephew who published it. Had Smith laid hands on it he would undoubtedly have burned it. The event had a profound effect on him: as Smith's own death approached he became extremely agitated about ensuring that his own papers were burnt, which was done a week before his death.

[73] See also Becker (1932: 79-81), for a similar story about Diderot.

It is well known that Adam Smith was a close friend of Hume's and admired his work enormously. He described Hume as the nearest possible to 'a perfectly wise and virtuous man' (**TMS** p383). Hume's words are in perfect agreement with Smith's project of prioritising reconciliation over investigation.

In Becker's view, the *Philosophes* faced

> "the ugly dilemma, emerging from the beautiful premises of the new philosophy: if nature is good, then there is no evil in the world; if there is evil in the world, then nature is so far not good .... Will they, closing their eyes to the brute facts, maintain that there is no evil in the world? In that case there is nothing for them to set right. Or will they, keeping their eyes open, admit that there is evil in the world?" (Becker, 1932: 69)

The philosophers were at a crossroads: reason pointed forwards, to atheism and to the project of rebuilding a haphazard, spontaneous and irrational society in the image of the order they had previously ascribed to nature; the alternative was the denial of reason and a return to medieval Christian faith. Open-eyed, they could adopt an empirical, materialist standpoint, recognising the need to take control of, and responsibility for, spontaneous human institutions; or with eyes closed they could take an *a priori* stance, imposing on the world a scheme derived from religious belief. 'Well, we know what the Philosophers did in this emergency. They found ... that reason is amenable to treatment. They therefore tempered reason with sentiment ...' (Becker, 1932: 69). 'Sometime about 1750, men of sense became men of sentiment ...' (Becker, 1932: 41).

None of this was written with Smith specifically to the forefront of Becker's mind – but the description fits like a glove. Smith is the epitome of this intellectual retreat of the enlightenment in the late eighteenth century, the retreat from rationalism to romanticism[74]. In every respect, reason is belittled and sentiment and religion brought to the fore[75]. At best, reason only confirms what we know anyway by means of sentiment and religion:

[74] Viner (1958: 216) speaks of 'the absolutism, the rigidity, the romanticism' which characterise **TMS**.

[75] This is not to criticise his rejection of a rationalist account of morality (see **TMS** VII.iii Ch II *Of those Systems which make Reason the Principle of Approbation)*, epitomised in the title of *The Theory of Moral Sentiments*, which so far as it goes, is undoubtedly correct, though he is mistaken in the reason he gives for such sentiments arising in the first place.

> "This reverence [for general rules] is still further enhanced by an opinion which is *first impressed by nature, and afterwards confirmed by reasoning* and philosophy, that those important rules of morality are the commands and laws of the Deity, who will finally reward the obedient, and punish the transgressors of their duty .... [R]eligion ... gave a sanction to the rules of morality, long before the age of *artificial reasoning* and philosophy. That the terrors of religion should thus enforce the natural sense of duty, was of too much importance to the happiness of mankind, for nature to leave it dependent on the slowness and uncertainty of philosophical researches. These researches, however, when they came to take place, confirmed those original anticipations of nature." (**TMS** II.5.3, my emphasis)

Reasoning, for Smith, is artificial, and only sentiment is natural:

> "That the Deity loves virtue and hates vice ... for the effects which they tend to produce ... is not the doctrine of nature, but of an artificial, though ingenious, refinement of philosophy. All our natural sentiments prompt us to believe [the opposite] ..."[76] (**TMS** p91 note, editions 1 and 2[77])

For Smith reason is 'the abstruse syllogisms of a quibbling dialectic', and sentiment, 'the great discipline which Nature has established' (**TMS** III.3.21).

The medieval view of the world, and the role of reason within it – the view of the world to which Smith and his contemporaries turned – is well summarised by Becker:

> "Existence was ... regarded by the medieval man as a cosmic drama, composed by the master dramatist according to a central theme and on a rational plan. Finished in idea before it was enacted in fact ... the drama was unalterable either for good or evil .... the duty of man was to accept the drama as written, since he could not alter it; his function, to play the role assigned .... Intelligence was essential, since God had endowed men with it. But the function of intelligence was strictly limited .... The function of intelligence was therefore to demonstrate the truth of revealed knowledge, to reconcile diverse and pragmatic experience with the rational pattern of the world as given in faith." (Becker, 1932: 7)

Smith, therefore, was in many ways typical of the philosophers of the period – on Becker's interpretation of the eighteenth century. Like Hume, who was a major influence on his philosophy, Smith regarded the preservation of the social order as of primary importance. Like his contemporary, Kant, who was also, though in a different direction, influenced by Hume[78], Smith wanted to place limits on the legitimate field of

[76] Smith, incidentally, here clearly ascribes to God his own hypostatisation of the intermediate, the valuing of the means 'for its own sake' over the end.

[77] Editions 3-5 read: '... of untaught nature but of an artificial refinement of reason and philosophy. Our untaught, natural sentiments ...' (ibid). The passage is omitted in edition 6.

[78] See Kant (1950: 5 ff; or Academy edition, Vol IV: 258 ff).

action of reason, to find a space for instinct and religious belief[79]. Perhaps the greatest overlap between Smith and his contemporaries lay in their application of the doctrine of natural law. This is the topic of the next subsection.

4.5.2 '*Nature*' *and the natural in Smith*

> "With Adam Smith and his disciples ... nature means the totality of impulses and instincts by which the individual members of society are animated; and their contention is that the best arrangements result from giving free play to those forces in the confidence that partial failure will be more than compensated by success elsewhere, and that the pursuit of his own interest by each will work out in the greatest happiness of all" (AW Benn, 1906, *History of English Rationalism in the Nineteenth Century* cited in **IEO**: 12 n 15)

The reader may have noticed the number of times, in the passages cited above, Smith uses the term 'Nature' interchangeably with that of the Deity. Far from being the passive background or substrate of our activities, nature is seen as a direct manifestation of the deity, as an active principle intervening in our lives. These citations illustrate Smith's adoption and adaptation of the archaic conception of natural law[80] so popular amongst eighteenth century philosophers (Becker, 1932: Ch II; Sabine, 1951: Ch XXVIIff).

The late eighteenth century French philosopher, Comte de Volney, defined natural law in eminently Smithian terms:

---

[79] 'Smith ... combined with Hume to prepare the way for Kant' (Macfie, 1967: 156). Smith has even been presented, though not, I think, completely successfully, as a direct precursor of Kant. Macfie (1967: 68 and n 24; 91 n 23) gives the references. In his 1959 paper, Macfie writes 'If, then, Hume lit the destructive spark which helped to inspire the Kantian reconstruction, Smith may well have contrived ... the revealing light which led to the *Critique of Practical Reason*.' (Macfie, 1959: 216) (In spite of the declaration in the 'Introductory' to Macfie's 1967 volume that 'I have made no changes in the texts' (Macfie, 1967: 12), 'the revealing light which led to' Kant's book is now reduced to 'a light reflected in' it (ibid: 68). He also wrongly claims that 'In the first five papers, I have kept to the order of appearance': Paper 3 appeared in 1961 and Paper 4 in 1959, while Paper 5 was not previously published (ibid: 12).) Elsewhere, Macfie writes 'Smith, as well as Hume, directly influenced Kant. What greater privilege!' (Macfie, 1961: 26)

[80] It is interesting in this context that the very reason that Smith started to study economics was his need to include the topic in his course of lectures on Moral Philosophy. This tradition 'stemmed from the treatment of natural law by Roman and medieval writers' (Raphael and Macfie, 1976: 24). Interestingly, Physiocracy, with which, of course, Smith had much to do, is a near-synonym for natural law.

> "What is natural law? It is the regular and constant order of facts by which God rules the universe; the order which his wisdom presents to the sense and reason of men, to serve them as an equal and common rule of conduct, and to guide them ... towards perfection and happiness." (Cited in Becker, 1932: 33, 45).

Here again we see the universe as an orderly system administered by a god. The order implicit in it, which is presented to both the senses and the reason of humans, issues in both factual statements about the way the world is, and normative statements as to how people are to behave, so as to correspond with the divine will. Again the god is utilitarian, maximising the happiness of mankind.

Becker cites this definition as typical of the eighteenth century philosophers, among whom he explicitly includes Adam Smith (Becker, 1932: 33). His commentary certainly applies well to Smith:

> "The language is familiar, but the idea, once we examine it critically, is as remote as that of Thomas Aquinas. Important if true, we say; but how comes it, we ask, that you are so well acquainted with God and his purposes? Who told you ... that there is a regular and constant order of nature? ... Indeed it is all too simple. It assumes everything that most needs to be proved and begs every question we could think of asking." (Becker, 1932: 45)

I keep stressing the primacy of *order* in Smith, and the same is true of the *Philosophes*: they wanted to be able to point to an ordered *natural* world in order to justify the conceptions of *social* order to which they variously subscribed:

> "Most eighteenth-century minds were too accustomed to a stable society with fixed ranks, too habituated to an orderly code ... to be at all happy in a disordered universe. It seemed safer, therefore, ... to retain God ... as a ... guaranty that all was well in the most comfortable of common-sense worlds." (Becker, 1932: 49-50)

And if a god did not exist, it would be necessary, as Voltaire (in)famously declared, to invent one. But a god in isolation, separate from the world, was not to the point. Their programme demanded that God directly reveal himself in nature:

> "God had revealed his purpose to men in a ... simple and natural ... way, through his works. To be enlightened was to understand ... that it was ... in the great book of nature ... that the laws of God had been recorded. This is the new revelation ... This open book of nature was what Jean Jacques Rousseau and his philosophical colleagues went in search of when they wished to know what God had said to them. Nature and natural law – what magic these words held for the philosophical century! ... Hume, Voltaire, Rousseau, Volney: in each of them nature takes without question the position customarily reserved for the guest of honor .... Search the writings of the new economists and you will find

> them demanding the abolition of artificial restrictions on trade and industry in order that men may be free to follow the natural law of self-interest .... controversialists of every party unite in calling upon nature as the sovereign arbiter of all their quarrels." (Becker, 1932: 51-52)

Perhaps we can best see the importance of this view of nature in the popular and scholarly response to a figure towering over the eighteenth century, that of Newton. During the course of the century, a large number of popular guides to Newton's philosophy were published. The point of interest was not the technical detail but the overall philosophy, in particular Newton's approach to the most fundamental of human problems – the relations between humanity, nature and God. Colin Maclaurin, Professor of Mathematics in the University of Edinburgh, set out the nature of these relationships in his own guidebook, *An Account of Sir Isaac Newton's Philosophical Discoveries*, published in 1775:

> "To describe the *phenomena* of nature, to explain their causes ... and to enquire into the whole constitution of the universe, is the business of natural philosophy .... But natural philosophy is subservient to purposes of a higher kind, and it is chiefly to be valued as it lays a sure foundation for natural religion and moral philosophy; by leading us, in a satisfactory manner, to the knowledge of the Author and Governor of the universe....
>
> "We are from his works, to seek to know God, and not to pretend to mark out the scheme of his conduct, in nature, from the very deficient ideas we are able to form of that great mysterious Being ....
>
> "Our views of Nature, however imperfect, serve to represent to us, in the most sensible manner, that mighty power which prevails throughout ... and that wisdom which we see displayed in the exquisite structure and just motions of the greatest and subtilest parts. These, with perfect goodness, by which they are evidently directed, constitute the supreme object of the speculations of a philosopher; who, while he contemplates and admires so excellent a system, cannot but be himself *excited and animated to correspond with the general harmony of nature*." (Maclaurin, 1775, cited in Becker, 1932: 62-63)

After citing this passage, Becker immediately adds: 'The closing words of this passage may well be taken as a just expression of the prevailing state of mind about the middle of the eighteenth century. Obviously the disciples of the Newtonian philosophy had ... deified nature' (Becker, 1932: 63).

The deification of nature led, as it was supposed to lead, to the sanctification of the particular model of human behaviour the philosophers wished to hold up as 'natural'. The Declaration of Independence, for example, invokes 'the laws of nature and of nature's God' (cited in Becker, 1932: 52) to sanction its particular programme. Macfie, speaking of the 'Scottish Tradition in Economic Thought', says that 'The main faith

which the Law of Nature and Stoicism inspired in Scotland was a faith in natural liberty in a natural society.' (Macfie, 1967: 26) In Smith we see frequent references to the 'sacred laws of justice' (**TMS** II.ii.2.3), a 'sacred regard to general rules' of morality (**TMS** III.5.2); 'by the wisdom of Nature, the happiness of every innocent man is ... rendered holy, consecrated, and hedged round against the approach of every other man' (**TMS** II.iii.3.4). And in **WN**, we read that Britain's trade policy with America, though in fact 'not very hurtful to the colonies' was, in diverting trade from its spontaneous course, 'a manifest violation of the most sacred rights of mankind' (**WN** IV.vii.b.44).

For Smith, therefore, as was commonly the case in natural law theorists, what is natural is god-given and therefore implicitly good. When Smith describes certain institutional arrangements in **WN** as 'natural', and others, on the contrary, as 'artificial' (as, for example, in **WN** IV.ii.3), he is saying that the former are not just spontaneous, but spontaneous *and therefore* an immediate expression of the will of God, whereas the latter must at the very least lie under the suspicion of sacrilege. There are many occasions where Smith invokes nature[81] in this way in **WN**. For example: 'All systems of preference or of restraint [of trade by the government] ... being ... completely taken away, the obvious and simple system of natural liberty establishes itself of its own accord' (**WN** IV.ix.5). '[V]iolations of natural liberty [are] ... unjust' (**WN** IV.v.b.16 [82]).

In his lectures as early as 1749 Smith was linking the ideas of an active, beneficent and rational nature – in short a *teleological* nature – to the policy prescription of *laissez-faire*: 'Projectors disturb nature in the course of her operations in human affairs, and it requires no more than to let her alone, and give her fair play in the pursuit of her ends that she may establish her own designs' (Smith, cited in Gay, 1969: 354). 'To let alone' is, almost certainly, a conscious translation of the phrase '*laisser faire*', which had been

[81] 'Natural' in Smith often, of course, also has the ordinary meaning of 'arising spontaneously in the normal course of events and from the circumstances of the case', rather than the special sense of 'divinely appointed' which I identify here. I do not wish to say that *every* use of 'natural' in **WN** has the second meaning, only that in the many invocations of nature in **WN** the penumbra of connotation is definitely intended to include this second meaning on many occasions. An excellent discussion of the meanings of 'natural' in Adam Smith appears in Waterman (1997).

[82] For 'natural liberty', see also **WN** I.x.c.59. For 'natural', in particular, the 'natural balance of industry', see, to mention only a few, **WN** I.vii passim, **WN** IV.i.12, **WN** IV.ii.3 and **WN** p453 editorial footnotes 7 and 8, containing references to further passages in **WN** and the *Early Draft* of **WN**.

in use in France since the end of the previous century to denote freedom from government interference.

But Smith extends the idea of what is natural to include human nature. What is instinct in us was implanted there by Nature, for a purpose – and this includes our weaknesses as well as our strengths. Thus, speaking of resentment and its issue in revenge, 'the most detestable of all the passions' (**TMS** II.i.5.8), he remarks that even here 'Nature ... does not seem to have dealt so unkindly with us, as to have endowed us with any principle which is wholly and in every respect evil, or which, in no degree and in no direction, can be the proper object of praise and approbation' (ibid). Thus resentment, like every other emotion, is divinely appointed, an endowment of 'Nature', but can become vicious when taken to an excess. This tactic, however logical in itself, involves Smith in inescapable contradictions once he attempts to derive his *laissez-faire* policy prescription from it, as we shall see in the next subsection.

So Smith has a similar approach to nature and the natural as his contemporaries. If anything, however, Smith is even more archaic than his contemporaries. Prior to the eighteenth century, according to Becker,

> "philosophers ... argued that, since God is goodness and reason, his creation must somehow be, even if not evidently so to finite minds, good and reasonable. Design in nature was thus derived *a priori* from the character which the Creator was assumed to have; and natural law, so far from being associated with the observed behaviour of physical phenomena, was no more than a conceptual universe above and outside the real one, a logical construction dwelling in the mind of God and dimly reflected in the minds of philosophers." (Becker, 1932: 55)

In the eighteenth century, however, – he cites Hume, in the person of Cleanthes in his *Dialogues*, as epitome – the logical process is reversed:

> "Cleanthes does not conclude that nature *must* be rational because God *is* eternal reason; he concludes that God *must* be an engineer because nature *is* a machine." (ibid: 56)
> "[T]he very foundation of the new philosophy was that the existence of God, if there was one, and his goodness, if goodness he could claim, must be inferred from the observable behaviour of the world. Following Newton, the Philosophers had all insisted on this to the point of pedantry" (ibid: 67).

Smith in this respect is out of step with his contemporaries. He clearly *starts* by deducing the nature of the world from a prior consideration of the 'necessary' qualities of

the deity, and only afterwards claims to be able to support his conclusions by reference to observations of nature itself:

> "The happiness of mankind, as well as of all other rational creatures, seems to have been the original purpose intended by the Author of nature, when he brought them into existence. No other end seems worthy of that supreme wisdom and divine benignity which we necessarily ascribe to him; and this opinion, which we are led to by the abstract consideration of his infinite perfections, is still more confirmed by the examination of the works of nature, which seem all intended to promote happiness, and to guard against misery." (**TMS** III.5.7)

There is no reason to believe that Smith would have seen any opposition between these two approaches – deductive versus inductive, *a priori* versus empirical – to the relation between God and nature. But he would certainly have rejected the latter as sole, or even major, support for his philosophy. Reason is 'artificial' and fallible, and our finite minds do not perceive the remote ramifications of things. Things, as he stresses in **Astronomy**, often appear to us to be discordant and unconnected. This is precisely why we need a 'soothing' scientific explanation of things, and God's will, manifested in natural law, is the most pleasing general explanation available. So it would be a mistake to deduce God's attributes from a finite and partial examination of nature: on the contrary, it is the assumption of God's omnipotence, omniscience and benevolence which makes the discordant world of appearances at once comprehensible and safe. Smith in this respect is thus conservative even with respect to his contemporaries.

Smith explicitly links the superiority of our natural feelings over the artificiality of reason, to the preservation of social order:

> "That kings are the servants of the people, to be obeyed, resisted, deposed, or punished, as the public conveniency may require, is the doctrine of reason and philosophy; but it is not the doctrine of Nature. Nature would teach us to submit to them for their own sake, to tremble and bow down before their exalted station ..." (**TMS**: I.iii.2.3)

The message is clear: the natural sentiments placed in us by a benevolent deity, expressed in established traditions, for example, of granting legitimacy to monarchs, are to be heeded in preference to whatever reason may tell us, so that social order may be preserved.

In conclusion of this sub-section, we may note how Heilbroner links the *philosophes*' promotion of sentiment over reason with the notion of an invisible hand in Smith:

> "all its [sc the Enlightenment's] leading thinkers – certainly Adam Smith – placed the 'passions' (feelings and emotions), not reason, at the center of human nature … Smith's critically important conception of an Invisible Hand – an indirect intervention of the Divinity into the mechanisms of social life – is based on the *inability* of human reason to achieve social harmony by itself." (Heilbroner, 1986: 2)

4.5.3 *Smith's contradictions*

There are many logical inconsistencies in Smith's theory[83], and we have noted some of them in passing. However, at base, there is one particular contradiction which confronts Smith, in various guises, at every turn. In his version of the stoic theory, everything is predestined for the maximisation of the 'quantity of happiness' in the world at every instant. In empirical reality, there is obvious suffering and injustice. How is the latter to be reconciled with the administration of the machine of the universe by a beneficent, omniscient and omnipotent god? To quote Hume: 'Epicurus's old questions are yet unanswered. Is he [sc God] willing to prevent evil, but not able? Then he is impotent. Is he able, but not willing? Then he is malevolent. Is he both able and willing? Whence then is evil?' (cited in Becker, 1932: 68).

Presumably even the most pious would concede that there must be *logical* restrictions on what a god can do – whether or not he can create a weight so heavy that he cannot lift it, for example, he is necessarily restricted to what is logically possible in what he can simultaneously achieve[84]. It is far less obvious, however, that suffering in general, let alone any specific instance of suffering, is a logical necessity for the achievement of God's presumed aims, and, indeed, Smith makes no attempt to put the case. Instead, its necessity for the good of the 'greater system' is simply asserted and assumed This is not a subtle point and neither is it new: it was certainly as well known in Smith's time as in ours that it was a problem for theories of this kind. He never addressed the issue, however, and failed to present any explicit theodicy going beyond these assumptions.

---

[83] 'Consistency was not his [ie, Smith's] shining virtue' (Macfie, 1959: 217). 'Smith's strengths lay in other directions than exactly logical thinking, and he displayed a fine tolerance for a generous measure of inconsistency' (Viner, 1958: 230).

[84] Thomas Aquinas concedes that 'although God's power is unlimited, he still cannot make … an unmade thing (for this involves contradictories being true together)' (cited in Rucker, 1995, p4; the reference is to Saint Thomas Aquinas, 1944, *Summa Theologiae* London: Blackfriars Ia, 7, 2-4), showing that he, for one, clearly accepted that God is constrained to what is *logically* possible.

Theodicy generally involves at some point an invocation of free will: God had to permit evil if he was to allow man free will and hence moral responsibility. Here again, Smith is on shaky ground, because he has made everything, including human nature, a part of nature; all behaviour, including human behaviour, is natural, and hence god-given. Our behaviour is prompted by the sentiments placed in our breast by 'a wise providence'. Since we do what we are led to do, what we are predestined to do, choice is presumably an illusion. Our judgement of the moral quality of an action, as we have seen, is for Smith essentially a sentimental judgement without rational content. Arguably, it was open to Smith to adopt the compatibilist position of Chrysippus, (whom, incidentally, Smith only mentions in order to flay, rather unfairly, as a traitor to, rather than exponent of, stoicism (**TMS** VII.ii.1.41)). But Smith carefully avoids addressing this issue, too, and the logic of his position, that we may admire 'the wisdom of God even in the folly of man', is surely that of determinism.

The problem for Smith is this: if God is maximising happiness, he cannot at the same time permit either evil and suffering or free will. If he allows suffering, then the quantity of happiness is presumably not at its logically possible maximum; if he allows free will, then he is again not maximising happiness, as he is leaving that to the outcome of the considerations of errant finite minds. Part of Smith's answer, no doubt, to both sides of this point, would be to say that God is not maximising the happiness of all *living* 'sensible and intelligent beings', but of the dead as well. Everyone has freedom of choice in their behaviour now, but they get their just deserts, in heaven, where also all unjust misery imposed in this world is undone. Happiness is maximised. However, this doesn't work. How does punishing sinners in the hereafter contribute to maximising happiness? It cannot have more remote consequences to counterbalance the pain of the sinners' tormented souls: it cannot deter sinners in this life; it is *futile* suffering.

Finally, the further consequence of the view that everything in the world is part of the great machine, playing its part in God's plan to maximise happiness, and that human nature and the behaviour to which man is led is a part of nature, is that regulation and state planning are just as natural and god-inspired as free trade and *laissez-faire*. Viner (1958: 233) asks, 'was not government itself a part of the order of nature, and its activities as 'natural' as those of the individuals whom it governed?' As Becker says,

> "if nature be the work of God, and man the product of nature, then all that man does and thinks, all that he has ever done or thought, must be natural, too, and in accord with the laws of nature and of nature's god. Pascal had long since asked the fundamental question: 'Why is custom not natural?' Why, indeed! But if all is natural, then how could man and his customs ever be *out of harmony* with nature?" (Becker, 1932: 66)

The concept of the natural only means anything – other than fatalistic acquiescence to anything and everything – if it is contrasted with something *else*, something *un*natural. This Smith attempts to do by referring to liberty as 'natural' and regulation as 'artificial' in **WN**, sentiment as 'natural' and reason as 'artificial' in **TMS**. But he cannot sustain this contrast on the basis of his theory. The category of the artificial has no meaning in a theory where the natural is already all-encompassing. This is clearly a critical contradiction for Smith's espousal of *laissez-faire*, but again, he makes no attempt to address the issue.

The contradiction can be seen particularly clearly in a paradoxical passage in **TMS** where he attempts, unsuccessfully, to reconcile his Panglossian view of the outcome of natural processes with the human attempt to remedy nature's faults. But if natural outcomes are the best which are logically possible, then such faults are inconceivable. Smith says that 'the general rules by which prosperity and adversity are distributed … appear to be perfectly suited to the situation of mankind in this life, yet they are by no means suited to some of our moral sentiments' (**TMS** III.5.9). In other words, God allocates prosperity by general rules which are designed to maximise human happiness, but the allocations which result, because of the finitude of human minds, do not always satisfy the moral sentiments which he has placed in us.

> "Thus man is by Nature directed to correct, in some measure, that distribution of things which she herself would otherwise have made. The rules which for this purpose she prompts him to follow, are different from those which she herself follows …. The rules which she follows are fit for her; those which he follows for him: but both are calculated to promote the same great end, the order of the world, and the perfection and happiness of human nature." (**TMS** III.5.9)

So nature[85] follows rules designed to maximise human happiness, and man, 'correcting' this, does the same. The inconsistency could not be clearer. If nature's rules lead to

[85] The 'nature' at issue here has, of course, nothing to do with fields and trees and flowers and wild animals: Smith is talking about the spontaneous outcomes of social processes, in particular the

optimising, happiness-maximising outcomes, then man's correction of nature must interfere with this and lead to a suboptimal outcome; if, on the contrary, man's correction of nature is happiness-maximising then nature's rules must themselves have been suboptimal. Smith cannot have it both ways. Or, rather, there is *one* interpretation which would allow him to have it both ways. If he were to say that nature *including* humanity were designed to optimise, but that nature *without* man were incomplete, imperfect, suboptimal, which is more or less what Hegel says, then he could reconcile both accounts. Then human action to correct spontaneous market outcomes and redistribute prosperity according to merit would be optimising as it would be the result of *both* the rules of nature and the rules of man.

To draw out the point, we may say that, while Smith's version of natural law formed a foundation for the invisible hand mechanism, it by no means follows that it undermines the case for a *visible* hand of state intervention. On the contrary, his *Weltanschauung* forms just as good a foundation for the latter as the former, and it is only Smith's prejudices, and not his theoretical system, which lead him to prefer one to the other. State intervention is a product of all the human strengths and frailties of those involved in the political process. On Adam Smith's account, those strengths and frailties are god-given and designed to lead individuals to act so as to maximise human happiness. There is nothing in the system of thought which Smith presents to say that the invisible hand active in the economic process will be inactive in the political process.

Smith cannot have been unaware of these inconsistencies in his standpoint. Yet there is a sense in which he, himself, is not inconsistent in neglecting them. Someone who kept faith with the Enlightenment ideal of following Reason wherever it may lead – a Ricardo, for example, a Marx, a Darwin, or an Einstein – would have concentrated attention on these contradictions and drawn the logical consequences. But we have already seen that Smith was not in this mould[86]. The late eighteenth century philosophers turned their back on reason and, instead, promoted sentiment. It was not Smith's goal to present an intellectually unified, logically coherent system of thought, but to paint as pleasing as

---

distribution of 'prosperity', or property, wealth and income in society.

[86] Writing of Smith, Macfie (1967: 126) says 'consistency was never the central aim or virtue of eighteenth century writers, especially of the Scottish sociological school.'

possible a picture of the world, such that the viewer would be 'animated to correspond with the general harmony of nature'.

4.6 Conclusion

The question we started with was, how Smith saw the articulation between individual behaviour at the micro level and social outcomes at the macro level. The answer I have given in this chapter is that the articulating mechanism consists in the agency of a deity. Our behaviours at the micro level are always just what is required for the optimal macro outcome because the deity's invisible hands always lead us, through the pursuit of our own interests, our own illusions and our own fellow feeling for others, to perform just those actions required to fulfil the divine plan. This is what Smith meant by the 'invisible hand'. The implication is that invisible hand theorists of more recent times, such as Hayek, to the extent that, as representatives of a secular age, they cannot rely on an interventionist god, need an alternative mode of articulation between levels. The most frequently invoked alternative – to the extent that the problem is addressed at all – is some kind of evolutionary mechanism, but that lies beyond the scope of the present chapter.

I have also argued in this chapter that Smith's 'invisible hand' mechanism is closely linked to the apologetic aspect in his political economy. While his belief in a harmonious universe allowed him to make real scientific progress in political economy, without fear that it would destabilise the social order, Smith's principal objective was, nevertheless, to reconcile humanity with the spontaneous social order and the status quo[87]. He invoked the idea of a divine teleological plan, of the universe as a machine administered by a god, in order to explain away suffering and evil as only the proximate manifestations of chains of connection whose distant ramifications would include more than compensatory benefits. The idea is to convince us that we need do nothing at the macro level. All we should do is pursue our own individual interests at the micro level, and display appropriate levels of patriotism and respect for our leaders. The rich, the powerful and the fortunate all ensure that the big decisions of society are for the best – because they

[87] That these, the spontaneous social order and the status quo, were not the same thing, did not in Smith's time present the acute problem it presented for subsequent writers, such as Ricardo, coming after the Industrial Revolution.

are taken by the hand and led by God to do so. All is for the best, then, in this, the best of all possible worlds.

But does Smith not 'protest too much'? Sometimes Smith's protestations seem to invite the speculation that the truth is just the opposite of what he says. Smith claims that the universe is a coherent and harmonic whole administered by a single intelligence. But we know that this is not the case. The world is a jungle, an arena of clashing interests: 'It is as though cheetahs had been designed by one deity and antelopes by a rival deity.'[88] (Dawkins, 1995: 123) Smith claims that human nature and human society are a part of this organic unity, 'all discord, harmony not understood'. But, of course, society was as riven by sectional interest then as it is now. His claim is to be understood, not as a positive statement of what is the case but as a normative statement of what is to be desired. He claims that spontaneous human institutions, 'the result of human action but not human design', such as the market, and the law, order and defence functions of the state, make an optimal contribution to human welfare because guided by the invisible hand of a beneficent, omnipotent and omniscient god. Again, we know of no reason to even suspect that any supernal agency exists, such that we can rely on its intervention to maximise social welfare[89]. Again, perhaps, Smith's claim is to be understood in a normative sense: what is required is a higher level *human* agency which will reconcile our differences and lead us through the pursuit of our own interests to the maximum achievable level of welfare:

> "the invisible hand is only one of the many names given in the *Moral Sentiments* to the Deity – great Author of Nature, Engineer, Great Architect, and so on .... Adam Smith did believe (as a matter of faith) in this final reconciler .... Now, there is little doubt that we today do not accept this kind of argument .... The inevitable reaction is that, if the supernatural control is abandoned, human societies must supply their own .... [T]he state ... must take the place of the invisible hand." (Macfie, 1967: 111)

[88] 'Alternatively, if there is only one Creator who made the tiger and the lamb, the cheetah and the gazelle, what is He playing at? Is He a sadist who enjoys spectator blood sports?' (ibid)

[89] And, even if there were such a power, some might argue, passing up all responsibility to it for our own actions and their consequences in this fashion, might scarcely be the best method of winning its approval.

Chapter 5 Friedrich Hayek: a Panglossian evolutionary theorist[90]

> "[C]omplex phenomena ... can be made intelligible only by ... a cosmology, that is, a theory of their evolution." (**SIP**: 76)

5.1 Introduction

Chapters 2 and 3 of this thesis looked at the mid- and late-twentieth century response of political economy to two anomalies which have been perceived as challenges to the invisible hand hypothesis: the prisoners' dilemma and Arrow's impossibility theorem. In Chapter 4 attention returned to the roots of this tradition in the writings of the eighteenth century father of nineteenth and twentieth century economics, Adam Smith. The question addressed there was, how Smith saw the articulation between individual behaviour at the micro level and social outcomes at the macro level. The answer I gave is that the articulating mechanism consisted in the agency of a deity. Our behaviours at the micro level were always just what was required for the optimal macro outcome because that invisible deity always led us by the hand, through the pursuit of an amalgam of our own interests, our own illusions and fears, and our own fellow feeling for others, to perform just those actions required to fulfil the divine plan. This is what Smith meant by the 'invisible hand'.

The implication I drew is that invisible hand theorists of more recent times, such as Friedrich Hayek, to the extent that, as representatives of a secular age, they cannot rely on an interventionist god, need an alternative mode of articulation between levels. The most frequently invoked alternative, in so far as an explicit alternative is presented at all,

[90] An earlier version of this paper appeared as Denis (1999b), and at point of writing material derived from it is under consideration at *Constitutional Political Economy* as Denis (2001b).

is some form of evolutionary mechanism. This amounts to replacing one form of Panglossianism with another. In replacing God with evolution we move from justifying the belief that 'what is, is best' by the claim that it was selected by God, to justifying it by the claim that it was selected by nature. We remain within the Leibnizian paradigm, that this is the best of all possible worlds, only replacing the explanation that it was selected by God all at once with the explanation that it was selected by Nature over a long period of time.

The purpose of the present chapter, therefore, is to investigate the deployment by Hayek of an evolutionary mechanism to argue that spontaneous aggregate level outcomes of our activity are intrinsically superior to any outcome we could achieve by conscious intervention at the macro level.

The thesis of the chapter will be that, whatever Hayek's intention, his mode of procedure in fact distorts the Darwinian theory of evolution and falsifies the standpoint of Adam Smith. This distortion and falsification then serve Hayek's goal of underpinning a pre-established policy prescription. He falsely claims that the Smithian economists can be described as 'Darwinians before Darwin', and that Darwin made his name simply by applying in the biological field the ideas that the Smithians had already established in social science. He systematically downgrades Darwin's contribution – apparently in order to render respectable a theory of social evolution which leads to *laissez-faire* conclusions. A theme of the chapter will be the care with which we have to read Hayek. Frequently we will find him saying one thing and doing another. Statements about his standpoint cannot necessarily be taken at face value. Two examples we will meet concern his supposedly individualist methodological stance, and contradictory statements as to the nature of the Darwinian theory of evolution.

As with the chapter on Smith, an initial caveat is in order. This chapter is in no way to be construed as an attempt to give an all-sided consideration of Hayek's contribution. On the contrary, the focus is on the specific concern of the research of which the chapter, and indeed the thesis, forms part: the question, that is, of how economic theorists have linked micro and macro levels, how these levels are coordinated or articulated. Given, as I argued in the previous chapter, that in Adam Smith individual behaviours are pre-coordinated by the invisible hand of a wise and benevolent Providence, and further that

the more secular twentieth century could not be expected to accept such an explanation, what alternative mechanism to the hand of God can be invoked by today's advocates of the invisible hand? With respect to Hayek, therefore, I am here concerned only with this one question: How viable is Hayek's evolutionary articulating mechanism as an alternative to Smith's invisible hand of God? The thesis proposed here is that, Hayek's theory is as Panglossian as Smith's – indeed, the two are in many ways very similar[91] – and that his obsession with countering 'socialism' and planning leads him, in effect if not intent, to major factual and theoretical distortions.

5.2 Hayek and Smith

As we saw in the previous chapter, twentieth century Smithians have had trouble with Adam Smith's reference to 'an invisible hand', at least in part because anyone who actually read Smith's work (and not just **WN** in isolation) could see that what Smith was referring to was not an 'analogy', a 'metaphor', but quite literally the hand of God: an unpalatable notion in a relatively secular age. Some instances of this anxiety about the phrase 'the invisible hand' on the part of modern claimants of the Smithian heritage were given in the last chapter.

Hayek was one of those who made it clear that he regarded the phrase as unfortunate. Smith, admittedly, had used the expression 'the invisible hand', but, according to Hayek, he had nevertheless also given an explanation of the phenomenon: 'Adam Smith and the other great Scottish individualists of the eighteenth century – even though they spoke of the "invisible hand" – provided ... an explanation [of how the interaction of the efforts of individuals can create something greater than they know]' (**CRS**: 392-393). Unfortunately, the explanation that Hayek refers us to is for something different: he has illicitly changed the subject.

The invisible hand – God's Hand – was the mechanism in Smith which ensured the perfect reconciliation of unconstrained individual motives and behaviour both with each other and with the social interest of maximising human welfare. Two strands can be discerned in this thought: spontaneous order and optimality. The spontaneous order

[91] Hutchison says that after the mid-30s, Hayek adopted 'a comprehensively "Smithian" approach'

strand says that individual self-seeking behaviour, unconstrained by central authority, may result, not in chaos, but in orderly collective behaviour. The optimality (efficiency, desirability) strand says that this order will be, in some sense, the best that we can get. Clearly, these are very different propositions. The spontaneity proposition is, I think, undeniable, while the far stronger optimality proposition is just false. If Smith had confined himself to the former, that would have been unexceptionable – and there would have been no call for divine intervention, nothing for an invisible hand to do. The consequences of individual action would just be the consequences: orderly but often sub-optimal collective behaviour. No explanation of the transmutation of base passions into golden outcomes would be necessary as no such transmutation would be assumed to take place.

Now the optimality proposition clearly encompasses the spontaneity proposition, and hence, when Smith attempts to sustain the former, he necessarily defends the latter. The bulk of **WN** is concerned with this defence of the idea of a spontaneous order. Optimality and the invisible hand are there but they tend to be implicit. It is therefore unsurprising that economists, reading Smith through nineteenth and especially twentieth century spectacles, have misunderstood 'the invisible hand' as a figure of speech representing the spontaneous emergence of *some* order, rather than, what it is, an assertion of the *optimality* of that emergent order. Hayek, however, has pretensions to be – more than merely a technical economist – an intellectual and a scholar. He has written widely on the Smith-Hume-Ferguson-Burke tradition – the 'antirationalist tradition' (**COL**: 61) – from which he claims intellectual descent. It is illegitimate for him to slur over the difference between spontaneity and optimality. The question is, how Smith explains his 'invisible hand', and the answer simply: God. Hayek, however, takes the question and immediately reduces it to the lesser, more innocuous question of the explanation of spontaneous order 'how the interaction of the efforts of individuals can create something greater than they know'. This might be relatively harmless, if Hayek then were satisfied with having supported the spontaneity proposition, but this is not what he does. Throughout his writings, Hayek adopts the same procedure: firstly, focus attention on the spontaneous order, and then slide over to an assumption of its optimality.

(Hutchison, 1981: 228 n 17).

The 'spontaneous order' in Hayek thus has a dual function. On the one hand, it is what it says it is, an account of the spontaneous emergence of orderly social behaviour in the absence of prior design and central direction of individual activity. On the other, it has also to smuggle in the idea of optimality, to act as a secular replacement for Smith's invisible hand of God. It is clear that God has to be replaced in the story: the question is, with what? Hayek's answer is that the role of God in sustaining the optimality of the spontaneous order is to be played by evolution.

> "If, in the form in which Adam Smith put it, the phrase that man in society 'constantly promotes ends which are no part of his intention'[92] has become the constant source of irritation of the scientistically minded, it describes nevertheless the central problem of the social sciences. As it was put a hundred years after Smith by Carl Menger, who did more than any other writer to carry beyond Smith the elucidation of the meaning of this phrase, the question 'how it is possible that institutions which serve the common welfare and are most important for its advancement can arise without a common will aiming at their creation' is still 'the significant, perhaps the most significant, problem of the social sciences.'[93]" (**CRS**: 146-147)

In this passage Hayek sets up the problem he is going to use a theory of evolution to solve. That problem is how individuals pursuing their own goals fulfil social objectives about which the individuals neither know nor care. This is the problem of the emergence of a spontaneous order. But Hayek immediately identifies this with the problem of the emergence of institutions which 'serve the common welfare'. This is the optimality assumption. This says that individuals fulfil social goals, and those goals are just the ones which serve the collective interests of the individuals in the society, the 'common welfare'. It is *assumed* by Hayek (and Menger) that this is so, the big question being not *whether* but *how* this comes about. The answer will be 'by means of evolution'.

According to Haworth, Hayek's main thesis, 'the thesis of 'spontaneous order' ... is the most subtle, interesting and credible version of the invisible hand argument there is' (Haworth, 1994: 114). Hodgson, on the other hand, says that Hayek's 'conception of socioeconomic and cultural evolution is the centrepiece of his mature theory' (Hodgson,

---

[92] Hayek gives no reference here, and the phrase as it stands does not appear in Smith. A similar passage is in **WN** IV.ii.9: the capitalist (not 'man in society'), Smith says, 'is in ... many ... cases, led by an invisible hand to promote an end which was no part of his intention'.

[93] The reference is to Menger (1883, LSE reprint 1933) *Untersuchungen über die Methoden der Sozialwissenschaft* p 163; trans Hayek.

1993: 153). These accounts seem to point in different directions, the one highlighting the 'invisible hand' aspect of Hayek's thesis of spontaneous order, the other focusing attention on the evolutionary aspect. In reality both are right, since the whole point of Hayek's evolutionary theory of spontaneous order is to provide an invisible hand explanation of that order. As we saw in the previous chapter, the deification of nature by the eighteenth century philosophers, including Smith, led – indeed, was intended to lead – to the sanctification of the particular model of human behaviour that they wished to hold up as natural. In just the same way, the notion of evolution deployed by Hayek is intended, not to provide a scientific understanding of the social order, warts and all, which has emerged from a blind evolutionary process[94], but to present that order as something with which it is beyond our competence to interfere. Flew gets the ideological, almost theological role of evolution exactly right in his discussion of Social Darwinism:

> "many people are inclined to believe, that whatever is in any sense natural must be as such commendable, and that Nature is a deep repository of wisdom, [so] for many the process of evolution by natural selection becomes a secular surrogate for Divine Providence; and ... for some the possibility, or even the duty, of relying on this benign and mighty force presents itself as a decisive reason why positive social policies must be superfluous, and may be wrong – indeed almost blasphemous!" (Flew, 1967: 15)

'Evolution', in this context, is just the latest stage in the evolution of natural law.

5.3 Holism and reductionism in Hayek

5.3.1 *Shenfield on collectivism and holism in Hayek*

It is a commonplace that the methodological standpoint of the Austrian school, including and, perhaps, especially Hayek, what they insist upon with a fundamentalist zeal which distinguishes them from their more pragmatic neoclassical cousins, is the reductionist principle of methodological individualism. See, for example Garrison and Kirzner (1989: 121-122), and Hodgson (1993: 153-157). With regard to Friedrich Hayek, this

[94] As Dawkins points out, evolution is a product of a universe in which 'there is, at bottom, no design, no purpose, no evil and no good, nothing but blind, pitiless indifference' (Dawkins, 1995: 155). To explore the consequences for our institutions of such a blind, pitilessly indifferent process in a social context would be a worthy enterprise indeed; needless to say, in my reading, this is not Hayek's goal.

commonplace is false: 'it may be that 'Hayek is by no means the champion of methodological individualism that he claims to be,' as Stephan Böhm (1989[95], p. 221) alleges. He is more a systems thinker: one moreover with strong traces of functionalism' (Hodgson, 1993: 157).

Hayek's standpoint is a combination of methodological holism and policy individualism. This is not to say that Hayek is entirely consistent or that it is impossible, with care, to pick out a methodological individualism in what he says. On the contrary, perhaps even more than Smith, whose legacy he claims, Hayek seems to have taken literally the enormously arrogant dictum of Emerson, that 'a foolish consistency is the hobgoblin of little minds'[96]. As we shall see, where the internal tensions of his standpoint, the stresses of combining holism and individualism in this way, prove too severe, consistency is the first casualty.

The profound intellectual disarray of the Austrian school on the question of holism and reductionism is shown by a remarkable passage in Shenfield (1977). The context for this needs to be made explicit. Machlup (1977a) is a book of *Essays on Hayek*, presented at a conference of the Mont Pelerin Society in 1975, devoted to the achievements of Friedrich Hayek. The Mont Pelerin Society was set up by Hayek and presided over by him for 12 years, after which he became its honorary president. Perhaps needless to say, the only hint of debate the book contains is on the level of 'Where do Hayek's greatest achievements lie?' (Machlup, 1977b: 50). Shenfield's essay is billed as an 'appraisal of Hayek's innovative work on the methodology of the social sciences'. (Machlup, 1977a: front flap). We may take it, therefore, that Shenfield's pronouncements here have something of the nature of a quasi-official Austrian statement of Hayekian methodology[97].

[95] Either the date or page number of this citation in Hodgson is incorrect.

[96] A saying which is reminiscent of an American millionaire's remark, when accused of tax evasion, that 'taxes are for the little people'.

[97] Shenfield also read the final text of Vol III of **LLL** and 'corrected there a variety of substantial as well as stylistic points' – showing, again, the close links between Hayek and Shenfield (**LLL**: xxi).

The topic of Shenfield's essay is Hayek's critique of 'scientism'[98], the allegedly inappropriate attempt to apply natural scientific methods in the social sciences. He argues that scientism depends on an unholy trinity of objectivism, collectivism and historicism. It is with the second of these that we are concerned. Collectivism, he says, is the same thing as holism: '*Collectivism* (perhaps better called holism ...) ... treats as wholes conventional constructs like "the economy" or groups like nations or classes, as if they were each invested with a single mind and acted accordingly' (Shenfield, 1977: 68).

This confused formulation says at least three things. Firstly it correctly recognises that holism, as its name suggests, attempts to grasp entities as a whole, rather than as a collection of parts, as a unity rather than a plurality. Secondly, however, focusing on particular social entities such as the economy, it suggests that the referent is not in fact an entity but a 'conventional construct', and the implication is clearly that holism is a misleading way to see things. Lastly, Shenfield suggests that holism ascribes rational behaviour (a 'single mind' which 'acts accordingly') to aggregate level social entities, '[p]artly,' he writes, 'collectivism is the result of the use of language for a system of relationships similar to that applied to living things' (ibid) – the implication being that such attribution is fallacious. This point is neither trivial nor superficial. It is interesting that already an inconsistency has crept into Shenfield's account: what only a few lines previously had been dismissed as a 'conventional construct' is now admitted to be 'a system of relationships'. And Hayekians are certainly ready to use language 'similar to that applied to living things' where it suits their purpose to do so:

> "Our social institutions, customs and rules have ... grown in an evolutionary way, persisting where they prove useful and fading out where they are not. The result is that society – like the physical structure of animals – has evolved to a point of much greater complexity than the human mind can understand, let alone attempt to redesign." (Butler, 1983: 7 – in 'Foreword' to **KES)**

There is a clear theoretical distinction between holistic accounts which view aggregate level entities as being, in some sense, 'organic', and those which regard them as organisms. The latter implicitly ascribes aggregate entities a greater degree of autonomy. This is both a bigger and a more specific claim. Both views, however, tend to be labelled as 'organicist'. Keynes, for example, clearly adopted an organicist standpoint, in which organic links between individuals lead to macroscopic effects of individual actions

[98] What Popper calls 'dogmatic methodological naturalism' (Popper, 1957: 60).

(Carabelli, 1988). Marx would undoubtedly have agreed, but went much further than Keynes, arguing that macro level entities, such as capital and the state, had the autonomy and self-interest of organisms: 'It is a great step forward [for Hegel] to have seen that the political state is an organism and that, therefore, its various powers are no longer to be seen as [merely] organic' (Marx 1843/1975: 66). Hayek is clearly in the Marxist, rather than the Keynesian camp on this particular issue: 'the state … ought to be only a small part of the much richer organism which we call "society"' (**IEO**: 22).

Shenfield's treatment of organicism here neglects this distinction and consequently is too extreme: not all versions need see all social level entities as organisms, though, to be sure, some certainly do see some such entities in this light. Adam Smith, for example, believed the world was an organism[99] with every part organically linked to every other and subordinated to the task of maximising human happiness. The world was ruled by a single mind, that of God, whose will was executed by the invisible hand. While we can share Shenfield's scepticism with regard to this kind of organicist fairy story, that does not mean to say that we should discard holism or organicism *sans phrase*. On the contrary, while each holism must be judged on its merits, according to its ability to identify the principle connections between the elements of the entity in question, reductionism can be rejected *a limine* since it itself rejects *a limine* the relevance of those connections.

Having said that collectivism was partly caused by organicism, Shenfield now goes on to examine the roots of collectivism in holism: 'Partly collectivism arises from the essential belief of philosophical holism, namely that wholes are more than the sum of their parts, and that the parts are less real than the wholes, being largely abstract analytical distinctions' (Shenfield, 1977: 69). While not being quite how an advocate of the holist standpoint might put it, this is not an unrecognisable description of holism. It would be

[99] That his *imagery* was mechanical rather than organic is a reflection of the idiom of the time, particularly the influence of the Newtonian scientific revolution, and has no bearing on the content of his theory in this respect. For centuries people have employed the trope of referring to animals as machines without there being any suggestion that they misunderstood the nature of an organism. Two examples spanning the last half-millennium are Leonardo da Vinci: 'A bird is an instrument working according to a mathematical law' (McCurdy: 1938); and Richard Dawkins's famous portrayal of individual organisms as 'gigantic lumbering robots' (Dawkins, 1989: 19).

more correct to say that, in the holist view, wholes are *different from* rather than *more than* the sum of their parts. It is not merely a quantitative relation. 'More than' is correct if used in a qualitative sense, such as in 'a home is more than a house'. And from a holistic point of view 'my leg' *could*, at a pinch, be construed as, in a sense, 'less real' than I am (if one really wants to start measuring things along a dimension of 'reality'), in that it does not and cannot exist without me, as a functioning leg, but I could exist without it, though less ably than with it. But this is not the main thrust of holism, which is merely to argue for seeing entities as wholes, as systems of relations, rather than as congeries of isolated parts. Hayek himself is clear on this:

> "That a particular order of events or objects is something different from all the individual events taken separately is the significant fact behind the [phrase of] ... 'the whole being greater than the mere sum of its parts' ... [I]t is only when we understand how the elements are related to each other that the talk about the whole being more than the parts becomes more than an empty phrase." (**TSO**: 47) "The overall order of actions in a group is ... more than the totality of regularities observable in the actions of the individuals and cannot be wholly reduced to them ... a whole is more than the mere *sum* of its parts but presupposes also that these elements are related to each other in a particular manner." (**SIP**: 70)

These are fine and unambiguous statements on Hayek's part of a holism entirely consistent with the definition of the term used in this thesis. So Hayek says that the whole is greater than the sum of the parts: to understand the entity in questions we have to understand the system of relations between its parts. And Shenfield says that believing the whole is more than the sum of the parts is holism. So presumably we can agree that Hayek's methodological approach is a holistic one. But no: according to Shenfield, holism is just what Hayek is combating. Or is it? Having identified holism as the philosophical basis of collectivism, itself one of the three legs supporting scientism, the target of Hayek's methodological critique, Shenfield immediately proceeds to admit that the holistic view is correct:

> "Of course it is not to be denied that the whole of a society is more than what the sum of its individuals would be if they had no contact with each other. Such a sum would not be a society at all. A society is not a collection of hermits. It is formed because individuals set up relations with each other. It is then not more than the total of its interconnected individuals, their systems of connection being a vital element in it." (Shenfield, 1977: 69)

So society is after all different from, 'more than', the sum of all its members *taken in isolation*. But that is exactly what holism says: the whole cannot be understood except on the basis of the 'systems of connection' of its parts. Shenfield simply concedes the

case. It is interesting, incidentally, to contrast Shenfield's view that 'society is not a collection of hermits' with Milton Friedman's previously mentioned view of economics as based on the study of 'a number of independent households – a collection of Robinson Crusoes' (cited in Haworth, 1994: 8). On this particular point, Friedman adopts a reductionist, and individualist, stance, while Shenfield, and, indeed, Hayek, adopt a holistic standpoint.

One of the distinguishing features of the Austrian tradition separating it from the neoclassical orthodoxy is supposed to be its greater emphasis on methodology, its more advanced epistemological self-consciousness, its greater sensitivity to the need to establish the philosophical preconditions for the practice of economics. What can we say, therefore when Shenfield, having conceded the case for holism, goes on in the very next sentence to conclude his critique of collectivism with

> "Thus to treat the whole as a reality and the individuals as an unreality, or as a lesser reality, is a fundamental misconception. As Popper says, '... the holistic way of thinking (whether about 'society' or about 'nature') ... is characteristic of a pre-scientific age'[100]" (Shenfield, 1977: 69)

Although introduced by 'thus', neither of these statements flows from anything that has preceded them. No case at all has been presented for regarding the holist standpoint as 'a fundamental misconception', and no reason given for regarding it as 'pre-scientific'. On the contrary, as noted above, where he actually addresses the issues, he concedes the case. The reference to the lamentable discussion of holism in Popper (1957: 76 ff) does nothing to help.

In an epilogue to his discussion of collectivism, a paragraph on collectivism and measurement, Shenfield manages to combine holism and opposition to it in a single sentence:

> "collectivism leads to an exaggerated belief in the importance of measurement as a mark of scientific status. But by treating the objects of its study as wholes, it is led to subject to measurement almost anything except their essence, namely the systems of connection between individuals which become visible only when the concept of wholes is abandoned." (Shenfield, 1977: 69)

[100] The reference is to Karl Popper (1944) 'The Poverty of Historicism', *Economica* XI (new series),

Shenfield says here, firstly, that the 'essence' of the object of study is 'the systems of connection between individuals', and then that these systems 'only become visible when the concept of wholes is abandoned'. He neglects to explain what might be the difference between 'wholes' and 'systems of connection'. The holistic view, the systems view of the world, says that we should examine a 'system of connection between individuals' as a whole, not as a congeries of parts, each considered in isolation, apart from its relationship with the whole. How such a system can only become visible when we *cease* to see it as a whole is just incomprehensible. This confusion is adopted almost word for word from Popper (1957: 76).

When Shenfield criticises 'collectivists' for measuring 'anything except their essence' (ie, the essence of the objects of study), the implication is clearly that that essence is the one thing that they should be measuring. This begs the question of how one is supposed to measure an essence. It also inverts Hayek's reason for scepticism about measurement in social science, which was basically a holistic one. According to Hayek,

> "The events which we must take into account in any attempt to predict the outcome of particular social processes are never so numerous as to enable us to substitute ascertained probabilities for information about the individual events .... in the biological and in the social sciences frequently we cannot rely on probabilities, or the law of large numbers, because unlike the positions which exist in the physical sciences, where statistical evidence of probabilities can be substituted for information on particular facts, we have to deal with ... organized complexity, where we cannot expect to find permanent constant relations between aggregates or averages." (**KES**: 25)

Here Hayek is making an essentially holistic point: the law of large numbers depends on the independence of the events in question; that is, that the mass of events is a congeries. In social matters, however, we have 'organized complexity' where the micro-level events we are concerned with are connected to each other in an organic way: they are not independent, and hence statistical inference is invalid.

To say that Hayek's opposition to measurement in social science is based in a holistic outlook, is not, however, to say that it is right. On the contrary, what this view ignores is that a principal feature of 'organized complexity' is, precisely, that it is organised, that is, that it displays a constancy or consistency over time in some of its key internal variables,

---

August, 126.

in spite of changes in environmental variables. This is summed up both in the concept of homeostasis and in Hayek's own concept of an order.

This, again, is something which Hayek is well aware of. In a subsection of *The Sensory Order* concerned with the evolution of the sensory order, Hayek briefly considers the (reductionist) argument that 'any attempt to explain the highly complex kind of purposive action made possible by a developed central nervous system may be premature so long as we do not possess a fully adequate biological theory of the comparatively simpler kind of purposive functioning.' (**TSO**: 82) His response to this view is to refer in this connexion to 'W.B. Cannon's concept of homeostasis and ... the most promising work of L. von Bertalanffy. His theory of 'open systems' in a steady state (Fliessgleichgewicht[101]) in which 'equifinality' prevails because the equilibrium that will be reached will in some measure be independent of the initial conditions, seems to provide the most helpful contribution to this problem' (**TSO**: 83).

This is essentially both a systems theoretical account of order emerging at the macro level of purposive behaviour and an assertion of our ability to understand it even in the absence of a 'fully adequate' theory of the substrate biological level. The idea of a dynamic steady state immediately suggests that certain variables will be in stable long-run mutual relationship, which in turn suggests the suitability of appropriate mathematical and statistical techniques. The body of techniques including cointegration, unit roots and error correction mechanisms springs to mind.

The only conclusion that can be drawn from Shenfield's account is thus that he, and Hayek with him, is extremely hostile to holistic approaches to economics, while semi-covertly recognising that such approaches are methodologically sound. Their opposition is due to the policy consequences which such approaches may entail – individual utility maximising behaviour is interdependent, and hence may not aggregate to collective welfare maximising outcomes. The tensions implicit in this inconsistent standpoint soon begin to emerge when Shenfield attempts to apply this approach to the question of the validity of macroeconomics. We should bear in mind that Friedman wrote the Foreword to the book (Shenfield, 1977: xxi-xxiv), and also chaired the session preceding

[101] ie, flow equilibrium.

Shenfield's, so Shenfield needs to be careful to avoid treading on toes. The question, then, is 'a problem which Hayek's *Scientism* did not deal with, namely the question of the legitimacy or virtue of macroeconomics. Is not the whole of macroeconomics infected with the collectivism that Hayek condemns?' (Shenfield, 1977: 69)

Now objectively the answer to this question is 'yes', and Hayek himself is absolutely clear on this:

> "I still believe that this [sc microeconomics] is the only approach which is entitled to regard itself as scientific .... [M]icroeconomic theory [is] the only legitimate economic theory." (**KES**: 21-22) "[M]y disagreement with that book [sc Keynes's *General Theory*] did not refer so much to any detail of the analysis as the general approach followed in the whole work. The real issue was the validity of ... macro-analysis". (**TBT**: 100)

Shenfield, however, is more cautious. He examines two aggregate level entities, national income and the general price level, for evidence of holism. While the concept of national income is – rather grudgingly – acquitted, the general price level, Shenfield finds, is 'a holistic fiction because it sets up an imaginary whole which is different from the individual transactions that are supposed to form it' (Shenfield, 1977: 70[102]). This immediately creates a major difficulty as conservative neoclassical economists, in particular the monetarist school of thought around Milton Friedman, have depended upon the quantity theory of money, which itself depends on a notion of the general price level. '[T]he concept of the general price level, and the quantity theory of money which goes with it [are] essentially holistic' (Shenfield, 1977: 70-71). How, then, to avoid division in the conservative ranks? How to avoid castigating Friedman as a rank collectivist – as Hayek's prejudices and his supposed theoretical position on the nature of macroeconomics would, indeed, imply? This is Shenfield's attempted way out:

> "one has to admit a qualification to Hayek's condemnation of collectivist notions .... there may be a few collectivist ... concepts which, when used by those who know the pitfalls of collectivism, may be enlightening. Macroeconomics can have value, but only in the hands of those who are first and foremost microeconomists." (Shenfield, 1977: 71)

Shenfield seems not to notice what he has said here. He has made the value and significance of concepts such as the general price level and the quantity of money

[102] Keynes goes further and finds major difficulties with both concepts, such that he decides not to use them in the *General Theory* (**GT**: 37-40).

depend, not on their content and inner logic, but on the personal characteristics of the theorist deploying them. If we were to take this approach seriously, all debate about ideas must cease and be replaced by discussion of personalities. It would be interesting to know what Hayek and Friedman made of this attempted reconciliation of their theoretical standpoints.

It is also interesting to note that, whereas earlier, on *methodological* terrain, we could characterise Hayek as adopting a holist and Friedman a reductionist stance, they have now – at the level of *theory* – swapped places. Friedman's stance is holist, dealing in entities such as the price level which only emerge at the macro level – something which Hayek's anti-holist prejudice (though not his real methodology) makes impermissible. This switch demonstrates how profoundly unconcerned these writers are with maintaining methodological and theoretical consistency. Both want to retain policy individualism in a holist world: where they differ is on where to make concessions with the world they actually inhabit. Friedman's strategy is to rely on a non-Keynesian macroeconomics to underpin a *laissez-faire* policy prescription, while to the same end Hayek would prefer to proscribe macroeconomic thought altogether.

5.3.2 *Hayek and holism*

Early in the final version of his *magnum opus*, Toynbee (1972), sets out definitions of his main terms. Perhaps the most important is his definition of 'society':

> "SOCIETY is the total network of relations between human beings. The components of society are thus not human beings but relations between them. In a social structure 'individuals are merely the *foci* in the network of relationships'. The famous frontispiece of Hobbes's *Leviathan*, displaying society as a gigantic human figure composed of a multitude of life-sized human figures, is an anthropomorphic misrepresentation of reality; and so is the practice of speaking of human beings as 'members' of society or one or other of its component institutions (e.g. a club, a church, a class, a family, a 'corporation'). A visible and palpable collection of people is not a society; it is a crowd. A crowd, unlike a society, can be assembled, dispersed, photographed, or massacred." (Toynbee, 1972: 43)[103]

This is a vivid presentation of a holist view of society. But what is interesting here for our present purposes is that the writer he cites, explaining that 'individuals are merely the

[103] For a similar criticism of the frontispiece to *Leviathan*, see Haworth (1994: 13).

foci in the network of relationships' of which society is composed, is none other than Friedrich Hayek (**CRS**: 59).

Further evidence of Hayek's fundamentally holistic methodology, in the context this time of the theory, not of society, but of mind, can be found in his work on theoretical psychology, *The Sensory Order*. It is extremely significant that **TSO** was read in draft, and commented on, by Hayek's friend, von Bertalanffy, and that there are several favourable references in the text to Ashby, Bertalanffy and Wiener, the pioneers of cybernetics and systems thinking, and to Cannon, a pre-cursor of cybernetics and inventor in 1932 of the term 'homeostasis'. Systems thinking is, as the name implies, a holistic standpoint emphasising the systematic interconnection of the substrate elements without considering the properties of those elements in isolation (Pask: 1961: 13). All of this underpins Hodgson's (1993: 157) remark cited earlier that methodologically Hayek is more a systems theorist than an individualist[104].

At the end of **TSO**, in the chapter on 'Philosophical Consequences', section on 'Dualism and Materialism', Hayek says the following:

> While our theory leads us to deny any ultimate dualism of the forces governing the realms of mind and that of the physical world respectively, it forces us at the same time to recognize that for practical purposes we shall always have to adopt a dualistic view .... [A]ny explanation of mental phenomena which we can hope ever to attain cannot be sufficient to 'unify' all our knowledge, in the sense that we should become able to substitute statements about particular physical events (or classes of physical events) for statements about mental events .... [W]e shall never be able to bridge the gap between physical and mental phenomena; and for practical purposes ... we shall permanently have to be content with a dualistic view of the world." (**TSO**: 179)

We should bear in mind that at the beginning of the book, in defining his terms, he refers to these 'realms' or 'orders', the objective or physical and the subjective or sensory orders respectively, as the 'microcosm' and the 'macrocosm' (ibid: 4[105]). Taking this into

[104] See also **LLL**: xviii-xix.

[105] Later in the book, Hayek forgets that he has defined the microcosm and the macrocosm in this way, and refers to them the other way round – the macrocosm as the physical order and the microcosm as the sensory order (**TSO**: 108, 127). The logic for this seems to be that the physical world, the macrocosm in this alternative definition, incorporates everything, including the microcosms of the sensory orders of organisms within it. This is just one of a host of errors and inconsistencies which will irritate and frustrate the reader of this book. A substantial proportion of the index entries are incorrect and the

account, the passage just cited is a lucid and succinct account of the holist standpoint applied to psychology. Hayek's vision is one in which macro level phenomena, the 'macrocosm', are in principle, if only we knew all the facts, reducible to the micro level substrate, or 'microcosm'; practical obstacles to knowing the detail requisite for this reduction, however, compel us to understand the macro level in its own terms. There is no hint of idealism or mysticism here: 'The order which we call mind is thus the order prevailing in a particular part of the physical universe – that part of it which is ourselves.' And: 'this order is formed by its physical elements' (ibid: 178). There is no 'ultimate dualism' but *in practice* we are forced to recognise a dichotomy between lower and higher levels, between 'microcosm' and 'macrocosm'. We can never hope fully to reduce mental phenomena to their micro level substrate in 'particular physical events'.

This section has established an important point, namely a clear recognition on Hayek's part, of the holistic nature of the world in which we live. This recognition is particularly important as it throws into sharp relief his assertion of individualist and reductionist methodological conceptions when policy issues loom. It also exposes a key aspect of his overall procedure: that, where necessary to defend his *laissez-faire* policy prescription, he is prepared to say one thing while doing another. Hayek's political philosophy is one of individualism, and to underpin that he realises he needs to assert an individualist methodology. However, such a methodology faces insurmountable incongruities with the way the world is: *at some stage* the fundamentally holistic nature of the world has to be taken on board if our interaction with it is to have any efficacy whatever. Hence the fundamental ambiguity in Hayek's methodology is the consequence of his partisan policy standpoint. In the next section we will see how this basic contradiction in Hayek's approach manifests itself in connection with his theory of evolution. We will find that something very similar occurs here, too: Hayek says one thing and does another – he is well aware of the implications of the Darwinian theory, but falsifies it when it clashes with his desired policy outcome.

5.4 Hayek and evolution

German orthography is haphazard. Elsewhere (**CRS**: 72), in an extended analogy between the physical and social domains, Hayek uses the terms *microcosm* and *macrocosm* to refer to the atomic substrate and aggregate outcome levels respectively.

5.4.1 *Darwinian evolution*

Hayek's theory of evolution can only be understood on the basis of its rôle in his overall intellectual project, and, hence, an understanding of the nature of that project itself. Hayek's theoretical objective throughout is to convince us of the validity of a particular policy prescription, namely *laissez-faire*[106]. Now to say this is not necessarily to condemn Hayek, although we shall see later that he can indeed be seriously criticised on this score. It is certainly no dishonour, however, to adopt a polemical stance or to allow that stance to inform and inspire one's theoretical researches. Many major figures in economics, from Smith and Marx to Keynes and Friedman have a had a particular policy axe to grind, and have attempted to establish an economic theory to provide underpinning for a pre-existing social philosophy. It is their engagement with policy implications, rather than any aloof, purely theoretical standpoint, which has given what they had to say its coherence and bite. The problems that arise when Hayek attempts to do this, however, will become apparent as we examine his theory.

The first problem concerns the issue of consistency in Hayek's writings. Hayek simply cannot be relied upon to give us an accurate description of his own mode of procedure. We have already seen an example of this in connection with methodological issues: Hayek repeatedly asserts his adherence to one approach, methodological individualism, while in fact adopting a contrary one, methodological holism. When unconstrained by potential adverse policy implications – in theoretical psychology, for example – Hayek adopts a thorough-going holist account, and traces of this view can be seen elsewhere. When he is discussing society, however, the shutters come down and methodological individualism is proclaimed. That what is key for Hayek is the goal of privileging the level of the individual, and not the methodological approach to be adopted, is shown by

[106] Hayek himself denies that he supports *laissez-faire*, and denies that Hume, Smith or Burke supported it (**COL**: 60). This is partly because the expression is French and hence 'rationalistic' – 'as the very words show', but mainly because in his habitual absolutism he interprets *laissez-faire* as the negation of *all* state activity – an 'antistate ... or anarchistic' argument (ibid). Hayek is also habitually inconsistent; elsewhere he makes *laissez-faire* mean, not an injunction to leave individuals alone, but to leave the existing *laws* alone (**IEO**: 135). On this interpretation, a removal of absolutist and mercantilist laws restricting trade would be a *denial* rather than an *implementation* of *laissez-faire*. *Laissez-faire* is understood in the same way in **RTS** (**RTS**: 13, 27). This, I contend, is an absurd misuse of words. On the meaning of the word which everyone else uses, Hayek is clearly a proponent of *laissez-faire*.

his adoption of holism in psychology, and his claim to adopt individualism in economics. The former implies a systems view of the individual personality, while the latter again, although now arbitrarily, privileges the level of the individual agent. Focus on the abstraction of the individual is the goal, and the selection of methodology is made to suit.

We may see another example of Hayek saying one thing and doing another in connection with the theory of evolution. He correctly sets out the principle elements of the Darwinian theory of evolution, but the theory he actually uses is different. We can see this by comparing what he says in **COL**, **SIP**, and **NSP**. In **COL**, Hayek sets out his evolutionary theory and links it to his major themes. Hume, Smith and Ferguson, he says, were able 'to comprehend how institutions and morals, language and law, have evolved by a process of cumulative growth' (**COL**: 57). The key word here is 'evolved'. Hayek wishes to attract to the tradition from which he claims descent the prestige associated with the Darwinian theory of evolution (Hodgson, 1993: 152). He immediately links the question of evolution to that of policy: 'Those British philosophers have given us an interpretation of the growth of civilization that is still the indispensable foundation of the argument for liberty. They find the origin of institutions, not in contrivance or design, but in the survival of the successful.' (**COL**: 56-57).

So the 'survival of the successful' – Hayek's, but by no means Darwin's, theory of evolution – is an 'indispensable foundation' for the case for *laissez-faire*, the 'argument for liberty'. I have argued, above, that, in order to reconcile us to the way things are, to the governmental and property systems which actually exist, Smith argued that all was for the best in this world as the invisible hand of God guided agents to those actions which achieved the socially most desirable outcomes. I said also that in the more secular twentieth century this goal of reconciliation could only be achieved if a plausible and secular replacement could be found for the invisible hand. Hayek's aim is to replace the supernatural invisible hand with a natural one, a version of evolutionary theory which can combine the minimal policy prescription of *laissez-faire*, 'liberty', with the intellectual respectability of a Darwin.

At the same time he re-writes history in an attempt to show, contrary to what we saw in the previous chapter, that Smith did *not* make God the linchpin of his system: on the

contrary, we are informed, Smith *et al* had *already* discovered evolution, and, indeed, had done most of the work for which Darwin later took the credit:

> "Since the emphasis we shall have to place on the role that selection plays in this process of social evolution today is likely to create the impression that we are borrowing the idea from biology, it is worth stressing that it was, in fact, the other way round: there can be little doubt that it was from the theories of social evolution that Darwin and his contemporaries derived the suggestion for their theories." (**COL**: 59)

This is indeed a staggering claim. Although, like much of what Hayek has to say, it is wrapped up in vague phrases – here about 'deriving the suggestion' for a theory – this is clearly a claim that the credit for the Darwinian theory should properly go to the Scottish eighteenth century philosophers, principally Smith and Hume. As Hodgson says, 'Hayek's attempt to belittle the importance of the Darwinian revolution by claiming multiple precedence is ... without foundation .... It betrays both a misreading of the sources and some misunderstanding on Hayek's part' (Hodgson, 1993: 160).

Contrary to what Hayek says, there is no truth whatever in the claim that anything even remotely approaching an anticipation of Darwin can be found in the works of Adam Smith. Many of the things whose origin Smith attributes to God, such as 'sympathy' as a fundamental feature of our social psychology, make a great deal of sense in an evolutionary, rather than a theological, perspective. But that doesn't mean that he himself had such a perspective, indeed he most certainly did not[107]. As far as Hume is concerned, however, it is the case that there is some evidence to support Hayek's claim. We will be better placed to evaluate that evidence after establishing Hayek's evolutionary theory and its relation to Darwin's.

In **COL** Hayek implies that the essence of Darwinism is 'the idea that similarity of structure might be accounted for by a common origin', and then, to bolster his assertion of a pre-Darwinian evolutionary tradition, immediately adds that this was 'a

[107] It is often asserted that Smith had an evolutionary approach. Heilbroner detects a 'theme of historic evolution' in **WN** (Heilbroner, 1986: 149), and speaks of 'the historical evolution of astronomy' in **Astronomy** (Heilbroner, 1986: 16). But the kind of 'evolution' being referred to here is nothing more than a chronological sequence of forms, merely a 'stadial treatment of history' (Heilbroner, 1986: 150). There is no sense of diversification of social forms coupled with selection of the fittest ones which would make this an evolutionary process in the modern, Darwinian sense.

commonplace in the study of social phenomena long before it was applied to biology' (**COL**: 59). In **SIP**, however, he says

> "Before we examine its character [sc the character of the Darwinian theory of evolution by natural selection] we must clear out of the way a widely held misconception as to its content. It is often represented as if it consisted of an assertion about the succession of particular species of organisms which gradually changed into each other. This however, is not the theory of evolution but an application of the theory" (**SIP**: 31).

He adds in a footnote: 'that particular species had common ancestors, or that similarity of structure always means a common ancestry ... is emphatically not the main content of the present theory of evolution.' (**SIP**: 31, n21). And this view, indeed, reflects Darwin's own: the hypothesis

> 'that species had not been independently created, but had descended ... from other species ... even if well founded, would be unsatisfactory, until it could be shown how the innumerable species ... have been modified ... It is, therefore, of the highest importance to gain a clear insight into the means of modification and coadaptation' (Darwin, 1859/1928: 18-19).

It was thus, not merely the assertion of the hypothesis of descent, but also the explanation of modification, that was the aim of *The Origin of Species* (Dennett, 1995: 39). So the theory of evolution espoused in **COL** is now, in **SIP**, 'emphatically not' what evolution is about. In his polemic against this 'widely held misconception', he gives no hint as to who might be guilty of holding it, and none at all that he himself had recently held this view. If the **COL** conception of evolution is incorrect, what is the correct conception?

> "The basic conception of the theory [of evolution by natural selection] ... is that a mechanism of reduplication with transmittable variations and competitive selection of those which prove to have a better chance of survival will in the course of time produce a great variety of structures adapted to continuous adjustment to the environment and to each other." (**SIP**: 32)

This formulation does indeed express the essence of Darwin's theory. It combines the two principal points, 'descent with variation' and 'natural selection'. As far as the *result* of evolution is concerned, we have organisms which are well-adapted to their environment, which includes each other ('continuous adjustment' is the *form* of that adaptation in some species).

The reader might be forgiven for thinking that, at least it is the later formulation which is correct: Hayek has changed his view for the better, even if he isn't explicit about having done so[108]. But this is not so. In 1978 Hayek (**NSP**: 249-266) reprinted without comment his 1966 British Academy Lecture on a Mastermind, 'Dr Bernard Mandeville', in which he says 'What I do mean to claim for Mandeville is that [his] speculations … mark the definite breakthrough in modern thought of the twin ideas of evolution and of the spontaneous order.' (**Mandeville**: 250). Moreover

> "Mandeville … made Hume possible … This development [of Mandeville's ideas] includes … Adam Smith and Adam Ferguson … [T]he tradition which Mandeville started includes also Edmund Burke, and … all those 'historical schools' which … made the idea of evolution a commonplace in the social sciences of the nineteenth century long before Darwin. And it was in this atmosphere of evolutionary thought in the study of society, where 'Darwinians before Darwin' had long thought in terms of the prevailing of more effective habits and practices, that Charles Darwin at last applied the idea systematically to biological organisms … Darwin is the culmination of a development which Mandeville more than any other single man had started … Darwin … finished what Mandeville had begun." (**NSP**: 264-265).

So, again, Hayek is claiming that the Scottish Enlightenment philosophers, this time building on what he regards as Mandeville's contribution, were 'Darwinians before Darwin', and had discovered the theory of evolution in the particular context of society. Then, 'at last', Darwin came along and applied a commonplace idea of social science to the biological context, and the rest was history[109]. Hayek's account is a caricature. Hayek's paper, and, indeed – and perhaps more importantly – Mandeville's works, contain not one point of evidence to back up his description of Mandeville as a 'Darwinian before Darwin'. It is true that Darwin gained some inspiration from political economy, in particular, from Malthus: 'The Struggle for Existence … is the doctrine of Malthus' applied to the biological domain (Darwin, 1928: 19). But, firstly, there is no question of Malthus having articulated this idea with the other necessary components – replication, descent and modification – to achieve a theory of evolution. And secondly,

[108] As John Toye remarks of Keynes in another context, 'Changing one's mind is not itself ground for criticism … The ground for criticism must be … the failure to acknowledge publicly and to explain honestly the reasons for his change of mind.' (Toye, 1997: 19)

[109] These ideas have been taken up and repeated: 'The idea that there can be order and regularity without design, and that social change can be explained without recourse to deliberate human actions, was discovered by the thinkers of the 18th-century Scottish Enlightenment, notably Adam Ferguson, David Hume and Adam Smith, long before Darwin in a not dissimilar way explained the biological world without the intervention of a Creator' (Barry, 1988: 44).

Hayek gives no credit to Malthus for his contribution to Darwinism. As is well known, Malthus strongly rejected the Panglossian optimism of the Smithian tradition which Hayek cherishes: 'Hayek's view of evolution is orientated more to harmony and equilibrium – as in the cases of Adam Smith and Herbert Spencer – than to the relentless discord to be found in the works of Malthus and Darwin' (Hodgson, 1993: 277-278, n4)[110].

So Hayek says different, and, indeed, contradictory things. When in 1960 he reduces Darwinian evolutionary theory to a hypothesis of common ancestry – what he 'emphatically' rejects in 1967 – it is so that he can link Darwin with the Scottish philosophers, particularly Smith and Hume, and bolster a particular notion of evolution in which 'institutions … evolved by a process of cumulative growth' based on 'the survival of the successful' (**COL**: 57) and in which 'such conceptions as "natural selection," "struggle for existence," and "survival of the fittest," … are not appropriate' (**COL**: 59). To have introduced an objective discussion of Darwin's theory at this point would have showed the gulf which existed between it and Hayek's own 'process of cumulative growth'.

In the passages cited from **SIP**, he is less constrained. He is citing evolution as an instance of 'pattern prediction', as an application of the 'theory of complex phenomena' (**SIP**: 31-35). His point is to argue that, contrary to physical theories, evolutionary theory cannot be used to make specific predictions: because of the limits on our ability to acquire the necessary information, it is not possible to predict the direction evolution will take or to verify such predictions. All we can do is to talk about certain *patterns* of outcomes which can or cannot be produced and the conditions which would affect them. His interest in this is to say that the study of social structures is of the same kind: a case of pattern, not specific, prediction. Evolution is, again, being used here ultimately to underscore a theoretical point required to generate his desired policy conclusions: the limitations on our knowledge of society preclude central planning[111]. Nevertheless, the

[110] See also the discussion of Malthus's decision to drop the theodicy from his (1798) *Essay on the Principle of Population* in the second (1802) and subsequent editions in Poovey (1998: Ch 6).

[111] The argument is illegitimate since, as *The Sensory Order* makes clear, the same limits to self-knowledge apply to individuals as to society – but this doesn't lead Hayek to deny that individuals can plan their activities.

link is here far less direct; hence in this context he has no reason to distort the theory and we get a relatively objective account.

When we come to 'Dr Bernard Mandeville', the case changes again. Instead of discussing evolutionary theory in general terms as an example of the study of complex phenomena, he is deploying it to underpin a particular policy agenda. He draws out a tradition going from Mandeville through Hume, Smith, Ferguson, and Burke, as well as Savigny and Herder (**NSP**: 265), to himself. These thinkers share two things: a notion of a 'spontaneous order', a social order that arises through people's efforts but independently of their intentions, and a belief in the optimality – in some sense (the individuals Hayek mentions vary) – of that spontaneous order[112]. The identification of the two is hinted at in Hayek's definition of evolution here as 'the prevailing of more effective habits and practices' (**NSP**: 256). Hayek's concern is to present the spontaneous emergence of a social order as tantamount to an evolutionary process. The corollary would be that the spontaneous order philosophers are pioneer evolutionary theorists, and Darwin is seen as getting all the credit simply by applying the idea to biological phenomena. One result is to confer the scientific authority of Darwinism on the spontaneous order tradition, and another is to present a fallible social process as something natural and hence right. The idea of optimality is smuggled in under cover of spontaneity.

In sum, then, when Hayek is discussing the theory of evolution in an objective way, he is able to give a reasonable account of the process; when he is actually *using* the theory to buttress his overall system we get something very different. The theory is treated with Procrustean instrumentality. This points to a great cynicism on Hayek's part: he surely

[112] That what Mandeville says in many places is open to this interpretation is very true. But it is also true, what Hayek never so much as hints at, that in many places he says exactly the opposite: 'private vices by the dexterous management of a skillful politician might be turned into public benefits' (cited in Landreth and Colander, 1994: 45). 'Mandeville argued … that social good could result from selfish acts if these actions were properly channelled by the government. As a mercantilist, Mandeville had no concept of a natural harmony, which is an essential ingredient in Adam Smith's advocacy of *laissez-faire*' (Landreth and Colander, 1994: 45). No final resolution of the question as to whether Mandeville was a mercantilist or a liberal seems likely: the fact is that he inconsistently adopted both points of view in different places.

knows what he is talking about, but it seems that he is willing to distort theory, just as he is willing to distort facts, in order to fit with his desired conclusions.

We have seen how Hayek is guilty of a major distortion of Smith, how he recruits Mandeville to his spontaneous order tradition by suppressing contrary evidence, and how he is willing to degrade and distort Darwin's theory when his purposes require it. We now need to examine the case for his claim that Hume anticipated Darwin. In the 'Legal and Political Philosophy of David Hume', Hayek writes

> "I [speak] of Hume's doctrine as a theory of the growth of an order which provided the basis of his argument for freedom. But this theory did more. Though his primary aim was to account for the evolution of social institutions, he seems to have been clearly aware that the same argument could also be used to explain the evolution of biological organisms. In his posthumously published *Dialogues on Natural Religion* he more than hints at such an application. He points out there that 'matter may be susceptible to many and great revolutions, through the endless periods of eternal duration. The incessant changes to which every part of it is subject, seem to indicate some such general transformation.' The apparent design of the 'parts in the animals or vegetables and their curious adjustment to each other' does not seem to him to require a designer, because he 'would fain know how an animal could subsist unless its parts were so adjusted? Do we not find that it perishes wherever this adjustment ceases, and that its matter corrupting tries some new form?' And 'no form can subsist unless it possess those powers and organs necessary for its subsistence: some new order of œconomy must be tried, and so on, without intermission; till at last some order which can support and maintain itself, is fallen upon.' Man, he insists, cannot 'pretend to an exemption from the lot of all living animals ... [the] perpetual war ... kindled among all living creatures' affects also his evolution. It was still another hundred years before Darwin finally described this 'struggle for existence'. But the transmission of ideas form Hume to Darwin is continuous and can be traced in detail." (**SIP**: 119)

There are three references to *evolution* in the passage but all are in Hayek's words, not Hume's – and even if Hume had used the term *evolution*, it would not have meant what *we* mean by 'evolution', for the simple reason that the concept did not exist in the public domain prior to 1859. Let us explore this issue further. The object of this passage is to convince the reader that Hume had a full-blown theory of evolution which he applied to institutions and which he toyed with applying to biological phenomena. Yet the article on David Hume in which it appears, contains no evidence or argument that Hume did in fact have an evolutionary theory of society. What we see instead is a demonstration that Hume regarded the emergence of the regular and lawful institutional structure of society as being largely the product of a spontaneous process. The impression that Hayek wishes to create is that the two are the same thing. Now there are certainly links between the theory of evolution and the emergence of a spontaneous human order; evolution is clearly both orderly and rule-governed on the one hand, and spontaneous, in

that it cannot be the result of anyone's intention, on the other. But, equally certainly, there are very significant differences between the two; the social structure, for example, arises, and must arise, on the basis of the deliberate actions of human beings (even if not as a result of their intentions), but this is clearly not necessarily the case for evolution. A simple identity between the theory of spontaneous orders and that of evolution cannot be assumed: on the contrary, the links between the two, and what divides them, have to be argued for.

So much for Hume's theory of evolution in society. Let us now examine Hayek's presentation of the evidence in Hume's *Dialogues* for a theory of evolution in nature. To help us we may call as expert witness the American philosopher Daniel Dennett, who has dealt with precisely this point in his book *Darwin's Dangerous Idea*. Part 4 of Chapter 1 is entitled 'Hume's Close Encounter' (Dennett, 1995: 28-34). After noting, what is surely correct, that Hume himself appears in the *Dialogues* in the person of Philo, Dennett says that Philo 'dreams up some speculations that come tantalisingly close to scooping Darwin by nearly a century.' (Dennett, 1995: 32). And it is true: if we read the *Dialogues*, or just the passages cited in Dennett, we can see that Philo did come close to the discovery of evolution by natural selection.

The extracts Dennett cites show (a) a prescient depiction of the origin of life (an 'economy or order') by random transpositions of matter; (b) a statement that such ordered matter will sustain itself and produce the 'appearance of art and contrivance' which we actually observe in living things; (c) the claim that defects in form will lead to removal of the form. The latter statement is clearly an idea of natural selection. What this account lacks, what distinguishes it from Darwinism, is the idea of reproduction as combining overall stability of form with some errors: imperfect replication, allowing selection to have something to work on. Hume doesn't mention reproduction – which must be a severe defect in any putative theory of evolution.

Moreover, Philo is ambiguous about the key issue of imperfect replication. Firstly he contrasts the stability of the orders or economies, with the random transpositions of disordered matter. Secondly he implies that there must be changes in natural forms in that defects in them are removed. But if the orders are stable, where do the defects come from? Again, because he says nothing about reproduction, descent with

modification, we are left with that version of selection in which an original endowment of variety is gradually whittled away by selection leaving only the forms we see today. A form of selection in which a given pool of objects is sifted, the imperfect forms removed and only the adapted ones remain. In this account nothing has happened to the adapted forms, only to the unadapted ones: the adapted forms were there from the start[113], selection only discovers them. As Dennett says, 'cute ideas about evolution had been floating around for millennia' (Dennett, 1995: 33), and, indeed, this particular version can be seen in Lucretius in the first century BC (Lucretius, 55 BC: Book 5, lines 837-877; 1969: 191; 1951: 196-198), and in Empedocles more than four hundred years before that (Barrow and Tipler, 1988: 34).

In arriving at a verdict on Philo, we need to take three things into consideration. Firstly, Philo's remarks constitute only a speculation, involving a brief passage, not a worked out theory. Secondly, Hume himself 'couldn't quite take Philo's daring foray seriously' (Dennett, 1995: 33). Hume's (and Philo's) own verdict was that 'a total suspense of judgment is here our only reasonable resource' (cited, ibid). And, finally and most seriously, without a source of variety for natural selection to work on, in the form of copying with errors, a fully functioning theory of evolution is not possible. Hume has thus not 'scooped Darwin'. It is not therefore the case that any of Mandeville, Hume or Smith anticipated Darwin.

To summarise the results of the discussion so far: Hayek is able to set out the theory of evolution concisely, but when he comes to use it, the theory of evolution he employs is something very different. Hayek claims that the 'Mandeville-Hume-Smith-Ferguson tradition' (**NSP**: 265, n58) developed the theory of evolution which Darwin then later applied to the biological sphere. This claim is false. In both cases his objective is to identify the emergence of a spontaneous social order with an evolutionary process, in order to present that order as something natural, something which we cannot and should not adapt to our own needs. The arguments in each case have been shown to be

[113] Philo's proposal seems to contradict this, with new, more perfect forms arising in place of old, more imperfect ones. But in Philo's account the origin of the new forms is obscure – if anything, it appears to be by the continual emergence of new life forms from inanimate matter rather than by adaptation of existing forms. In this sense it is true that adapted forms do not acquire their adaptation from the evolutionary process but acquire it at their origin as life forms.

illegitimate. Ultimately the rôle of evolution in Hayek is just that of the invisible hand in Smith: an inscrutable agency which arranges that the unintended outcomes of our actions will be benign. Again, we have an interface between a holistic world and a reductionist policy prescription: our activities as human beings are organically linked to each other but can be treated as if they were independent: if the actions of each individual are directed towards the good of that individual, then the actions of all individuals must tend to the good of all. All the hard work of reconciling disparate interests and behaviours is accomplished behind the scenes by the process of evolution.

5.4.2 *Hayekian evolution*

We have seen that the theory of evolution which Hayek deploys is not Darwin's. Now we need to investigate the merits of Hayek's theory in its own right. What does Hayek mean when he says that institutions 'evolved by a process of cumulative growth', when he locates 'the origin of institutions … in the survival of the successful' (**COL**: 56-57), and when he defines evolution as 'the prevailing of more effective habits and practices' (**NSP**: 256) or as 'the idea that similarity of structure might be accounted for by a common origin' (**COL**: 59)? This subsection will investigate these questions further.

In order to do so we first need to introduce the distinction between ontogeny and phylogeny – terms used by Hayek himself (**TSO**: 42, for example), but without any apparent consciousness of the significance of the distinction for his theory. Ontogeny is the development of the individual of a species, while phylogeny is the evolution of the species. Now the term *evolution* can be used in both senses: both '[t]he process of developing [of animal and vegetable organisms] from a rudimentary to a complete state' and 'the origination of species conceived as a process of development from earlier forms' are included under the catchword *evolution* in the *Shorter Oxford English Dictionary* (Onions, 1973: 693)[114]. But there is a key difference between these ideas: ontogeny, the development of the individual from embryo to maturity, has a goal, whereas phylogeny has none. Ontogeny is teleological: the phases which the immature organism normally passes through are *means* to the *end* of creating a reproductively mature adult. Once

[114] The two definitions are dated to 1670 and 1832 respectively. The second definition still is not the same thing as Darwinism – as, indeed, it could not be, prior to 1859 – lacking as it does the key notion of natural selection as the cause of the modification of species.

that goal, adulthood, is achieved, development ceases. Deviation from the normal ontogeny is in general pathological. Phylogeny is very different: there is no goal in the development, and can be none, since no agency exists to implement the interest of the species. Each individual attempts to survive and to pass on its DNA to successive generations, the best it can, given the character of its environment, including the other individuals of its species. The history of the species is just a list of the ways in which this happens to have been done in the species' past. Since the underlying nature of the world is permanent change, the process of adaptation to the environment can never be complete. Ontogeny terminates, but the only possible terminus to phylogeny is extinction.

It must, then, be clear that the Darwinian theory of evolution is a theory of phylogeny, not ontogeny. Hodgson (1993: 161) makes the case, however, that Hayek's evolutionary theory is ontogenetic and not phylogenetic:

> "When Hayek (1967, p.72 [**SIP**: 72]) writes that 'the whole of economic theory ... may be interpreted as nothing else but an endeavour to reconstruct from regularities of the individual actions the character of the resulting order' he is letting the cat out of the bag. Biological ontogeny is precisely the endeavour to explain the development of organisms from the regularities of their genetic endowment, in contrast to phylogeny which considers the sifting and changing of the gene pool through natural selection or drift. Hayek's statement clearly suggests ontogeny rather than phylogeny." (Hodgson, 1993: 161)

Now this argument could easily be extremely unfair to Hayek. Ontogeny and phylogeny are not alternative theories in biology, but complementary. It is true that species evolve, and it is also true that individuals develop to adulthood. Hodgson has clearly identified an ontogenetic strand in Hayek's exposition, but that does not prove that Hayek uses this ontogeny *in place* of phylogeny. Hayek's statement, cited by Hodgson, that the *whole* of economics is ontogeny is evidence for Hodgson's view, but it is well known that Hayek, like Popper, exhibited great anxiety and dogmatism about the boundaries between one science and another and between science and other related activities. It would be quite plausible for Hayek to relegate phylogeny to some other cognate discipline. We will see shortly that this is not the case.

The implication of Hodgson's interpretation is that Hayek's account is not genuinely evolutionary, since evolution concerns phylogeny not ontogeny. This is very important for Hayek, because the 'evolution' of an individual organism leads to a predetermined

outcome: there is a logical goal, and once you've got there, there is nowhere else to go – the 'evolutionary' process terminates, there is no more evolution. We reach the end of history. The ontogenetic version of evolution is fully compatible with the Whig interpretation of history, in which the essence and goal of history is the development of liberal institutions (Marwick, 1970: 45). The apologetic character of Hayek's enterprise, in Hodgson's interpretation, is very clear.

Hodgson also links Hayek's ontogenetic version of evolution with his persistent tendency to slight Darwin's achievement, discussed in the previous subsection:

> "Thus, in implicitly comparing his theory to the kind of economic ontogeny found in the writings of Walras or Smith, Hayek makes the addition of the idea of 'natural selection' a mere appendage. Darwin is then reduced in stature because he is not significant for the Hayekian theory. Without further clarification, the latter can easily be reduced to the post-Humean ontogeny of the emergence of the coherent social order... [O]ntogeny was well established before Darwin. It is thus no accident that Hayek simultaneously upgrades ontogenesis and downgrades Darwin's contribution." (Hodgson, 1993: 161)

Here we have further evidence that Hayek attempts to annex the prestige of Darwinian evolution whilst simultaneously distorting the Darwinian theory. Hayek's theory is ontogenetic; ontogenetic theories of development in nature and society were extant long before Darwin, and, moreover, are not theories of evolution at all, but a theory of the maturation of an organism. Clearly, twentieth century capitalism has emerged from *some* kind of developmental process. The question is, whether that process was ontogenetic or phylogenetic. Hayek's case is that the relevant concept of development is that of ontogeny: we have reached the terminus of history, and the system we have now must be considered optimal. If, on the contrary, phylogeny is the appropriate concept then no such assumption of optimality can be justified: on the contrary the desirability or otherwise of the institutional framework has to be determined by reference to the facts rather than to theological, or evolutionary, invisible hand apologetics.

Even Norman Barry, a writer with great sympathy for Hayek's project, regards Hayek's invocation of evolution as a failure:

> "Hayek has suggested ... that the process of evolution (not quite the same as 'Social Darwinism') will 'select out' those aggregate structures that are more successful in meeting man's needs. Most writers have found this implausible and indeed the historical evidence hardly points to an ultimate triumph of the market order ... the evidence for

> Hayek's solution is rather meagre and hardly sufficient to establish it as even the beginning of a genuine theory." (Barry, 1988: 46-47)

5.4.3 *The assumed optimality of evolved institutions*

There is a key passage in **COL** in which Hayek links the Scottish Enlightenment thinkers, social evolution, and the question of optimality:

> "While the rationalist tradition assumes that man was originally endowed with both the intellectual and the moral attributes that enabled him to fashion civilization deliberately, the evolutionists made it clear that civilization was the accumulated hard-earned result of trial and error; that it was the sum of experience, in part handed from generation to generation as explicit knowledge, but to a larger extent embodied in tools and institutions which had proved themselves superior − institutions whose significance we might discover by analysis but which will also serve men's ends without men's understanding them. The Scottish theorists were very much aware how delicate this artificial structure of civilisation was which rested on man's more primitive and ferocious instincts[115] being tamed and checked by institutions that he neither had designed nor could control. They were very far from holding such naïve views ... as the "natural goodness of man," the existence of a "natural harmony of interests," or the beneficent effects of "natural liberty" (even though they did sometimes use the last phrase). They knew that it required the artifices of institutions and traditions to reconcile the conflicts of interest. Their problem was how "that universal mover in human nature, self love, may receive such direction in this case (as in all others) as to promote the public interest by those efforts it shall make towards pursuing its own."[116] It was not "natural liberty" in any literal sense, but the institutions evolved to secure "life, liberty, and property," which made ... individual efforts beneficial .... [I]t was not some sort of magic but the evolution of "well-constructed institutions," where the "rules and principles of contending interests and compromised advantages"[117] would be reconciled, that had successfully channeled individual efforts to socially beneficial aims." (**COL**: 59-60)

There are many points to comment on in this passage. Firstly, as we saw in the previous chapter, Adam Smith, for one, certainly did hold the 'naïve view' of a 'natural harmony of interests' and the 'beneficent effects of natural liberty'. On the 'goodness of man' his view was a little more complex: it was true, in his view, that man was imperfect – but even the imperfections were god-given and contributed to the overall plan. The implication of Smith's account is therefore that the imperfections were only construed as such by men because of their finite minds: the infinite mind of God, seeing all the

---

[115] These 'primitive and ferocious instincts', Hayek later explains, are *altruism* and *solidarity*: 'It is these two instincts ... which remained the great obstacle to the development of the modern economy .... I could write the whole of economic history in terms of the subduing of these good natural instincts by culturally developed rules of conduct' (**KES**: 31).

[116] The reference is to Josiah Tucker (1755) *The Elements of Commerce* in RL Schuyler (ed) (1931) *Josiah Tucker: A Selection* New York: Columbia University Press, p 92.

[117] The reference is to Edmund Burke *Thoughts and Details on Scarcity* in *Works* Vol VII: 398.

'connexions and dependencies' of things, would realise the contribution of man's apparent imperfections to the perfection of the whole. In this sense, then, Smith certainly did believe in the natural goodness of man, despite all appearances contradicting that belief. Again, Hayek distorts what the Scottish Enlightenment thinkers actually said in order to present them as the initiators of an evolutionary trend in social thought.

But the main point of the passage lies in Hayek's acknowledgement of the need to reconcile conflicting interests. The central question of economics and of social science in general does indeed concern how individual self-seeking behaviour is to be led into socially desirable channels. Hayek's account here is extraordinarily abstract. We are assured that institutions evolved to secure life, liberty, and property, and that they have the effect of directing self-interest into the service of the community. But we are given absolutely no details – no examples, no analysis of the evolution of particular institutions or types of institution. Hayek ought to show, as an example, how, and in what circumstances, some of the institutions we have inherited from the past reconcile self-interest with the interests of others. But there is no word here on how institutions accomplish this task.

There are two sides to Hayek's account. Firstly, he contrasts the theory of the spontaneous emergence of a social order with the extreme rationalist account[118] in which every instance of institutional progress occurs as a direct result of the conscious intentions of some social reformer. In this Hayek is correct. It was in the interest of the absolutist regimes and their mercantilist literary representatives to argue that the only alternative to a consciously constructed and imposed centralised order was a chaotic anarchy. The Smithians performed a signal service by arguing that social orders could also arise spontaneously on the basis of self-seeking behaviour.

Secondly, however, we have the hypothesis of optimality. The institutions which issue from an extensive process of evolution are said to reconcile individual interests and channel individual efforts into socially beneficial directions. Clearly, they must do this to *some* extent, or we would not be able to observe spontaneous orders, but the key question is, how *well* they do this. On this, Hayek is content to remain silent and to

[118] This is not the place to discuss whether his treatment of the continental rationalist trend is any nearer the mark than his treatment of the Smithian trend.

allow the presumption of optimality to remain unexamined. In his theory, the institutions which are handed down by our forbears show their superiority in their very survival: they survived because they 'proved themselves superior'. In two other places in the passage he argues that individual interests are rendered socially beneficial by the evolution of institutions. But this begs the question: institutions, like genes, survive the selection process and lever themselves into the future by 'proving themselves superior' – at *surviving*, not necessarily at serving our interests. An institution may well survive because, although harmful to the majority, it serves the interest of a minority who happen to be powerful enough to override the majority's wishes – the Mafia, for example, or Thuggee.

The *clear* implication of Hayek's account is that anything which is old is superior. If it is old it has survived the selection process of imitation and learning, by means of which institutions, ideas and skills are transmitted to new generations, on the basis of their success in serving the interests of previous generations. The *hidden* implication is that when one of two institutional forms is selected, then it can be regarded as unambiguously superior to the alternative form. In some sense it reconciles the interests of the individual agents more efficiently. There is no hint that one group of agents may have an interest in one institutional setup, while another group's interests are better served by an alternative. Once this is allowed, institutional change involves victory and defeat for different groups. In that case it becomes impossible to describe such change, without qualification, as improvement. Hayek's account, therefore, contains the hidden implication, that there are no fundamentally conflicting interests: everyone has, let us say, 'congruent' interests: and it is unproblematic to compare two institutions and select the 'superior' one. Amartya Sen, discussing this point in Hayek, makes the point that the Smithian argument which Hayek relies on

> "rests precisely on the ability of the market to *achieve* the results intended by individuals ... and *then* some more. I want bread and will happily give some money for it, and the baker wants money and will give me a loaf of bread in exchange. When we carry out the exchange, we do achieve what we set out to achieve, and in the process we have helped each other ... the market works on the basis of congruence of interests of different participants. That is the essence of the Smithian perspective: different people have a common interest in exchange and the market gives them the opportunity to pursue their common interests ... In most economic problems the interests of the different people involved are partly congruent, partly conflicting. The market mechanism on its own confines its attention only to issues of congruence, leaving the interest conflicts unaddressed." (Sen, 1983: 4-6)

As we saw in the previous chapter, market activity in Smith does not reconcile individual agent interests – because these interests have been pre-reconciled by the invisible hand of God: exchange *realises* that pre-existing reconciliation. Now, we can see that Hayek adopts a comparable stand. Agent interests are assumed not to be mutually contradictory: individual agents hold common but merely spontaneously uncoordinated interests. Explaining the evolution of new modes of conduct Hayek says that the 'new manners of conduct ... were adopted because somebody who acted on them profited from it and his group gained from it' (**KES**: 32). In other words, there was a pre-existing harmony of interest between the individual and the group: the individual benefits himself – and simultaneously the group – by his actions. This is the basic assumption on which Hayek's whole philosophy is raised. Agent interests, in Sen's terminology, are 'congruent'. Inherited institutions then *realise* the agents' congruent interests by bringing about a coordination of their activities. This allows the optimality of the institutional framework resulting from the evolutionary process to be assumed.

Since I have placed considerable emphasis on the optimality assumption in Hayek, it is important to consider the occasions on which it is explicitly raised in his writings. I will refer to two such passages, one in **IEO** and one in **COL**.

> "If each man is to use *his* peculiar knowledge and skill with the aim of furthering the aims for which *he* cares, and if, in so doing, he is to make as large a contribution as possible to needs which are beyond his ken, it is clearly necessary ... that the relative importance to him of the different results he can achieve must correspond to the relative importance to others of the more remote and to him unknown effects of his action." (**IEO**: 17) "[A]ny workable individualist order must be so framed ... that the relative remunerations the individual can expect from the different uses of his abilities and resources correspond to the relative utility of the result of his efforts to others ... An effectively competitive market satisfies [this] condition" (**IEO**: 21).

This is just an obscure way of saying that the ordinal ranking of the possible outcomes of his action ('the relative importance of the different results he can achieve'[119]) must be the same for him and for other people affected by his actions; ie, if he faces a choice between spending and saving, and he decides to save, society in general must not prefer that he spends. In other words, there must be a harmony of interests. If this condition were always satisfied, then self-serving behaviour would *always* lead to socially optimal

[119] Hayek, as so often, is obscure: by 'relative importance' he may well mean something more than an ordinal ranking. Whether it does or not, it certainly means *at least* that, and hence the argument in the text follows.

outcomes. So we urgently need to know when we can expect this harmony to hold. Remarkably, Hayek simply asserts that it holds in an 'effectively' competitive market: in such a market the individual making a decision has the same ranking of outcomes of his behaviour as everyone else does. Hayek, in other words, blandly asserts that there is no such thing as an externality. This is just to assume the whole problem away. It also eliminates any qualitative diversity between individuals: each individual has the same ranking of outcomes as everyone else.

It is true that Hayek allows himself an escape route: to speak of 'effectively competitive' markets could imply that any market in which the condition fails is defined as 'ineffectively competitive', thus allowing the statement to stand as a tautology. If the claim is empirical, he should justify it empirically – which, of course, he cannot do. If it is a definition, he should tell us under what conditions competition will be 'effective'.

In **COL**, Hayek allows that the points he has previously made

> "do not prove that all the sets of moral beliefs which have grown up in a society will be beneficial .... [A] group or nation [may] destroy itself by the moral beliefs to which it adheres. Only the eventual results can show whether the ideals which guide a group are beneficial or destructive .... It may well be that a nation may destroy itself by following the teaching of what it regards as its best men .... There would be little danger of this in a society whose members were still free to choose their way of practical life, because in such a society such tendencies would be self-corrective: only groups guided by "impractical" ideals would decline, and others, less moral by current standards, would take their place. But this will happen only in a free society in which such ideals are not enforced on all." (**COL**: 67)

So, although he admits that suboptimal systems may evolve, firstly, this can only be judged retrospectively, by 'eventual results': it is thus impermissible for governments to rationalistically step in beforehand to avert the catastrophe. Secondly he is able to assert that there would be 'little danger' of suboptimal results in a 'free society' – by appeal to an argument … which itself *assumes optimality*: 'groups guided by "impractical" ideas would decline, and others … would take their place'. The assumption is that what is good for individuals is good for their group and what is good for the group is good for the nation. But of course the behaviour which is Nash for agents within a society – whether they be individuals or groups – cannot be assumed to be optimal for the society as a whole. Individuals and groups do not achieve pre-eminence in a nation by following rules which it would be in the interest of the nation for everyone to follow, but by

following rules which are well adapted for gaining power and influence within a nation's establishment.

To conclude, therefore, Hayek does clearly subscribe to the optimality assumption. A market system in a 'free society' based on private property (**IEO**: 20) will generate social outcomes which are the best we can get. His theory of social evolution is intended to provide underpinning for this assumption.

In view of the link drawn in the previous chapter between the assumption of optimality and the natural law tradition in Smith, it is interesting to note here the very favourable view that Hayek takes of that tradition. In **Mandeville**, pp 131 ff, he details how the idea of spontaneous order was maintained by theorists of 'the law of nature' – ie, natural law – from Greek times up to the present. He postulates a connection between freedom, natural law, and a belief in the agency of a benign deity: 'There appears to have existed in all free countries a belief that a special providence watched over their affairs which turned their unsystematic efforts to their benefit' (**Mandeville**: 130)[120]. 'Belief in special providence', that is, the belief that a kindly God would tie up all the loose ends and heal all ills, is the central tenet of Smith's Panglossian *Weltanschauung*. Hayek summarises the history of the spontaneous order tradition as follows:

> "This tradition was handed on, chiefly through the theories of the law of nature; and it is startling how far the older theorists of the law of nature … penetrated into the secrets of the spontaneous development of social orders .... [I]t [sc this tradition] led to a systematic questioning of how things would have ordered themselves if they had not otherwise been arranged by the deliberate efforts of government; they [sc the older theorists of natural law] thus produced what I should call the first modern theories of society" (**Mandeville**: 131).

Thus Hayek claims intellectual descent from the ancient[121] and medieval tradition of natural law. In particular he says of the tradition from which he claims intellectual descent that 'Savigny and his older historical school, largely based on the conception of a

[120] The *preceding* sentence reads 'It is remarkable how close … some of the ancient thinkers came to an understanding of the evolutionary processes that produced social institutions.' So according to Hayek, holding a belief in the ministrations of a special providence illustrates an understanding of social evolutionary processes. Hayek draws an explicit link between 'evolutionary processes' and 'special providence' – yet more evidence of the quasi-theological rôle that 'evolution' plays in his overall thesis.

[121] Two examples he gives are Aristophanes and Cato (**Mandeville**: 130-131).

grown order elaborated by the Scottish philosophers of the eighteenth century … continued or resumed the aim of the older natural law theorists' (**SIP**: 103-104).

5.4.4 *Group selection*

The idea that institutions, constructed on the basis of individual micro behaviour, are selected on the basis of their success at the macro level, leads immediately to the idea of group selection, so reviled by contemporary biology. This subsection explores the theme of group selection in Hayek's theory of social evolution.

Hayek sets the scene in 'Notes on the Evolution of Systems of Rules of Conduct' (**SIP**: Ch 4: 66-81). Firstly, he explains that *rules* of individual behaviour constitute the basic units of evolution in society: it is rules which play in social science the role played by genes in biology. 'The term 'rule' is used for a statement by which a regularity of the conduct of individuals can be described, irrespective of whether such a rule is 'known' to the individuals in any other sense than that they normally act in accordance with it.' (**SIP**: 67) Hayek's *rules* can be transmitted genetically or culturally, and they may be embodied in humans, non-human animals, or even self-replicating von Neumann machines (**SIP**: 66). In so far as they are genetically transmitted, they are identical to the genes for the behaviour they specify; in so far as they are culturally transmitted, they are synonymous with Dawkins's *memes*, 'the new replicator ... a unit of cultural transmission, or a unit of imitation' (Dawkins, 1989a: 192). Hayek distinguishes 'between the systems of rules of conduct which govern the behaviour of the individual members of a group ... and ... the order or pattern of actions which results from this for the group as a whole.' This distinction is well-taken. The system of rules of conduct are the social equivalent of the genotype, let's call it the *memotype*, while the order of actions corresponds to the *phenotype*. The system or order of rules is the set of instructions; the system or order of actions is the outcome. They describe the micro behaviour of the individuals composing the society and the macro social outcome, respectively.

Contrary to any reductionist view, which would imply a simple, mechanical, and aggregative relationship between memotype and phenotype, between system of rules and order of actions, Hayek stresses the contingent nature of the macro level system arising on the basis of any particular micro level behaviour. 'The interplay of the rules of

conduct of the individuals with the actions of other individuals and the external circumstances in producing an overall order may be a highly complex affair' (**SIP**: 71). He cites the entropy principle embodied in the second law of thermodynamics as an instance of regular micro level behaviour leading to perfect disorder at the macro level (**SIP**: 67), and a society in which fixed rôles were filled by individuals selected by lot, as an instance of irregular behaviour at the micro level supporting a perfectly orderly macro outcome (**SIP**: 69). Moreover,

> "Not every system of rules of individual conduct will produce an overall order of the actions of a group of individuals ... and it is at least conceivable that the same overall order of actions may be produced by different sets of rules of individual conduct ... The same set of rules of individual conduct may in some circumstances bring about a certain order of actions, but not do so in different external circumstances." (**SIP**: 67-68)

These perfectly correct – indeed, valuable – points, however, are leading up to an *incorrect* theory. An order is 'a steady structure (showing 'homeostatic' control)' – in other words, an organism. And what could be more *holist* than the assertion that 'systems of rules of conduct will develop as wholes ... the selection process of evolution will operate on the order as a whole' (**SIP**: 71)? Having set out a holistic view of the relationship between the micro and macro levels, in which the link between the two is complex, indirect and mediated, rather than simple, direct and immediate, as it would appear in a reductionist view, he proceeds to break the link between the two altogether. Macro level objects, 'orders', now become independent entities in their own right, owing nothing to their material bases in individual behaviour. It is in this sense that Hayek departs from what Bunge would call 'systemism', and slips into a holism which ignores the substrate-dependence of macro-level phenomena: he forgets that there are 'no relations without relata'. Having severed the link between levels, Hayek strays into mysticism: evolution in this view can operate on the order as a whole even in the absence of any mechanism tying the interest of the individual to that of the whole. By eliminating the tie between macro and micro, Hayek obscures the necessity of such a mechanism. His thesis is that social evolution is evolution which occurs at the level of society, at the level of the group: the cultural

> "*transmission* of rules of conduct takes place *from individual to individual*, while what may be called the natural *selection* of rules will operate on the basis of the greater or lesser efficiency of the resulting *order of the group*." (**SIP**: 67) "The evolutionary selection of different rules of individual conduct operates through the viability of the order it will produce" (**SIP**: 68). 'T[h]e origin of institutions [is to be found] ... in the survival of the

> successful.' (**COL**: 56-57). Evolution is 'the prevailing of more effective habits and practices' (**NSP**: 256).

Cultural rules are, indeed, transmitted from individual to individual in the sense that only individuals can execute cultural instructions (whether on their own account as principals, or as agents for others such as firms and organisations), although the transmission is mainly *mediated* by cultural artefacts – telephone messages, letters, newspapers, magazines, journals, books, films, TV and radio programmes, e-mail and websites[122]. However, *efficiency* and *viability*, *successful* and *more effective* are left undefined. Do we mean efficiency for the order or for the individuals working within its framework? It is clear that Hayek means 'efficiency for the group of individuals'. Logically, however, it can only mean 'efficiency for, and viability of, the order'.

We have a system of instructions. Executing those instructions has consequences. Some consequences are more favourable than others to the continuation and expansion of the system of instructions. Those systems of instructions which we actually find are likely to be those whose execution leads to their own successful replication. What does lead to successful replication depends on the environment. The environment of the system of rules includes its own substrate. Just as a person can only do what their limbs are able and willing to do – most people cannot hold their hand in a flame or look at the sun, and no one can fly by flapping their arms – a society can only do what its constituent individuals are willing and able to do. The difference is that in the person, the individual components all have the same set of instructions, the same genotype; and their interest is to lever a copy of that genotype into the future. The only way they can do it is by aiding the production and care of offspring. This can only be done by each part playing its rôle in the activity of the whole person. Every part has an interest in cooperating with every other part to fulfil the aims of the whole person. In society it is otherwise. Everyone has their own interest to follow, largely based on tastes and preferences selected for because of their likelihood of leading to successful genetic propagation. So the parts of an organism play the rôle they are required to play if they get the right information and resources – they already have the necessary incentive for playing their part in the overall scheme. The parts of a society play the required rôle – the rôle required for the successful adaptation of the social system – if they get the required information,

[122] ie, Popper's 'World 3': see Magee (1973: 59-61).

resources, *and incentives*. So the system will have a selective advantage if it contains instructions which allocate to individuals (including here every kind of economic agent – individual households, firms etc) the appropriate incentives, information and resources.

Consider the following thought experiment to show that we cannot assume that the 'efficiency' which is selected for in the evolution of institutions is necessarily the efficient satisfaction of human interests: A well-adapted social system, one which has survived the selective process, must include a system of incentives for individuals to follow. In following these incentives, the individual is necessarily doing something which is both in his own interest, or he wouldn't do it, and in that of the system, or it wouldn't have given him the incentive to do it (we assumed at the start that it was well adapted). So, on the assumption that a set of such incentives can in fact exist, there is at least that much mutuality between the individual and the collective. Nevertheless, while it may be in the individual's interest to behave in this way – call it action A – *given what everyone else is doing* (ie, action A is Nash), it may well be that the collection of all individuals could all do better if they were somehow all coordinated to behave otherwise – say to carry out action B. In other words, the individual agents find themselves in a multi-player prisoners' dilemma. In this case, we have two systems of rules competing for the allegiance of the population. One system directs them to carry out action A, acting as an isolated human atom, the other requires action B, where the population acts as a collective agent and each individual achieves a better satisfaction of his interests. While none of this may in fact be the case, it is clear that there is no reason in principle for us to assume that the characteristics which are spontaneously selected for in social systems will be more desirable to the members of those social systems than the characteristics with which they might wish consciously to endow them.

Yet this is Hayek's key assumption throughout – that the selection process sifting institutions is one which endows those systems of rules which are better able to satisfy human interests with a higher probability of survival and propagation. The thought experiment shows that this assumption is unwarranted. On the contrary, selection of institutional forms of society may well throw up systems of rules of conduct in which the behaviour of each is Nash, but the outcome for all is suboptimal. And, of course, it is precisely the contention of many that that is precisely the situation we face. Hence Hayek's argument begs the most important question. To take a single example, it is

perfectly clear that the order we have involuntarily constructed is not beneficent towards animals, or we might all be vegans; on what grounds are we to believe that it is beneficent towards humans?

Necessarily, Hayek is extremely vague[123]. He cannot specify the mechanism by which mutation and selection is to take place. It may be that there is a set of rules such that if a given society were to implement it, it would have a competitive advantage over other groups. But it may also be the case that the individuals of the society cannot reach that set of rules from their present set by each individual following his own spontaneous interest. It may need coordination at the macro level to achieve it. Selection will never be able to work on this set of rules as there is no spontaneous mechanism which will allow the society to adopt it, without it being imposed on the whole society by central command.

To illustrate his thesis that macro level orders are systems of behaviour that have been selected for because they are optimal for the micro level agents, Hayek turns to zoological examples:

> "The most easily observed instances in which the rules of individual conduct produce an overall order are those where this order consists in a spatial pattern such as will occur in the marching, defence, or hunting of a group of animals or men. The arrow formation of migrating wild geese, the defensive ring of the buffaloes, or the manner in which lionesses drive the prey towards the male for the kill, are simple instances in which presumably it is not an awareness of the overall pattern by the individual but some rules of how to respond

[123] 'Hayek repeatedly associates evolution with the existence of some kind of selection mechanism, although its specification, along with that of the unit(s) of selection and the criteria of fitness, are somewhat vague' (Hodgson, 1993: 163). Indeed, much of what Hodgson has to say about Hayek is on a note of frustration: Hayek is so obscure, allusive, vague, contradictory and downright wrong that often little can be made of what he says. Hodgson refers to Hayek's 'errors and travesties', his 'recurring historical error' (Hodgson, 1993: 159), to 'not unique cases of a casual attitude to sources and scholarship in Hayek's work' (160); 'Hayek is unclear as to the criteria on which evolutionary selection takes place' (164), 'Hayek's own standpoint [on the unit of selection is] ... vague and rarely elaborated' (164) ; 'The mechanisms of rule replication are not clarified or explained. The mere suggestion of imitation is not enough', 'it is not entirely clear', 'Hayek writes vaguely', 'Here there is another gap in his theory and it is necessary to interpolate and to conjecture so as to attempt to understand his theoretical system as a whole' (165); 'Despite having written on cultural evolution for over 20 years ... Hayek's evolutionary theory remains incomplete' (166). I can only sympathise.

to the immediate environment which co-ordinate the actions of the several individuals." (**SIP**: 69)

This is a heterogeneous list of examples. It would be out of place here to launch into a critique of Hayek's views on the behaviour of lionesses. Suffice it to say that even if his description of their hunting methods were true, it is most unlikely that their tactics could be reduced to a stereotype summed up in a few simple rules, in the same way as geese and buffalo. But that is not the main point here. What is interesting is how these examples contradict his thesis. Firstly, the arrow formation of the geese. In Hayek's model, the arrow formation would be an order of actions which would have evolved because it was optimal for the geese. A better explanation, or candidate explanation, is that perhaps each goose gains by flying in the slipstream of another, and so it follows the rule of doing so, where possible. If this is true, then the arrow formation is an epiphenomenon of following this rule: it confers no cost or benefit on the flight of geese. Now we know that this is not true for Homo sapiens: the patterns of our collective behaviour have a major impact on the fate of individuals.

The defensive circle of the buffalo seems to fit Hayek's pattern a little better (again, assuming that his facts are correct): it is quite likely that each animal follows a simple rule in certain circumstances of danger – the result is a defensive circle which no individual animal intended but which serves all their interests optimally – and presumably this behaviour evolved because groups of buffalo which reacted thus were at a selective advantage. The point is that although Hayek's thesis is plausible in individual cases, it can not be assumed to hold in general. On the contrary, in situations where it may well hold, such as the buffalo circle, there has to be a mechanism by which it is in the interest of individuals to behave in the way required for the macro level outcome. In the case of the buffaloes this incentive may lie in the consanguinity of the group: each buffalo is a vehicle for much the same assemblage of genes, and so (the gene complex embodied in) each gains if it aids the survival of its fellows. The externality is internalised. Or perhaps it just makes an immediate gain in security for little extra cost by joining a defensive circle. Such links from self-seeking micro behaviour to desirable social outcomes cannot be assumed in groups of humans, but have to be demonstrated on each occasion where it is thought to hold.

Hayek thus believes that people do what they do, not because it is in their interest to do so, but because it is functional for the society for them to do so[124]. A conception which clearly denies the necessity of incentives to underpin any posited pattern of individual behaviours. Speaking of the rules of conduct in primitive human societies, he says that

> "the 'functions' which these rules serve we shall be able to discover only after we have reconstructed the overall order which is produced by actions in accordance with them ... all the individuals of the species which exist will behave in that manner because groups of individuals which have thus behaved have displaced those which did not do so." (**SIP**: 70)

When we act, what we do is describable, if sufficiently regular, by a rule. But the question is whether the rule is an epiphenomenon, like the arrow formation of the flying geese, a pattern which emerges from generalising a large number of instances of the particular action, or whether the individual actions are executed *because of* the rule. In the second case, the actions of individuals are functional for the purposes served by the rule. The use of the term 'functions' in the passage cited – albeit in scare-quotes – only illustrates Hayek's functionalism.

> "Vanberg (1986[125], p.83) is right to suggest that Hayek's argument has a functionalist quality; it assumes that the contribution of a rule to the maintenance of a system is sufficient to explain the existence of that rule. Absent in Hayek's argument is the specification of a process by which a rule that is advantageous to the system is sustained in operation within that system." (Hodgson, 1993: 168)

It is a basic *assumption* in Hayek that individual actions serve a 'function' for the collective, that is, that in carrying through one's own interest, one is simultaneously (and more importantly) carrying through the interest of society; that actions performed by individuals are *automatically* **functional** for society. This is to assume all our problems away.

Hayek was talking about 'primitive human societies' but by the end of the paragraph from **SIP** just cited, this has become 'all the individuals of the [human] species which exist'. This implies that all the individuals that exist face identical circumstances, and therefore that the same set of rules is appropriate to all. All the individuals now alive behave in the same way: there *was* variety in behaviour in the past, but not now, as less

[124] As we shall see, this involves an impressive negation of the autonomy and rationality, indeed, of the *value*, of individuals.

[125] The page number is inconsistent with any of the works of Vanberg cited in Hodgson's bibliography.

well adapted behaviours have been eliminated. In the past some individuals behaved this way, now all do. The implication is that evolution is convergence to a destination state, rather than permanent flux. This brings together two threads: firstly, the Lucretian vision of evolution which Hume attributes to Philo in the *Dialogues*, the vision in which selection eliminates less well adapted forms, allowing the pre-existing fitter forms to displace them. Secondly, Hayek's ontogenetic version of evolution which sees evolution as the approach to a destination state. On the big issues, the emergence of the market and common law, evolution, and hence history, comes to an end with capitalism. Hayek is revealed as just as much a 'historicist' as thinkers such as Hegel, upon whom he has poured so much abuse ('Comte and Hegel', **CRS**: 365-400).

In a footnote to the passage just cited, Hayek refers the reader to the '[a]mple further illustrations of the kind of orders briefly sketched in this section ... in V.C. Wynne-Edwards, *Animal Dispersion in Relation to Social Behaviour*, Edinburgh, 1962' (**SIP**: 70 n 7). Hayek's point of view is exactly the group selectionist argument criticised by Richard Dawkins:

> "A group, such as a species or a population within a species, whose individual members are prepared to sacrifice themselves for the welfare of the group, may be less likely to go extinct than a rival group whose individual members place their own selfish interests first. Therefore the world becomes populated mainly by groups consisting of self-sacrificing individuals. This is the theory of 'group selection' [expressed] in a famous book by V.C. Wynne-Edwards [*Animal Dispersion in Relation to Social Behaviour*]... [But if] there is just one selfish rebel, prepared to exploit the altruism of the rest, then he, by definition, is more likely than they are to survive and have children. Each of these children will tend to inherit his selfish traits. After several generations of natural selection, the 'altruistic group' will be over-run by selfish individuals, and will be indistinguishable from the selfish group." (Dawkins, 1989a: 7-8)[126]

To illustrate the point we may cast our minds back to the case of the elephant seals, discussed in Chapter 2 of this thesis. We saw that the elephant seal species and populations within it are less efficient in exploiting their habitat than they could be, were their genes not caught up in a multi-player iterated prisoners' dilemma. Species do go extinct and many mammal species are endangered: it is not beyond the bounds of plausibility that one day the elephant seals will teeter over the edge of extinction when

[126] It is relevant to the title of this chapter to note that John Maynard Smith, adopting the same standpoint as Dawkins, chose to denounce the Hayekian, Wynne-Edwards group-selection argument as 'the old Panglossian fallacy that natural selection favours adaptations that are good for the species as a whole, rather than acting at the level of the individual' (cited in Dennett, 1995: 239).

they would not have done, had that prisoners' dilemma not existed. A cooperative solution to the dilemma cannot emerge and replicate, and displace the Pareto-inefficient defection solution, because no mechanism, no incentive structure exists to make cooperative behaviour Nash.

Evolution may very well take place at the group or species level; but for that to happen, there *has to be* a mechanism within the group which gives individuals adequate incentive to behave in the manner required for the group to prosper. Hayek attempts to suggest such an incentive:

> "The properties of the individuals which are significant for the existence and preservation of the group, and through this also for the existence and preservation of the individuals themselves, have been shaped by the selection of those from the individuals living in groups which at each stage of the evolution of the group tended to act according to such rules as made the group more efficient." (**SIP**: 72)

This contains two formulations; let us examine the first – 'the properties of the individuals which are significant for the existence and preservation of the group [are] *through* this also [significant] for the existence and preservation of the individuals themselves'. This formulation is incorrect: the properties of a *particular* individual, although they may well be significant for the group – when taken together with all the other similar individuals – can only be significant to that individual *as a result of its significance for society* to a vanishingly small degree, if the group itself is of any significant size. And, moreover, selection cannot distinguish between the effects on an individual's survival chances due to its own contribution to society and those due to someone else's.

Now we may turn to the second formulation – 'the properties of the individuals which are significant for the existence and preservation of the group ... have been shaped by the selection ... from the individuals living in groups which at each stage of the evolution of the group tended to act according to such rules as made the group more efficient'. This says that individuals' *behaviour*, that is, the rules that they follow, even if not the individuals themselves, is descended from the behaviour of individuals in groups which have been successful. Again, behaviour can be successful in two different ways: it can fulfil the objectives of those carrying out the activity, or it can be successful in terms of propagating itself, of getting itself copied by other individuals and groups. We have to

ask, what is the selection mechanism here? Some groups contain individuals which 'act according to such rules as made the group more efficient', others don't, and it is the former which constitute the template for future expansion. This is the group selection argument, and it is vulnerable to the Dawkins critique mentioned above: this other-regarding behaviour may be optimal for the group, but without individual incentives it is not Nash-ESS for individuals. Hence Hayek's attempt to provide an incentive structure to underpin the group selectionist argument assumes exactly what was to be demonstrated.

Now if groups were run by a central authority, then that authority might be able to set up an incentive structure to preserve the optimal behaviour pattern; the central authority changes the payoffs to individual actions to make the socially optimal outcome consistent with individual self-seeking behaviour. Such a group is also in a position to observe and copy collectively desirable behaviour patterns from other groups, which individual agents certainly are not able to do. As we have seen the group selectionist argument simply cannot work for spontaneous human societies.

Hayek recasts his evolutionary argument in teleological terms. There is, he says,

> "a sort of inversion of the relation between cause and effect in the sense that the structures possessing a kind of order will exist because the elements do what is necessary to secure the persistence of that order. The 'final cause' or 'purpose', i.e., the adaptation of the parts to the requirements of the whole, becomes a necessary part of the explanation of why structures of the kind exist: we are bound to explain the fact that the elements behave in a certain way by the circumstance than this sort of conduct is more likely to preserve the whole – on the preservation of which depends the preservation of the individuals, which would therefore not exist if they did not behave in this manner. A 'teleological' explanation is thus entirely in order so long as it does not imply design by a maker but merely the recognition that the kind of structure would not have perpetuated itself if it did not act in a manner likely to produce certain effects, and that it has evolved through those prevailing at each stage who did.
>
> "The reason why we are reluctant to describe such actions as purposive is that the order which will form as the result of these actions is of course in no sense 'part of the purpose' or of the motive of the acting individuals. The immediate cause, the impulse which drives them to act, will be something affecting them only; and it is merely because in doing so they are restrained by rules that an overall order results, while this consequence of observing these rules is wholly beyond their knowledge or intentions. In Adam Smith's classical phrase, man 'is led to promote an end which is no part of his intentions' just as the animal defending its territory has no idea that it thereby contributes to regulate the numbers of its species" (**SIP**: 77).

This is a revealing passage. Firstly, Hayek had already said that different macro orders are compatible with the same rules of micro behaviour, and different rules compatible

with the same order. Now he says that the preservation of the individuals depends on the preservation of the order such that they would not exist without it. Well, no doubt their existence depends on there being *some* order, but it by no means follows that their survival depends on the maintenance of the particular order obtaining at the present moment. This is the optimality assumption. Indeed, it is an extreme version of the optimality assumption, in which the present order is not just the best we can get, but the only one in which we could survive. This is repeated elsewhere:

> "If there exist recurrent and persistent structures of a certain type (i.e., showing a certain order), this is due to the elements responding to external influences which they are likely to encounter in a manner which brings about the preservation or restoration of this order; and on this, in turn, may be dependent the chances of the individuals to preserve themselves." (**SIP**: 71)

Hayek says here that an order exists because individuals behave in a way which preserves it. Perfectly true. But then he goes on to imply that their survival depends on the preservation of the order. This prompts the comments that (a) an order which allows a number of individuals to survive is not definitionally the best they can get: the social outcome could be suboptimal even if the behaviour of each individual in such an order is Nash, and (b) even if their survival depended on the individuals maintaining the order, the individuals concerned would still not do so unless it were *individually* rational to do so.

Secondly, we can see here how thin Hayek's individualism is: the 'final cause' or purpose of a society is 'the adaptation of the parts to the requirements of the whole' – the adaptation of individuals to the requirements of society. We explain the way people behave 'by the circumstance that this sort of conduct is more likely to preserve the whole'. But this completely neglects any reason why self-seeking individuals should do any such thing: 'that the order which will form as the result of these actions is of course in no sense 'part of the purpose' or of the motive of the acting individuals'. The implication is that they don't know what they are doing, they are ignorant or irrational: the 'consequence of observing these rules is wholly beyond their knowledge or intentions'. It is interesting that, quite illegitimately, Hayek now draws a distinction between the actions individuals would take to satisfy their own drives in the absence of rules, and the rules themselves, which take on the status of a restraint. This breaks with his own procedure, which regards all the actions of the individuals, in so far as they are

regular, as the execution of rules. In the version which sees rules as restraints, they can only be imposed by an external authority or force.

Thirdly, Hayek claims that teleological explanation is in order, as long as it does not imply a deity. However, so long as the explanation includes no motivation for individuals to behave in this socially desirable way, then it must depend on supernatural intervention. Hayek replicates Smith's schema with all individual behaviour revolving round and organised in accordance with the general interest of society. But it is Hamlet without the prince: the supernal agency arranging all this is missing. Hayek's group selectionism cannot stand in for this agency, indeed, without that agency it makes no sense: for group selection *assumes* individual and general interests already to be reconciled. Contrary to Hayek's declaration, teleological explanation is only valid if it is, at least in principle, possible to specify a mechanism by which actions at the level at which the teleology is thought to hold can be reconciled with the interests of substrate levels.

Lastly, it is very fitting that Hayek should cite Smith, at he does at the end, and explicitly link him to a group selectionist standpoint: a footnote at the end of the last sentence of the passage cited refers us again to Wynne-Edwards (1962). The mythical animal fulfilling its own interests at the same time as it fulfils those of a higher entity of which it is the unwitting subordinate part[127] is just that – mythical. This is a version of Hegel's cunning of history.[128] Here, it is not even the interest of the population or the species which is the unintended beneficiary, but the whole ecosystem. Smith's 'famous phrase' was actually 'every individual ... is ... led by an invisible hand to promote an end which was no part of his intention.' Hayek censors out the 'invisible hand' as it exposes what he is doing, namely assuming a mystical force which will reconcile our conflicting plans spontaneously to produce a beneficent order at the macro level.

---

[127] Lest the reader think that Hayek's claim about animals regulating their numbers is a slip, I should point out that he repeats this claim – this time about humans – in **KES**: the growth of world population, he says, 'is a process which is self-regulating' (**KES**: 52).

[128] '[S]tates, nations, and individuals arise animated by their particular determinate principle ... While their consciousness is limited to these and they are absorbed in their mundane interests, they are all the time the unconscious tools and organs of the world mind at work within them' (Hegel, 1952: §344). '[I]ndividuals ... are the living instruments of what is in substance the deed of the world mind ... though it is concealed from them and is not their aim and object' (Hegel, 1952: §348).

This subsection has shown that Hayek's account of evolution is essentially the group selectionist argument which has been rightly subjected to considerable criticism in the biological domain. We have also seen that this is closely associated with the Smithian Panglossianism of which Hayek is a principal twentieth century representative. Hayek's twists and turns, his falsifications and distortions, his vagueness and obscurity, can be accounted for by his dedication to a particular political programme and his apparent willingness cynically to prefer the support of that programme to all scholarly values of truth and consistency.

5.4.5 *Have Sober and Wilson rescued group selection?*

It is now necessary to turn to an important issue which has been raised in connection with the points made above (Vromen, personal communication), namely whether Sober and Wilson (1998) have rescued the notion of group selection deployed by Hayek. Clearly, if they have, then much of the argument of the present chapter collapses. Now, Sober and Wilson certainly perform a signal service by clarifying the conditions under which what we might legitimately call 'group selection' could be sustained, and ruling out crude, reductionist readings of Darwinian writers such as Dawkins. Whether they have anything sensible to say about those Darwinian writers themselves, is another question which it would be inappropriate to explore here. Suffice it to say that Sober and Wilson significantly muddy the water by their systematic usage, from the front cover of their book to its last page, of the terms 'unselfish' and 'altruistic behaviour' when they mean *cooperative* behaviour. If by altruism we mean, as we clearly should mean, other-regarding behaviour engaged in to one's own cost, then altruism is not selected for, full stop. Sober and Wilson's book is *not*, what its title implies, about the evolution of *unselfish* behaviour: it is about the evolution of cooperation, that is, the circumstances in which cooperation is the outcome of individual self-seeking behaviour. It is individually advantageous to engage in cooperation when the benefits accruing to oneself as a result of the cooperative behaviour exceed its costs. This can happen, for example, when membership of the group of cooperators is a function of one's own propensity to cooperate. In economic terms, cooperation can arise spontaneously when the externality generated by cooperation is internalised.

Once this is understood, a sensible investigation of 'group selection' can begin. Group selection can take place when there is variation between groups, with some exhibiting more, and some less, cooperative behaviour on the part of individual group members, where such cooperative individual behaviour is underpinned by some mechanism ensuring that cooperation is in the interest of selfish individuals. Then more cooperative groups will tend to displace less cooperative ones: cooperative behaviour is selected for. There is nothing in this which contradicts Darwinianism of the Dawkins variety. Indeed, selfish organism theory is a variety of group selection where the group is the community of genes embodied in the organism, as Dawkins makes clear – and as Sober and Wilson concede he makes clear: 'group selection is a question about vehicles' (1998: 92) – though, somehow, Dawkins's, view here is transmuted into a 'concession' to group selectionism. The opposition of Darwinians to 'group selection' is not to the idea that such mechanisms, reconciling individual and group interests may exist, but to the Panglossian notion propagated by Wynne-Edwards and his followers, such as Hayek, that group selection can exist in the absence of such mechanisms.

The critical link between individual and group interest is what I will call *connation*. It is worth quoting Dawkins, from *The Extended Phenotype*, at length on the issue:

> "each gene is fighting only its alleles at the same locus, and it will unite with genes at other loci only in so far as doing so assists it in its selfish war against its own alleles. A fluke gene may 'unite' with other fluke genes in this way but, equally, if it was convenient to do so, it might unite with particular snail genes. And if it remains true that snail genes are in practice selected to work together with each other and against an opposing gang of fluke genes, the reason is only that snail genes tend to gain from the same events in the world as do other snail genes. Fluke genes tend to gain from other events. And the real reason why snail genes stand to gain from the same events as each other, while fluke genes stand to gain from a different set of events, is simply this: all snail genes *share the same route into the next generation* – snail gametes. All fluke genes, on the other hand, must use a different route, fluke cercariae, to get into the next generation. It is this fact alone which 'unites' snail genes against fluke genes and vice versa. If it were the case that the parasite genes passed out of the host's body inside the host's gametes, things might be very different. The interests of host genes and parasite genes … would then be very much closer than in the case of fluke and snail." (Dawkins, 1989b: 221-222, my emphasis)

This 'sharing the same route' – what I call *connation* – is absolutely critical. The difference between parasitism and symbiosis, between a liver fluke and a mitochondrion, rests on shared destiny. The mitochondrion can only place copies of itself in the next generation by aiding its host, the animal cell; the liver fluke is not so restricted and exploits its host, damaging its host's interests to its own advantage. So we need to think about the routes to the future available to social structures, the phenotypic expression of

meme complexes. Clearly, those routes are utterly different from the route by which humans reproduce and so those meme complexes cannot be relied upon spontaneously to share interests with humans. The meme complexes embodied in the social institutions which emerge spontaneously can be expected sometimes to be symbiotic with, and sometimes parasitic on, their human hosts. Some interesting examples are given in Blackmore (1999).

If the social institutions which emerge spontaneously from the evolutionary process can be parasitic, then the presumption in favour of a *laissez-faire* policy framework is undermined. Spontaneously emerging forms may need to be modified or replaced by institutions adapted to human interests. In this context, we can return to the question this section seeks to answer: is Hayek a *Panglossian* evolutionary theorist? Hayek invokes 'group selection' – but is it a form of 'group selection' consistent with Darwinian, Dawkinsian thought, or does it rather, as Hayek himself clearly believes, stand in the Wynne-Edwards tradition? Given Hayek's belief, his assumption without explanation or justification, that institutions are selected 'on the basis of their human survival-value', rather than on the basis of their meme-complex survival value, the conclusion has to be that nothing in Sober and Wilson gives cause to modify our verdict, that Hayek's evolutionary theory is, indeed, Panglossian. The last word here goes to Sober and Wilson: against 'group-level functionalism' they warn that

> "One can never simply assume that higher-level units such as cultures, societies or biological ecosystems must be well-functioning organic wholes. Higher-level functionalism always requires special conditions and is vulnerable to subversion from within." (Sober and Wilson 1998: 11)

5.5 Hayek's anti-individualism

A theme of the previous chapter, on Adam Smith, was that although his policy prescription was one of individualism, this was linked to a methodological holism and combined with some distinctly anti-individualist social attitudes. This section will present the case for a very similar verdict on Hayek. Hayek's basically holistic methodology has already been discussed. Here I want to draw attention to some strongly anti-individualist sentiments in Hayek's position.

We saw in the previous section how Hayek made the 'final cause' or purpose of human activity 'the adaptation of the parts to the requirements of the whole' – the adaptation, that is, of individuals to the requirements of society. Hayek repeats this elsewhere:

> 'Frequently the behaviour of the individual is determined by his success in maintaining himself as part of a certain system within which it is … the whole system to which he has adapted that determines his behaviour. For this reason, value … can only be understood as the determinant of what people must do to maintain the overall structure." (**KES**: 36)

So the individual has to adapt to the system, and it is the system that determines the behaviour of the individual. What we think of as *value* is just a signal to us to act in accordance with the needs of the system. As statements of fact, these assertions, like Smith's on government as a means of defence of the rich against the poor, sound very radical and subversive. As normative statements, as statements of what *should* be the case, they take on a more sinister coloration. It is only Hayek's Panglossian and harmonic world view, in which individual interests are illegitimately *identified* with those of the system, which allows him to say this. If the system is optimal for the individuals composing the society, then it is acceptable to require the individual to adapt to it. If the system is necessarily optimal, *then there are **no** social problems*. Apparent problems are problems for social scientists only – a challenge for them to explain away. To assure individuals suffering the consequences of macro level social problems that such problems do not exist, is to treat those individuals with contempt.

We also saw that in the non-harmonic world in which we actually live, the group selectionist argument that individuals behave in the social interest rather than in their own interest implies that such individuals are unaware of the consequences of their actions: they don't know what they are doing – they are ignorant or irrational[129] (or both). If individuals were rational and adequately informed, they would follow their own interests and the group interest would remain unimplemented. '[T]he order which will form as the

[129] Along with anti-individualism, anti-rationalism is a major theme in Hayek: he endorses 'the true, antirationalistic … individualism' (**IEO**: 11). 'Hayek … retained a belief in the inescapably limited nature of human knowledge of society, which would seem to owe much to Kant … and hence the inappropriateness of trying to reconstruct society as if such knowledge existed' (Tomlinson, 1990: 10). I argued in the previous chapter that in many ways Kant could be seen as complementing Smith and Hume, particularly in terms of the desire, shared by all three, to limit the legitimate scope of reason. However, I cannot explore this theme further here.

result of [individual] actions is of course in no sense 'part of the purpose' or of the motive of the acting individuals', Hayek asserts in the passage cited in the previous section, the 'consequence of observing these rules is wholly beyond their knowledge or intentions' (**SIP**: 77). Institutions arise, he says, 'from the separate actions of many men *who did not know what they were doing*' (**COL**: 58-59, emphasis added).

The irrationality thesis is stated even more bluntly in **KES**:

> "to preserve rules of conduct whose functions they [the groups which obeyed these rules] did not understand, they drew upon the aid of supernatural sanctions ... we owe it to mystical beliefs, that we preserved a tradition which was beneficial to us. Thus we owe our civilisation to beliefs which are not true" (**KES**: 48). "The antirationalistic approach … regards man not as a highly rational and intelligent but as an irrational and fallible being" (**IEO**: 8)

This is a very close parallel to the argument in Smith, that individuals are deceived by nature for the good of society. Smith says that the individual is deceived into thinking, for example, that wealth will make him happy, and as a result slaves to accumulate wealth, without gaining the expected increase in happiness. This aided society by keeping the wheels of industry turning. Just so, individuals in the Hayekian story are misled by bizarre superstitious notions which upset their lives and cause them untold suffering, so that civilisation may prosper.

When we are looking at the status of individuals in a proposed or actual social system, a key diagnostic turns on the question, what constitutes the ends and what the means of the system. Hayek is clear on this: 'Like all other values, our morals are … part of the ends which the instrument of our intellect has been developed to serve … the system of values into which we are born supplies the ends which our reason must serve' (**COL**: 63). Again, this is an extremely illiberal and anti-individualist standpoint. Contrary to what Hayek says, the ends we serve arise within us as our natural desires, they are not something we meet externally as something to which we must subordinate ourselves. The system of values within which we work tells us the constraints our actions must satisfy; it tells us one way of coordinating our efforts to fulfil our individual goals so that we all may fulfil them more effectively. It is merely a means to the end of human happiness.

To argue that the culturally determined system of what at some point in time is taken as right and wrong, 'the system of values into which we are born', is an *end*, rather than a

means to our ends, does two things. Firstly, and this is Hayek's goal here, it makes the system of values existing here and now absolute instead of relative. Its survival, the prevention of its succession by another, alternative system becomes a good in its own right, an end which we should pursue regardless of its utility for us. Secondly, instead of seeking to pursue our own goals, instead of seeking to maximise our own welfare, we are enjoined to pursue, and to subordinate our own interests to, the interest of some non-human entity. This is precisely the establishment of 'Society' 'as an abstraction *vis-à-vis* the individual', which Amartya Sen, echoing Karl Marx, warned against (Marx[130] cited in Sen 1970: 1 n 1).

Further evidence on Hayek's anti-individualist outlook emerges in his criticism of (unnamed) rationalists for their alleged tendency to replace the word 'moral' with 'social' as in 'social conscience' instead of 'moral conscience' or just 'conscience':

> "They are in effect saying that our action should be guided by a full understanding of the functioning of the social process and that it should be our aim, through conscious assessment of the concrete facts of the situation, to produce a foreseeable result which they describe as the "social good" … The curious thing is that this appeal to the "social" really involves a demand that individual intelligence, rather than rules evolved by society, should guide individual action – that men should dispense with the use of what could truly be called "social" (in the sense of being a product of the impersonal process of society) and should rely on their individual judgment of the particular case. The preference for "social considerations" over the adherence to moral rules is, therefore, ultimately the result of a contempt for what really is a social phenomenon and of a belief in the superior powers of individual human reason." (**COL**: 65)

What is remarkable, in the present context, about this passage, is that it criticises a supposed 'appeal to the social', not on the grounds that we should not appeal to the social, but because it is really an appeal to the individual. He is criticising these 'rationalists' for their individualism, for their 'contempt' for the social! Once again, he is expressing a very anti-individualist view. Individuals are not to be trusted to work things out for themselves, but have to submit to traditional rules whose rationale is unknown. This is very similar to the anti-Protestant view of the Catholic Church, as can be seen by re-casting the passage in terms of the Catholic critique of Protestantism:

> Protestantism involved a demand that individual religious conscience, rather than the doctrines of the Church, should guide individual action – that men should dispense with

[130] The reference is to Karl Marx (1844/1959) *Economic and Philosophical Manuscripts of 1844* p104; reprinted in Marx (1975: 350).

> what could truly be called 'religious' (in the sense of being a work of God, rather than of humans) and should rely on their individual judgement of the particular case. The preference for 'considerations of religious conscience' over adherence to religious rules, is therefore ultimately the result of a contempt for genuine religion and a belief in the superior powers of individual human reason.

Hayek's argument, as well as being explicitly anti-individualist, is also highly misleading. Although the thinking involved in review, revision and reform of the rules under which we are to live, must take place in the brains of individual people, they are by no means acting *as* individuals, as isolated atoms, but rather as members of the community, debating with others in ways supervised by the public and according to rules invigilated by the public, many individuals contributing factual and analytical material to the discussion so that the whole thing is far more complex and profound than any one individual could have managed in isolation. Reform is a truly social enterprise. It is the superior powers of social, not individual, human reason which are being relied upon.

Indeed, when we get to the bottom of it, we find that Hayek's individualism really has nothing at all to do with … individualism:

> "[I]ndividualism ... is primarily a *theory* of society, an attempt to understand the forces which determine the social life of man, and only in the second instance a set of political maxims derived from this view of society. This fact by itself should be sufficient to refute the silliest of the common misunderstandings: the belief that individualism postulates (or bases its arguments on the assumption of) the existence of isolated or self-contained individuals, instead of starting from men whose whole nature and character is determined by their existence in society." (**IEO**: 6)

Just about everything Hayek says here is false. Contrary to what he says, individualism is indeed primarily a political doctrine, and only secondarily a theoretical underpinning for it. And whatever its status in general, it certainly is primarily political rather than theoretical in Hayek's own case. It is also false to claim that individualism does not start from the assumption of 'the existence of isolated … individuals'. The implication, that the whole nature and character of people is determined by their social being, their being in society, is of the essence of the holist standpoint: it is very far indeed from the individualism of neoclassical economics, which starts from isolated individuals – 'a collection of Robinson Crusoes' as Friedman says. The starting point of 'men whose whole nature and character is determined by their existence in society' echoes his statement in **CRS** quoted by Toynbee, that individuals are but foci in the network of relationships constituting society.

Smith and Hayek are essentially conservatives[131], and from this all else flows. In both cases nostalgia for stability plays a key rôle in their psychologies. Smith wanted to preserve the fixed orders and ranks of society; Hayek feels the same about traditions:

> 'There probably never has existed a genuine belief in freedom, and there has certainly been no successful attempt to operate a free society, without a genuine reverence for grown institutions, for customs and habits ... [A] successful free society will always in a large measure be a tradition-bound society' (**COL**: 61). '[S]ubmission to undesigned rules and conventions whose significance and importance we largely do not understand ... [and] reverence for the traditional ... is indispensable for the working of a free society' (**COL**: 63).

Hayek's vision of the free society is thus one in which individuals, as lost and uncomprehending of the world they inhabit as children, voluntarily submit themselves to the greater wisdom embodied in tradition, or, failing such voluntary submission, one in which submission is brought about by compulsion. The individualist order requires

> "that the individual ... must be willing ... to submit to conventions ... which to him will often appear unintelligible and irrational .... The willingness to submit to such rules ... is ... an indispensable condition if it is to be possible to dispense with compulsion." (**IEO**: 22) "[T]he fundamental attitude of true individualism is one of humility toward the processes by which mankind has achieved things which have not been designed or understood by any individual and are indeed greater than individual minds." (**IEO**: 32)

Hayek's 'fundamental attitude', that is, is a humbling of ourselves before the fetish of the spontaneous.

The verdict, then, on the relation of Hayek's philosophy to individualism, is this. In the sphere of policy, he is indeed an individualist, regarding individual freedom as a paramount value. Just as Smith regarded social order as more important than justice, as 'of more importance than even the relief of the miserable', so Hayek, too, regards

[131] As O'Hear says, 'Hayek's epistemology thus leads to a defence of moral and institutional conservatism' (O'Hear, 1995: 336) – but only because that is where it is intended to lead. 'Hayek is a conservative most obviously in his veneration of tradition, and in his belief that most of the forms of social organisation that exist in capitalist economies are the product of a long process of evolution. Thus private property is not something to be defended purely on intellectual or rational grounds, but also as the successful product of a process of evolutionary 'sifting' of institutions over the centuries. Hayek also tends to conservatism in social affairs, notable in his attitude to religion and the family ...' (Tomlinson, 1990: viii-ix).

individual freedom as more important than social justice, and frankly says so: 'We must face the fact that the preservation of individual freedom is incompatible with a full satisfaction of our views of distributive justice' (**IEO**: 22). On the other hand, there is nothing specifically individualist about the methodology which Hayek employs to investigate the world. On the contrary, and again like Smith, and despite equivocation and inconsistency, he does recognise the fundamental necessity of a holistic scientific methodology to understand the world. Finally we have seen that the two writers, Smith and Hayek, also share a set of authoritarian conservative[132] social attitudes which are extremely thin on respect for the rationality, the autonomy, and the fate of individuals.

5.6 Conclusion

Just as Smith faced, and refused to face up to, the problem of how, given that human nature is natural, anything that humans do – especially state activity – could be unnatural, so Hayek faces a problem of how any human behaviour, including state intervention, can fail to be the result of an evolutionary process. And this is fatal for his policy prescription, just as it was for Smith. Smith's Panglossianism depended on the view that everything natural was God-given and hence good, while everything artificial was human-made and hence fallible. Hence his opposition to the visible hand of state intervention and his belief in the optimality of the outcomes supplied us by the invisible hand of a benign deity. But if all human strengths and weaknesses alike are themselves natural and God-given, then state intervention, too, must play its necessary part in the scheme of things. Hayek's Panglossianism depends on the view that spontaneous evolution of institutions automatically generates optimal outcomes while rationalistic intervention is both unnecessary and perverse in its effects. But the institutions by means of which the society as a whole acts to coordinate agent actions, and to improve on the inefficiencies of spontaneous outcomes, are themselves the outcome of a spontaneous process of evolution – and thus, in a consistent Hayekian view, optimal. To intervene on principle to prevent them from doing their job, without regard to the actual content of

[132] Hayek's authoritarian conservatism is shown, for example, in his disdain for democracy: 'We have no intention ... of making a fetish of democracy. It may well be true that our generation talks and thinks too much of democracy and too little of the values which it serves … Nor must we forget that there has often been much more cultural and spiritual freedom under an autocratic rule than under some democracies … democratic government might be as oppressive as the worst dictatorship' (**RTS**: 52).

what they were doing, would be a clear instance of the rationalism which he has spent his life combating.

Appendix: Bibliographical note

Hayek's books exist in several editions and papers are reprinted in various collections. Navigating his works can be confusing. This appendix will set out a bibliography of Hayek's works and explain which versions have been consulted for this chapter. This is not a comprehensive bibliography of Hayek's work, which would be a much more ambitious task, but an indication of what has been consulted for this chapter, an explanation of the abbreviations used in the text, and an explanation of which version page numbers of works by Hayek cited here refer to. For fuller bibliographical information, refer to:

a Machlup (1977: 51-59). Although only going up to 1977, this gives details of 173 publications by Hayek;
b a comprehensive bibliography of Hayek's writings in J Gray (1985, 2e: 1986) *Hayek on Liberty* Oxford: Blackwell;
c a more brief but up to date listing contained in the bibliography to Hodgson (1993: 325-326).

Only items referenced in the text are included. Each of the following items is preceded by the abbreviation used for it in the text, where relevant. Only the books are included in the bibliography at the end of the thesis.

Books by Hayek

RTS 1944 *The Road to Serfdom* London: George Routledge & Sons

IEO 1948 *Individualism and Economic Order* Chicago: University of Chicago Press; (reprinted, 1949, London: Routledge). References are to the 1949 reprint.

TSO 1952 *The Sensory Order. An Inquiry into the Foundations of Theoretical Psychology* London: Routledge & Kegan Paul. Originally drafted in approx 1920.

CRS 1e: 1952 *The Counter-Revolution of Science. Studies on the Abuse of Reason* Glencoe, Illinois (second edition, 1979, Indianapolis: Liberty*Press*). Reprinted articles. References are to the second edition of 1979, which contains additional prefatory material.

COL 1960 *The Constitution of Liberty* London: Routledge & Kegan Paul

SIP 1967a *Studies in Philosophy, Politics and Economics* London: Routledge & Kegan Paul. Reprinted articles.

NSP 1978a *New Studies in Philosophy, Politics, Economics and the History of Ideas* London: Routledge & Kegan Paul. Reprinted articles.

TBT 2e: 1978b (ed SR Shenoy) *A Tiger by the Tail* London: Institute for Economic Affairs. An edited collection of extracts from other works.

LLL (1982) *Law, Legislation and Liberty. A new statement of the liberal principles of justice and political economy* London: Routledge. 'Hayek's major statement of political philosophy ... Hayek's comprehensive study on the basic principles of the political order of a free society', according to the publishers comments. One-volume version with corrections and revised preface of: *Law Legislation and Liberty* (1973-1979) in three volumes:

Vol 1 (1973) *Rules and Order*
Vol 2 (1976) *The Mirage of Social Justice*
Vol 3 (1979) *The Political Order of a Free People*

KES 1983 *Knowledge, Evolution and Society* London: Adam Smith Institute. Four lectures given in the United States in the late 1970s, plus prefatory essays by Eamonn Butler and Arthur Shenfield.

Journal articles and lectures

1935 'The Nature and History of the Problem' in FA Hayek (ed) (1935) *Collectivist Economic Planning* London: George Routledge & Sons. Reprinted as Ch VII, 'Socialist Calculation I: *The Nature and History of the Problem*', of **IEO**: 119-147. References are to the 1949 reprint of **IEO**.

1941 'The Counter-Revolution of Science' *Economica*. Reprinted as Part II of **CRS**: 183-363. References are to the 1979 edition of **CRS**.

1942-44 'Scientism and the Study of Society' *Economica*. Reprinted as Part I of **CRS**: 17-182. References are to the 1979 edition of **CRS**.

1945 'Individualism: True and False' Finlay Lecture, University College, Dublin; first published, 1946, Dublin: Hodges, Fidges & Co, and Oxford: Blackwell; reprinted as Ch 1 of **IEO**: 1-32. References are to the 1949 reprint of **IEO**.

1951 'Comte and Hegel' *Measure* June. Reprinted as Part III of **CRS**: 365-400. References are to the 1979 edition of **CRS**.

Mandeville 1967b 'Dr Bernard Mandeville' *Proceedings of the British Academy* Volume LII: 125-141 (British Academy 'Lecture on a Master Mind' read 23 March 1966) London: Oxford University Press. Reprinted as Ch 15 of **NSP**: 249-266. References are to the *Proceedings* version..

1978a 'Coping with Ignorance' *Imprimis* Vol 7 No 7 (Hillsdale College, Michigan). Reprinted in **KES**: 17-27. References are to the **KES** version.

1978b 'Science and Socialism' American Enterprise Institute for Public Policy Research. Reprinted in **KES**: 28-37. References are to the **KES** version.

1978c 'The Reactionary Nature of the Socialist Conception' Hoover Institution, Stanford University. Reprinted in **KES**: 38-44. References are to the **KES** version.

1983 'Our Moral Heritage' Heritage Foundation. Reprinted in **KES**: 45-57. References are to the **KES** version.

Chapter 6 Maynard Keynes's methodological standpoint and policy prescription

6.1 Introduction

In the previous chapters of this thesis I have tried to show two things: Firstly, that in a world of partially overlapping and partially conflicting interests there is good reason to doubt that self-seeking behaviour at the micro-level will spontaneously lead to desirable social outcomes at the macro-level. And, secondly, that some sophisticated economic writers who would like us to rely on the spontaneous interaction of self-seeking agents, writers advocating a *laissez-faire* policy prescription, have proposed various 'invisible hand' mechanisms which can, in their view, be relied upon to 'educe good from ill'. Smith, I argued, defended the 'simple system of natural liberty' as giving the greatest scope to the unfolding of God's will and the working out of 'natural', providential processes free of interference by 'artificial' state intervention – the expression not of divine order but of fallible human reason. Hayek, adopting a similar policy stance, based it in an evolutionary process in which those institutional forms best adapted to reconciling individual interests would, he believed, spontaneously be selected for in the inter-group struggle for survival.

The purpose of the present chapter is to cast a light on this issue from another direction by displaying an example of the policy consequences of adopting an alternative methodological stance. The argument of the chapter is that (a) staying within the holistic framework of Smith and Hayek, but (b) rejecting their invisible hand mechanisms, leads (c) to the rejection of their reductionist *laissez-faire* policy stance as well.

The structure of the chapter is as follows. In Section 6.2 I show Keynes's view of the historical role of capitalism and his analysis of its pathology, rooted in what we would

now refer to as the prisoners' dilemma. The section draws attention to the fundamental significance of his methodological standpoint. This lays the basis for a consideration of his policy prescription in the following two sections. Section 6.3 looks at two very important aspects to the question, spelling out Keynes's call for planning, and explaining exactly what he meant by this. Before doing so however – and this is the other key aspect to the question – the section examines Keynes's *class* standpoint, showing the critical role he expected his own class, the 'educated *bourgeoisie*', to play in the reform process he mapped out. Section 6.4 adds a further layer to the consideration of Keynes's policy prescription, drawing out the distinction , but also the intimate connection, between, on the one hand, micro-level individualism (the 'Manchester System'), and, on the other, the macro-level collective action ('planning') required to preserve it. Section 6.5 concludes by considering Keynes in relation to the themes of the thesis – Smith and Hayek, holism, reductionism and the invisible hand.

6.2 Keynes's historical perspective

Whereas, for Smith and Hayek capitalist individualism is the terminus of an *ontogenetic* process, for Keynes it is something transitional, something with a historical and conditional validity. Keynes's historical perspective is thus consistent with a *phylogenetic* evolutionary stance. *Laissez-faire* in Keynes's conception performed a vital historical role, carrying us from an Era of Scarcity to an Era of Abundance. It was precisely because it had substantially fulfilled that role that it had become counter-productive. The point is controversial. Joan Robinson claimed that Keynes 'saw the capitalist system as ... a phase in historical development' (Robinson, 1962: 71); Geoffrey Pilling, on the other hand, criticising both Keynes and Robinson, writes that 'It is just this historical conception of capitalism which in absent in Keynes' (Pilling, 1986: 35, and n1). The purpose of this section is to show that Robinson was right, and Pilling wrong[133], on

[133] It has been argued (Geoffrey Kay, personal communication) that both Robinson and Pilling were right, here, since they meant different things by 'historical'. However, Pilling clearly believed he was talking about the same thing as Robinson and that she was wrong: 'One cannot ... accept Joan Robinson's confident assertion about Keynes' – ie the assertion that Keynes viewed capitalism as a phase in historical development (Pilling, 1986: 57). On this latter point Pilling was clearly wrong and Robinson right. However, one might wish to go further and argue that there was a sense in which Pilling was right about Keynes being unhistorical because he, Keynes, had a notion of capital which was

this point: to establish Keynes's conception of the historical role of capitalism – though not, I should emphasis, necessarily to defend it.

In order to establish Keynes's view of the historical and historically limited role of *laissez-faire*, we must say something about his conception of the historical context, that is, about his periodisation of history. I have attempted to reconstruct Keynes's view here by rereading 'Am I a Liberal?' (Keynes, 1925b) in the light of his later works, in particular 'Economic Possibilities for our Grandchildren' (Keynes, 1930c), and Book VI of the *General Theory*, 'Short Notes Suggested by the General Theory' (**GT**: 313-384).

The first great era in Keynes's scheme takes in prehistoric, ancient and medieval times. In the 'Era of Scarcity' (**EP**: 304) production is overwhelmingly production for the sake of consumption, indeed subsistence purposes: 'The economic problem, the struggle for subsistence, always has been hitherto the primary, most pressing problem of the human race – not only of the human race, but of the whole of the biological kingdom from the beginning of life in its most primitive forms' (**EP**: 326-7). During the Era of Scarcity, there is an overwhelming obstacle to the accumulation of capital in the form of uncertainty driving the marginal efficiency of capital (*MEC*) below the rate of interest (*i*):

> "The destruction of the inducement to invest by an excessive liquidity-preference was the outstanding evil, the prime impediment to the growth of wealth, in the ancient and medieval worlds. And naturally so, since certain of the risks and hazards of economic life diminish the marginal efficiency of capital while others serve to increase the preference for liquidity. In a world, therefore, which no one reckoned to be safe, it was almost inevitable that the rate of interest ... would rise too high to permit of an adequate inducement to invest." (**GT**: 351)

At – so to speak – the other end of history from the era of scarcity, in the near future, lies 'our destination of economic bliss' (**EP**: 331), 'economic paradise' (**EP**: 268), 'the age of leisure and abundance' (**EP**: 328). 'The economic problem may be solved, or be at least within sight of solution, within a hundred years ... the economic problem ... is not the permanent problem of the human race' (**EP**: 326). The essence of the era of abundance is that 'needs are satisfied in the sense that we prefer to devote our further energies to non-economic purposes' (**EP**: 326). Thus, we may note in passing, production here, too,

---

in some sense timeless. But this goes beyond the question of whether Keynes had a historical view of capitalism, and slides over into another question, namely whether he had what one believes to have been the *correct* historical view of capitalism. That is not an issue I am addressing here.

is for the sake of consumption in the broadest sense: 'for the first time since his creation man will be faced with his real, his permanent problem – how to use his freedom from pressing economic cares …. to live wisely and agreeably and well' (**EP**: 328).

The following year, in the Preface (dated 1931) to *Essays in Persuasion* (**EP**), a collection of essays spanning a dozen years, this approaching liberation from economic care has become his 'central thesis throughout': 'the day is not far off when the economic problem will take the back seat where it belongs, and that the arena of the heart and head will be occupied ... by our real problems – the problems of life and human relations, of creation and behaviour and religion' (**EP**: xviii). This messianic strand, though expressed in more sober language, still plays a central, and, indeed, even more urgent, role in the *General Theory*. The age of abundance now appears as the 'quasi-stationary community' (**GT**: 220), and is to be attained, not in our grandchildren's time, but 'within a single generation' (**GT**: 220), 'say within twenty-five years or less' (**GT**: 324).

The modern period, the third historical division or 'economic order' (**EP**: 304) in Keynes's schema, is the age of capitalism. This period does not have the fundamental, self-sufficient character of the other two epochs, but is simply the period of transition from the one to the other. As such it is not an end in itself but a means to an end lying beyond itself, namely our entry into the 'economic paradise'. Hence our judgement of capitalism must refer, not to how pleasant or otherwise it may be, but to its efficacy in achieving that end:

> "Many people, who are really objecting to capitalism as a way of life, argue as though they were objecting to it on the ground of its inefficiency in attaining its own objects … For my part I think that capitalism, wisely managed, can probably be made more efficient for attaining economic ends than any alternative system yet in sight, but that in itself it is in many ways extremely objectionable." (**EP**: 294)

Capitalism achieves this end, in Keynes's view, by means of the accumulation of capital, and, for Keynes, the rate of capital accumulation is the measure of the rate of our approach to the economic paradise. To denote the motives to this accumulation of capital, Keynes spoke of 'compound interest' (**EP**: 324) and 'purposiveness' (**EP**: 329). By this latter peculiar expression, 'purposiveness', he merely means money-making as an end in itself, saving, ostensibly for future consumption, but actually for the sake of

accumulating claims on future production; saving not in order to enjoy the deferred consumption later, but in order to secure a stream of unearned income. Keynes analyses 'purposiveness' psychologically as an attempt to gain immortality by projecting one's actions into an indefinite future by means of an infinite regress:

> 'purposiveness means that we are more concerned with the remote future results of our actions than with their own quality ... the purposive man is always trying to secure a spurious and delusive immortality for his acts by pushing his interest in them forward in time." (**EP**: 330)

Just as Marx, in the Communist Manifesto, for example, was outspoken in his praise for the achievements of capitalism (Marx and Engels, 1976: 489), Keynes, too, paid tribute to those achievements: 'In the nineteenth century this epoch culminated gloriously in the victories of *laissez-faire* and historic Liberalism.' (**EP**: 304) The accumulation of capital depended upon the freedoms of *laissez-faire* – in particular, private properly in the means of production and unrestricted scope for the operation of market forces:

> "The system worked, throughout Europe, with an extraordinary success and facilitated the growth of wealth on an unprecedented scale. To save and invest became at once the duty and the delight of a class. The savings were seldom drawn on, and, accumulating at compound interest, made possible the material triumphs which we now all take for granted. The morals, the politics, the literature, and the religion of the age joined in a grand conspiracy for the promotion of saving." (**EP**: 62)

In one of his essays on Liberalism, where he applied to his own views the term 'New Liberalism' (**EP**: 305), he remarked that 'old-fashioned individualism and laissez-faire ... contributed to the success of the nineteenth-century ... I should have belonged to this party [sc the Liberal Party] if I had been born a hundred year earlier' (**EP**: 300-301).

While recognising the historical necessity and legitimacy of the *laissez-faire* system, and appreciating the benefits of its 'material triumphs', Keynes nevertheless deprecated this subversion of morals: 'we have exalted some of the most distasteful of human qualities into the position of highest virtues.' (**EP**: 329) Keynes is here protesting against the fact that capitalism requires, and *laissez-faire* permits, the transformation of the economy from production for the sake of consumption to production for the sake of profit, for the sake of the accumulation of wealth. Saving for the sake of future consumption Keynes can put up with; saving in order 'to exploit the scarcity value of capital' (**GT**: 376) is morally reprehensible.

That mankind has had to depend on this sort of egoistic materialism in order to raise itself from scarcity to abundance had, according to Keynes, had widespread deleterious consequences. Defining capitalism as 'egotistic atomism', he complains that: 'modern capitalism is absolutely irreligious, without internal union, without much public spirit, often ... a mere congeries of possessors and pursuers' (**EP**: 267). 'I think that Capitalism ... in itself is in many ways extremely objectionable' (**EP**: 294).

> "[T]he moral problem of our age is concerned with the love of money, with the habitual appeal to the money motive in nine-tenths of the activities of life, with the universal striving after individual economic security as the prime object of endeavour, with the social approbation of money as the measure of constructive success, and with the social appeal to the hoarding instinct as the foundation of the necessary provision for the family and for the future." (**EP**: 268-9)

The ultimate problem with Capitalism, however, was when it became ineffective as a means to the end which justified it: 'Capitalism ... is not intelligent, it is not just, it is not virtuous – and it doesn't deliver the goods' (**CWXXI**: 239). Nevertheless, despite these criticisms of capitalism, Keynes was anxious not to throw out the baby with the bathwater:

> "It is common to hear people say that the epoch of enormous economic progress which characterised the nineteenth century is over ... I believe this is a wildly mistaken interpretation of what is happening to us. We are suffering, not from the rheumatics of old age, but from the growing pains of over-rapid changes, from the painfulness of readjustment between one economic period and another." (**EP**: 321)

The fundamental, underlying problem in this period is that production is not directly production for the sake of consumption, as it is in the two great eras of scarcity and abundance, instead we have production for the sake of profit, of accumulation, for the sake, that is, of production itself. A comparison of the category 'consumption' as it appears in 'Economic Possibilities for our Grandchildren' (Keynes, 1930c) with that in **GT** shows it in two diametrically opposed rôles. Consumption today is consumption for production: it does not matter what it is consumption *of* so long as it contributes to aggregate demand and hence keeps the accumulation of capital going. In the future, in the 'economic paradise', production is a mere means, and consumption the end: in that context consumption means 'learning to live wisely and agreeably and well' (**EP**: 328), solving 'the problems of life and human relations, of creation and behaviour and religion'

(**EP**: xviii). The critical importance of this view of consumption, and its methodological implications, will be taken up in the next section.

The problems of this period of 'capitalistic individualism' are for Keynes precisely those arising from its transitional nature. The *MEC* is falling precisely because it has fulfilled its purpose. Its purpose was to promote the accumulation of capital and, in general, the wealth of society: the falling *MEC* (and marginal propensity to consume, *MPC*) are the inevitable result of that accumulation. Indeed, for Keynes, the definition of the 'economic paradise' is that the *MEC* has fallen to zero. There is nothing pathological about this – on the contrary, it is to be expected and desired.

The trouble arises from the institutional context within which the transition was taking place, namely that of *laissez-faire*. Under *laissez-faire*, Keynes believed, and believed he had demonstrated, the *MEC* falls faster, and further, than the rate of interest (*i*). This is due to a peculiarity of money that it can act as a store of value for the individual but not for the community – what is true for each individual taken separately is not true for all the individuals taken together. '[T]here is no such thing as liquidity of investment for the community as a whole' (**GT**: 155). If the community tries to convert part of its aggregate income into a hoard of money, total income simply declines to the point where the community no longer tries to do so.

Keynes's argument is that the natural tendency is for the *MEC* to decline with increasing abundance of capital and *MPC* to decline with increasing income. *i* should therefore decline *pari passu*: the opportunity cost of investment – that is, the foregone or postponed consumption – should fall to zero, since that portion of income is saved anyway. Given an adequate institutional framework this is what will happen. The *MEC* can then decline to zero without falling below *i* and hence without investment being brought to a standstill. Once the *MEC* has fallen to zero, capital goods are essentially free and we have entered the economic paradise.

However, the institutional framework is *not* adequate: the *laissez-faire* system introduces an intolerable level of uncertainty. If every agent were in some way linked up to every other so that they could act in concert, each would realise that it is in the interest of all to make sure that their saving and investment correspond. No-one could have any interest

in a beggar-thy-neighbour policy of hoarding money. But *laissez-faire* means, precisely, that this coordination is lacking. Every agent must now be in ignorance as to what his fellows are going to do. Instead of assessing real economic conditions each agent must now devote himself to guessing what all the *other* agents think of those conditions, or, rather, to guessing what each other agent guesses every other agent guesses... A rational saver may know that it would be best for all if he (and everyone else) were to restrain himself from hoarding money; he may even assume that everyone else knows this in theory, but he cannot be certain that everyone will have the necessary restraint not to save money 'just in case'. Any such suspicion means that he would be well advised to increase the liquidity of his assets a little. But if he, as a rational agent, finds that necessary, then so presumably do other agents. Every increase in the demand for money (or liquidity preference, as Keynes calls it) is a reduction in aggregate demand (*AD*). A reduction in *AD* means a fall in the *MEC*. The agent must now believe, correctly, that a severe economic recession is on the way, and would be foolish not to build up as large a pool of liquid wealth as possible, thereby driving up *i* yet further. Even if the agent is fully conscious that he is contributing to the crisis, exacerbating it, there is absolutely nothing that he, as an isolated individual, can do about it[134].

> "Many of the greatest economic evils of our time are the fruits of risk, uncertainty, and ignorance ... these ... factors are ... the cause of un-employment ... Yet the cure lies outside the operations of individuals; *it may even be to the interest of individuals to aggravate the disease*. I believe that the cure for these things ... would involve Society in exercising directive intelligence ... over ... private business." (**EP**: 291-292; my emphasis)

The essence of the prisoners' dilemma is that the prisoners are compelled to pursue their partially overlapping and partially conflicting interests rationally but without collaboration. It shows how rationality at the individual (micro) level necessarily leads to irrationality at the collective (macro) level under these conditions. The essence of Keynesian opposition to *laissez-faire* is that by artificially dividing economic agents from each other it compels them, in individual self defence, to act in a manner detrimental to

[134] In the context of a discussion of the prisoners' dilemma Douglas Hofstadter has introduced a concept of 'reverberant doubt' which describes exactly what Keynes is concerned with here: 'Isn't this an amazing and disturbing slide from certain restraint ... It is a cascade, a stampede, in which the tiniest flicker of doubt has become amplified into the gravest avalanche of doubt. That is what I mean by 'reverberant' doubt' and one of the annoying things about it is that the brighter you are, the more quickly and clearly you see what there is to fear' (Hofstadter, 1985: 753).

themselves as a group. Keynesian agents thus find themselves in what we in retrospect can see was a prisoners' dilemma. Not, indeed, a one-shot game, but an indefinitely repeated one. As we have seen, players in an indefinitely repeated game may under certain circumstances – a sufficiently large probability of further rounds of the game together with a sufficiently low rate of discount of future payoffs – find their way to a cooperative outcome. However, as we have also seen, while this is the case for two-player games, the achievement of such desirable outcomes rapidly becomes extremely difficult as the number of players rises above two. With any significant number of players, it becomes impossible to discriminate between cooperators and defectors, leading to the collapse of reciprocity: defection is once more the dominant strategy. And in the Keynesian case we have a multi-player game with the number of players being the number of wealth owners who need to determine the proportions of money and other assets to hold in their portfolios.

Keynes says something remarkable, in this connection, in his 'Notes on the Trade Cycle' (**GT**: Ch 22: 313-332). The crisis, he says, is due to an 'error of pessimism' in which

> "the investments, which would in fact yield 2 per cent in conditions of full employment, are expected to yield less than nothing; and the resulting collapse of new investment then leads to a state of unemployment in which the investments, which would have yielded 2 per cent in conditions of full employment, in fact yield less than nothing." (**GT**: 322)

Clearly this is not an 'error' on the part of the *individual* investors: they expected yields to fall by more than two percentage points and that is exactly what happened; their expectations were quite rational. The 'error' is on the part of the investors *as a whole*: it was sheer insanity for them to be pessimistic as it was precisely that pessimism which led to the collapse in new investments, the consequent unemployment and hence the collapse in yields. The institutional framework of *laissez-faire* dictates individual decision-making on an issue which is fundamentally not an individual matter.

Laissez-faire divides economic agents from each other and leads to uncertainty; uncertainty leads to increased liquidity preference; raised liquidity preference leads to recession; and recession leads to unemployment: 'A monetary economy ... is essentially one in which changing views about the future are capable of influencing the quantity of employment' (**GT**: xxii). 'Unemployment develops ... because people want ... money' (**GT**: 235). Because of this irreducible uncertainty associated with the *laissez-faire*

system, Keynes believed that capitalism would settle down to a normal condition of under-employment equilibrium: 'Unemployment ... apart from brief intervals of excitement is associated – and in my opinion, inevitably associated – with present day capitalistic individualism' (**GT**: 381). 'We oscillate ... round an intermediate position appreciably below full employment' (**GT**: 254).

This outcome has two particularly deleterious consequences, other than the obvious one that unemployment and a fall in aggregate income is in no-one's interest. Firstly, the regular fall in the *MEC* towards zero, and what that is an index of, namely, the accumulation of capital up to the desired level of intensity, is broken off. For as long as *MEC* is below *i*, this process cannot continue. Hence our entry into the 'New Jerusalem', as Lambert (1963: 358) puts it, is postponed for as long as we remain in this rut of under-employment.

Secondly, and this is critical for Keynes, unemployment may lead to damaging, revolutionary changes, either in the direction of fascism or of communism: 'it is certain that the world will not much longer tolerate [this] unemployment' (**GT**: 381). 'If [income deflation] occurs, our present regime of capitalistic individualism will assuredly be replaced by a far-reaching socialism' (**TM**: 346). On another occasion, he took comfort from a general willingness to drop the philosophy of *laissez-faire* for similar reasons to his own – fear that the existing institutions would otherwise be jeopardised. He could discern, he claimed, 'a general conviction that the stability of our institutions absolutely requires a resolute attempt to apply what perhaps we know to preventing the recurrence of another steep descent' (**Times**: 65).

Keynes thus wants reform in order to forestall revolution. Things must change so that things may remain the same. The next section examines in more detail Keynes's methodological holism which laid the basis for his policy prescription, and subsequent sections examine what Keynes believed had to change, what he wanted to retain, and how it should be done.

6.3 Keynes and holism

In the previous section, I argued that, for Keynes, the underlying problem with capitalism was that production was not for the sake of consumption, but for the sake of production itself. To elucidate the relation between production and consumption in Keynes, we need to consider a number of passages from the *General Theory* and early drafts. In a draft chapter of the *General Theory* (**Sup**), Keynes adopted Marx's formulae for simple commodity circulation and capitalist circulation, $C - M - C'$ and $M - C - M'$. The first says that a commodity, C, is exchanged for money, M, and the latter used to purchase another commodity, C'. The difference between C and C' is qualitative: they are different commodities. The second says that a quantity of money, M is invested in commodities, C, and the latter sold for a quantity of money, M', greater than the original quantity ($M' = M + \Delta M$, $\Delta M > 0$)[135]. The mistake of the classical economists, Keynes says, was to assume that money has the role only of means of exchange, as it does in simple commodity circulation, rather than store of value, as in the circulation of capital. In simple commodity production, production is still for consumption: the original commodity is produced in order to sell it and with the proceeds purchase the commodity desired for consumption. In capitalist production, the purpose of production is to augment the value of the capitalist's wealth, and consumption is reduced to a means to this end. In the one case, money is a convenience allowing the commodity owner to translate his commodity, produced only for the market, into the one he wants to consume. In the other, money is money capital, money is the goal and criterion of production.

> "Karl Marx … pointed out that the nature of production in the actual world is not, as economists seem often to suppose, a case of C – M – C', ie of exchanging commodity for money in order to obtain another commodity. That may be the standpoint of the private consumer. But it is not the attitude of *business*, which is a case of M – C – M', ie of parting with money for commodity in order to obtain more money." (**Sup**: 81)

While the formula for the circulation of capital expresses the 'standpoint of business', and the structure of incentives under capitalism, in Keynes's view this involves contradictions: although we might behave as though production were carried out for its

[135] In Marx, the second formula is true of merchant capital, which buys cheap in one market and sells the same commodities dear in another market; in capitalist production proper the original money capital is invested in means of production – constant and variable capital (c, v) – which are then consumed in the process of production, generating new commodities which are subsequently sold for more than the value of the means of production: M – C: MP (c, v) … P … C' – M' (Marx, 1974: 25).

own sake, this cannot literally be true: 'capital is not a self-subsistent entity existing apart from consumption' (**GT**: 106); 'the expectation of consumption is the only *raison d'être* of employment'(**GT**: 211); 'consumption – to repeat the obvious – is the sole end and object of all economic activity' (**GT**: 104). The point Keynes is insisting on here is that production has to be validated by consumption to *count* as production: output must be sold to convert it back into money, and, indeed, more money than was started with. The subordination of consumption to production implicit in classical *laissez-faire* capitalism sets up a continually re-emerging barrier to accumulation in the form of under-consumption and failures of aggregate demand.

Keynes's approach here illustrates the methodological significance of his critique of classical economists, from Ricardo to Pigou. For the individual household, 'the standpoint of the private consumer', we have $C – M – C'$: consumption is the immediate goal of economic activity. So does this mean that $C – M – C'$ is valid for society as a whole? That is what 'economists seem often to suppose'. But that is reductionist: it is saying that what is true of the parts is therefore true of the whole. '[T]he nature of production in the actual world' is the opposite: $M – C – M'$: economic activity is directed towards the accumulation of claims on future production.

This rejection of reductionism is evidenced over and over again in Keynes's writings. In the famous passage from the 'Preface' to the French edition of the *General Theory*, cited in the Introduction to this thesis, and in the Conclusion to the present chapter, Keynes criticises the classical economists for erroneously 'extending to the system as a whole conclusions which have been correctly arrived at in respect of a part taken in isolation' (**GT**: xxxii).

The same line of criticism is apparent in Chapter 2 of the *General Theory*. Here Keynes criticises Ricardo for focusing on microeconomic problems concerning relative prices and the allocation of resources between different uses, and his denial of the desirability, indeed possibility, of macroeconomic analysis of the level of economic activity as a whole. Keynes cites Ricardo's letter to Malthus of 9 October 1820:

> "Political economy you think is an enquiry into the nature and causes of wealth – I think it should be called an enquiry into the laws which determine the division of the produce of industry amongst the classes who concur in its formation. No law can be laid down

> respecting quantity, but a tolerably correct one can be laid down respecting proportions. Every day I am more satisfied that the former enquiry is vain and delusive, and the latter only the true objects of the science." (cited in **GT**: 4)

Later in the same chapter he criticises the classical school for its reductionist approach to the wage bargain. Keynes sets out his famous 'two postulates of classical economics: that 'The wage is equal to the marginal product of labour', and that 'the utility of the wage when a given volume of labour is employed is equal to the marginal disutility of that amount of employment'. The first says that firms are optimising in the labour market, the second that households are. Keynes conceded the first but denied that the second held as a rule. Classical economists who assumed it to be true forgot, he claimed, firstly, that the relationship between real and money wages was different for the individual industry and the whole economy:

> "In the case of a change peculiar to a particular industry one would expect the change in real wages to be in the same direction as the change in money wages. But in the case of changes in the general level of wages … the changes in real wages associated with a change in money wages … is almost always in the opposite direction." (**GT**: 10)

Keynes is clearly reiterating the point that the whole cannot be understood as the sum of its parts: the relationship between real and money wages is transformed as we change levels.

Secondly, according to Keynes – and this is really the same point made another way – the classical economists forgot that the principle, that unemployed workers can always underbid the employed and so bring supply and demand into equilibrium in the labour market, is 'intended … to apply to the whole body of labour and do[es] not merely mean that a single individual can get employment by accepting a cut in money-wages which his fellows refuse' (**GT**: 11). The point is, that if one worker 'considered in isolation' were to accept a cut in wages, this would be relative to a given price level, which would remain unchanged by his actions – and so his, or her, real wage would decline in the same proportion as the money wage. The actions of a single worker, in an economy of any significant size, have a vanishingly small impact on the general price level. For the actions of the workers as a whole this is no longer true: an attempt to reduce the general level of money wages, via their impact on firms' marginal costs, would lead to reductions in the general price level of about the same magnitude, leaving real wages where they were (**GT**: 12). Again, it is clear that Keynes is making a point about the relationship

between phenomena at the system and substrate levels, and criticising the classical economists for failing to see it. At the substrate level, the general price level is a parameter, at the system level it is a variable.

Numerous further examples from Keynes could be cited. At the risk of labouring the point, just two more instances will be considered here, both from the *General Theory*. In Chapter 19, on 'Changes in Money Wages', Keynes once again takes the 'classical economists' to task, in a passage of such clarity as to render exegesis redundant, for impermissibly transferring unexceptionable micro statements to the macro context:

> "In any given industry we have … the demand schedule for labour in the industry relating the quantity of employment to different levels of wages … This conception is then transferred without substantial modification to industry as a whole; and it is supposed by a parity of reasoning, that we have a demand schedule for labour in industry as a whole relating quantities of employment to different levels of wages … [S]urely [this] is fallacious. For the demand schedules for particular industries can only be constructed on some fixed assumption as to the nature of the demand and supply schedules of other industries and as to the amount of the aggregate effective demand. It is invalid, therefore, to transfer the argument to industry as a whole … But if the classical theory is not allowed to extend by analogy its conclusions in respect of a particular industry to industry as a whole, it is wholly unable to answer the question what effect on employment a reduction in money wages will have." (**GT**: 258-260)

Finally, and for exactly the same reasons, we may note that in the chapter of the *General Theory* on 'The Theory of Prices' (**GT**: Chapter 21, 292-309), Keynes rejects the classical dichotomy 'between the theory of value and distribution on the one hand and the theory of money on the other hand' (**GT**: 293).

> "The right dichotomy is, I suggest, between the theory of the individual industry or firm and of the rewards and the distribution between different uses of a *given* quantity of resources on the one hand, and the theory of output and employment *as a whole* on the other hand." (**GT**: 293)

Keynes is again clearly articulating a holist conception here. The classical dichotomy distinguishes between a real supply side and a purely nominal demand side – a standpoint which, as he points out, implies that at the macro level 'the elasticity of supply must have become zero and demand proportional to the quantity of money' (**GT**: 292). This classical standpoint tacitly – and illicitly – assumes that what is '*given*' at the micro level, namely the quantity of resources which is employed in the economy as a whole, must also be given at the macro-level, the level to which monetary theory applies. This leaves money with no real effects (the real and monetary sectors are dichotomous): from our

standpoint as observers it is a mere veil over the real workings of the economy. In opposition to this classical dichotomy, Keynes proposes his own micro-macro dichotomy: a micro sphere of analysis in which conclusions can be 'correctly arrived at in respect of a part … taken in isolation' (**GT**: xxxii), and a macro sphere to be analysed as a whole, as a system, and in which money attains critical importance for real outcomes.

6.4 Keynes's policy prescription

The overview, in the previous two sections, of Keynes's dynamic and historical view of the pathologies of capitalism, and of his clear sighted articulation of the micro and macro levels in economics, lays the basis for an understanding of his policy prescription. Only one further point is required. In utter contrast to Smith's invisible hand of god, and Hayek's evolutionary theory of group selection, Keynes never for a moment assumes that we live in a world endowed with providential, pro-human qualities. His standpoint is entirely consistent with that of A.E. Houseman's 'heartless, witless Nature' (cited, Dawkins, 1995:155). If good is to be found in the world, it must be the results of our own activity. In such a world, a policy of *laissez-faire* is a non-starter.

So what was Keynes's policy prescription – and, moreover, who was to execute it? A careful reading of Keynes makes it quite clear what he was prepared to sacrifice, and what he was determined at all costs to retain – what it was about 'our institutions' and 'the kind of system in which we actually live' (**GT**: 247) which he thought worth keeping. What Keynes was concerned to defend was the liberties, the privileges, the prestige, the security, the standard of living, and in short the whole mode of life of the class of which he was a member: 'If I am going to pursue sectional interests at all, I shall pursue my own. When it comes to the class struggle as such ... the *Class* war will find me on the side of the educated *bourgeoisie*' (**EP**: 297). Far from expressing any narrow, sectarian point of view, however, Keynes was able to take this stance because of the *universality* he ascribed to his class. As we shall see, the 'educated bourgeoisie' was a universal class in the sense that, by following its own interests, it would lead the whole population to the destination of economic 'bliss'.

The 'educated bourgeoisie' comprised for Keynes all those sections of society that his own activities made him part of – business management and public administration, and

the worlds of academia and the arts. It excluded the actual owners of the means of production, the rentier capitalists, and it excluded the 'ordinary' people who 'sell themselves for the means of life' (**EP**: 328). Keynes invented a rather grotesque *raison d'être* for this stratum, which runs as follows. The big problem with the approach of the 'economic paradise' is that *ordinary* people will not know what to do with themselves:

> "I think with dread of the adjustment of the habits and instincts of the ordinary man, bred into him for countless generations, which he may be asked to discard within a few decades ... must we not expect a general 'nervous breakdown'?" (**EP**: 327) "There is no country and no people, I think, who can look forward to the age of leisure and abundance without a dread ... It is a fearful problem for the ordinary person, with no special talents, to occupy himself ..." (**EP**: 328).

Fortunately, however, there are strata of the population who are not 'ordinary', who do have 'special talents'. (These 'talents', however, turn out to be of the monetary variety.)

> "It will be those people, who can keep alive, and cultivate into fuller perfection, the art of life itself, and do not sell themselves for the means of life, who will be able to enjoy the abundance when it comes ... the wealthy classes in any quarter of the world ... are, so to speak, our advance guard – those who are spying out the promised land for the rest of us and pitching their tent there ... those who have an independent income but no associations or duties or ties" (**EP**: 328).

Keynes immediately takes the opportunity of castigating the *idle* rich, the rentiers: 'most of them have failed disastrously ... to solve the problem which has been set them' (**EP**: 328). Keynes is attacking them for failing to live up to the role he ascribes to the rich – the development of a good life of culture and consumption, rather than 'purposiveness' and hoarding – and thereby undermining the *raison d'être* of the class system. Keynes's attitude towards the rentier class is ruthless. If they will not spend their money, take it off them. He regarded the inheritance of fortunes as a specially pernicious, feudal institution, and favoured high death duties to counter its effect on the *MPC* (**EP**: 299; **GT**: 95, 372-3; **Times**: 72). Meanwhile, to the rest of us he addresses an appeal for another chance: 'I feel sure that with a little more experience we shall use the new-found bounty of nature quite differently from the way in which the rich use it today, and will map out for ourselves a plan of life quite otherwise than theirs' (**EP**: 328).

Keynes's belief in the importance of class distinctions comes out clearly when he states his differences from communism :

> "How can I adopt a creed which, preferring the mud to the fish, exalts the boorish proletariat above the *bourgeois* and the intelligentsia who, with whatever faults, are the quality in life and surely carry the seeds of all human advancement ... It is hard for an educated, decent, intelligent son of Western Europe to find his ideals here ... It exalts the common man and makes him everything." (**EP**: 258-259)

We should be quite clear, here, what Keynes means by the *bourgeoisie*. Just as he borrowed the designation 'classical economists' from Marx and then proceeded to use it in a completely different – almost opposite – sense[136], his use of the term bourgeois has little or no overlap with Marx's conception of the capitalist class. The 'educated bourgeoisie', the 'bourgeoisie and intelligentsia', in Keynes do not include the owners of the means of production, the capitalists *per se* (whether holders of debt or equity). When he explicitly sides with the bourgeoisie, by no means is Keynes erecting an apology for the rentier. The latter he regards as a parasitic excrescence on the productive apparatus of society, and one which is in the course of quiet liquidation by the spontaneous development of the economy itself:

> "The rentier aspect of capitalism [is] a transitional phase which will disappear when it has done its work ... the euthanasia of the rentier, of the functionless investor, will be nothing sudden, merely a gradual but prolonged continuance of what we have seen recently in Great Britain, and will need no revolution.' (**GT**: 376)

This particular terminology adopted by Keynes means that when he speaks of 'private initiative and responsibility', 'the traditional advantages of individualism', 'personal liberty', and so on, as desirable attributes of capitalism which will be retained in the new society (**GT**: Ch 24 passim), he is not referring to private property in the means of production. Indeed, the question of private or public ownership of the means of production was a non-issue as far as Keynes was concerned.

> "It is not the ownership of the instruments of production which it is important for the state to assume." (**GT**: 378) "There is no so-called important political question so really

[136] For Marx, 'classical' political economy referred to scientific economics – ie economics which, he felt, tried to explain, rather than to explain away, the nature of capitalistic production – from Petty in the late 17th century on, and culminating in Smith and Ricardo. Subsequent mainstream economists Marx designated 'vulgar', and considered to be only interested in explaining away the undesirable features of capitalism. For Keynes, on the contrary, 'classical' economists are those mainstream economists, since Ricardo, who, like Ricardo, adopt Say's Law: he names JS Mill, Marshall, Edgeworth and Pigou as examples (**GT**: 3). Hence, for Keynes, the labour theory of value of Smith and Ricardo is 'pre-classical' (**GT**: 213).

> unimportant, so irrelevant to the reorganisation of the economic life of Great Britain, as the Nationalisation of the Railways." (**EP**: 290)

Nationalisation was a non-issue for Keynes because the 'educated bourgeoisie' was in fact taking, or had already taken, control of the bulk of industrial – and, indeed, non-industrial – institutions. This theme in Keynes – the separation of ownership and control, leading to the hegemony of the managers in industry and state – has since become a major tradition in its own right. The theme originally had two aspects, distinguishing between holders of debt and equity. In 1923 Keynes described what he calls 'the Investment System, in these terms:

> "Under this phase of capitalism, as developed during the nineteenth century, arrangements were devised for separating the management of property from its ownership ... Contracts to receive fixed sums of money at future dates must have existed as long as money has been lent and borrowed ... But during the nineteenth century they developed a new and increased importance, and had, by the beginning of the twentieth, divided the propertied classes into two groups – the 'business men' and the 'investors' – with partly divergent interests ... business men might be investors also, and investors might hold ordinary shares; but the division was nevertheless real." (**EP**: 61-62)

The second phase, so to speak, occurs when the rentiers, or 'investors', buy up the shares, too, and leave the managers without any ownership stake in the enterprise.

> "A point arrives in the growth of a big institution ... at which the owners of the capital, ie the shareholders, are almost entirely dissociated from the management, with the result that the direct personal interest of the latter in the making of great profit becomes quite secondary.' (**EP**: 289)

Keynes calls this 'the tendency of big business to socialise itself' (**EP**: 289), and describes it as 'a natural line of evolution. The battle of Socialism against unlimited private profit is being won in detail hour by hour' (**EP**: 290). A significant example cited by Keynes in this context, and – significantly – prior to its nationalisation, is that of the Bank of England: 'there is no class of persons in the kingdom of whom the Governor of the Bank of England thinks. less when he decides on his policy than of his shareholders. Their rights, in excess of their conventional dividend have already sunk to the neighbourhood of zero' (**EP**: 290).

This conception of what has since been sensationalised as a 'managerial revolution' is crucial to the understanding of Keynes's policy prescription. Keynes's aim is an adequate policy framework for '[t]he transition from economic anarchy to a regime which

deliberately aims at controlling and directing economic forces' (**EP**: 305). He wants 'a somewhat comprehensive socialisation of investment' (**GT**: 378); he wants 'planning' (**Times**: 72, 77); he wants, as he told Hayek, 'more planning' (**CWXXVII**: 387). He was enthusiastic about the proposals for a national plan contained in the Mosley Manifesto (Keynes, 1930b). 'The central debate in politics, he [sc Keynes] wrote, was between planning and laissez-faire' (Skidelsky, 1975: 241).

But what sort of planning does Keynes want? Planning by whom? For whom? We have already seen his contempt for 'the ordinary man with no special talents', and for 'the mud ... the boorish proletariat' who 'sell themselves for the means of life'. He is even less sympathetic towards those at the opposite pole of the social spectrum. Keynes clearly believed that the 'beastly', 'avaricious' Jews were over-represented, to put it no more strongly than that, among the rentier capitalists. He writes in 'A Short View of Russia' (Keynes, 1925a) that 'the mood of oppression ... in Russia ... is the fruit of some beastliness in the Russian and Jewish natures ...' (**EP**: 270), and in the same article remarks that the Russian Revolution has failed to make the Jews any less avaricious (**EP**: 259). In a highly sinister passage, Keynes daydreams about the fate of Jewish financiers in the economic paradise to come:

> "The love of money as a possession ... will be recognised for what it is, a somewhat disgusting morbidity, one of those semi-criminal, semi-pathological propensities which one hands over with a shudder to the specialists in mental disease ... Perhaps it is not an accident that the race which did most to bring the promise of immortality into the heart and essence of our religions has also done most for the principle of compound interest and particularly loves this most purposive of institutions." (**EP**: 329-330)[137]

Keynes's vision of planning, therefore, is one in which the main role is taken by the 'educated bourgeoisie', excluding the 'common people' on the one side, and the (mainly Jewish) rentiers on the other. While expressing no particular desire to dispense with parliamentary democracy, he clearly regards it as simply irrelevant: 'in the future the Government will have to take on many duties which it has avoided in the past. For these purposes Ministers and Parliament will be unserviceable' (**EP**: 301-302).

[137] It should, I think and hope, be clear that presenting my interpretation of economists' writings implies no endorsement of their stance, but so that there should be no possibility of misinterpreting my motives in reporting Keynes's views here, I completely dissociate myself from his racist remarks and standpoint regarding supposedly 'avaricious' Jews and 'boorish' proletarians alike.

So, in answer to the question of what it is that Keynes wants to preserve in 'existing economic forms', what he means when he refers to the preservation of 'individual initiative' and its 'successful functioning' (**GT**: 380), I have argued that Keynes is anxious, above all, to preserve the status and privileges of his own class, the 'educated bourgeoisie', the 'advanced guard ... spying out the promised land'. It is *their* 'individual initiative' which he wishes to defend. 'Private self interest', expressed through consumer preferences on the market, and the exercise of 'enterprise and skill in the estimation of prospective yields' (**GT**: 221) on the part of the entrepreneur, 'will determine what in particular is produced, in what proportions the factors of production are combined to produce it, and how the value of the final product will be distributed between them' (**GT**: 379). Just as when Smith and Hayek speak of the rights 'the individual' they refer to the individual owner of wealth, the individual as vehicle for capital, and in every other context the individual is the servant of 'society', so for Keynes the 'individual' who is truly efficacious, who enjoys freedom, opportunity and activity, is the 'educated *bourgeois*', the gentleman of independent means and public spirit.

It seems fair to summarise Keynes's vision of planning as a network of 'semi-autonomous bodies' – quangos and quagos linked together and to the national bank by a board of public investment:

> "Progress lies in the growth and the recognition of semi-autonomous bodies within the state ... bodies which in the ordinary course of events are mainly autonomous within their prescribed limitations, but are subject in the last resort to the sovereignty of the democracy expressed through Parliament." (**EP**: 288-289)

Now, even the private enterprise firms of the *laissez-faire* period were 'in the last resort' subject to parliamentary sovereignty. To say this of the 'semi-autonomous bodies' is to say very little. In practice, what we have is a new *laissez-faire*, differing from the old in being collective rather than individualistic. The managerial class, which has quietly triumphed in both the formally private and the formally public sectors, is to be allowed to get on with it, free – in the ordinary course of events – of effective parliamentary supervision, regulation or restraint.

Keynes's articles in *The Times* in January and March, 1937, provide perhaps the most explicit statement of how these 'semi-autonomous bodies' are to be linked, as well as a useful ostensive definition of 'semi-autonomous body':

> 'Now is the time to appoint a board of public investment ... to make sure that detailed plans are prepared. The railway companies, the port and river authorities, the water, gas and electricity undertakings, the building contractors, the local authorities, above all, perhaps, the London County Council and the other great corporations with congested population, should be asked to investigate what projects could be usefully undertaken if capital were available at certain rates of interest – 3½%, 3%, 2½%, 2%. The question of the general advisability of the schemes and their order of preference should be examined next. What is required at once are acts of constructive imagination by our administrators, engineers, and architects, to be followed by financial criticism, sifting and more detailed designing' (**Times**: 72)

This is one half of the strategy. The other half is to ascertain from the mass of information obtained in this, and every other conceivable way, what rate of interest would be compatible with a flow of new projects just sufficient to absorb what the nation chooses to save:

> "The rate of interest must be reduced to the figure that the new projects can afford. In special cases subsidies may be justified. But in general it is the long-term rate of interest which should come down to the figure which the marginal project can earn ... We have the power to achieve this. If we know what rate of interest is required to make profitable a flow of new projects at the proper pace, we have the power to make this rate prevail in the market." (**Times**: 73)

There are three points worth noting here, as to why Keynes is so confident about the rate of interest, when, after all, he had only recently proclaimed himself 'somewhat sceptical of the success of a merely monetary policy directed towards influencing the rate of interest' (**GT**: 164).

The first point is that the Bank of England had already been cited by Keynes as a progressive example of 'semi-autonomous body', and of course the Treasury consists entirely of administrators and economists. The two institutions could therefore be depended upon, once Keynesian ideas had made themselves felt, to take the side of the 'educated bourgeoisie' against the rentiers, on the one side, and interference by the electorate via parliament, on the other.

The second point is that due to the institutionally powerful position in the market of the Treasury and central bank, 'it lies within their power ... to make the long-term rate of

interest what they choose within reason' (**Times**: 73). The channelling of savings through the national bank, too, would give the state additional leverage against the rentier class and undermine the ability of the latter to dictate absolutely the rate of interest on borrowed capital (**Times**: 73, **GT**: 376).

Thirdly, the plan has been drawn up by the business community themselves, and in a collective rather than individualistic way. Everyone knows what the rest of the economy is doing and no-one has any incentive to increase his liquid reserves. Consequently, there is nothing to force *i* up above *MEC* at full-employment. Uncertainty has been eliminated at the outset by removing the artificial isolation of economic agents imposed by the anachronistic *laissez-faire* approach to policy. The payoffs to holding money and bonds have been changed so that the prisoners' dilemma has been removed.

6.5 Did Keynes reject *laissez-faire*?

In conclusion, it is worth addressing the vexed question as to whether Keynes rejected *laissez-faire*. In the past three answers have been given – *yes*, *no*, and *yes and no* – all of them false. The first has tended to be associated with more left-wing interpreters of Keynes. such as Joan Robinson, and the second both with more conservative interpreters, and with left-wing anti-Keynesians, such as Geoffrey Pilling. The third alternative, that Keynes was inconsistent in his attitude to *laissez-faire*, has been a very common one, and in the 1930s cartoons used to appear in the press of Keynes as a double-jointed man supporting, for example, both free-trade and protection.

These approaches fail to do Keynes justice. It is true that Keynes did not make the final break with classical economic theory, until around 1933, and he himself aptly summarised his life to this point as 'a long struggle of escape' (**GT**: xxiii). In spite of this, his general social and political philosophy was consistent throughout his productive life and, I would argue, the changes in his economic theory were designed specifically to supply a theoretical underpinning for his political attitudes: 'The field of social philosophy is the field in which Keynes remained consistent throughout his career' (Lambert, 1963: 344).

> "While the *General Theory* marks a sharp break in economic theory, the 'social philosophy' implications he drew from the work [in Chapter 24] are consistent with his earlier views. In fact the *General Theory* can be viewed as giving an economic theoretic rationalisation for views that Keynes's ethics and intuition had led him to." (Minsky, 1975: 145)

While they can thus agree that Keynes was consistent, commentators are anything but agreed on what it was that Keynes was (supposedly consistently) saying: 'Keynes [is] essentially an economic liberal arguing for specific non-liberal measures solely in periods of unemployment' (Corry, 1978: 26). 'When the whole question of seeing that potential savings are not run to waste in unemployment … is added to the *agenda* [of government], it seem as if there is precious little *non-agenda* left' (Robinson, 1962: 81).

The reason why these views are mistaken is that they take the supposed Keynesian rejection of *laissez-faire* (whether they assert or deny that rejection) to be a rejection *a limine*. It is not. An implication of the present chapter is that it is a *critique* – a concrete negation with a concrete result. Keynes's view of *laissez-faire* is not absolute but conditional and historical. His call for state intervention to equilibrate saving and investment is, in his own view by no means timelessly or universally valid.

The difference between himself on the one hand, and, on the other, the old-fashioned Liberals as well as the classical neoclassical schools whose theories underlay the *laissez-faire* approach was that Keynes 'explained the phenomena, which the old Liberal school attributed to the unchanging and universal character of natural law, in terms of positive and therefore changeable laws and of the particular conditions obtaining at a given time and a given place' (Lambert, 1963: 345).

Lambert here is commenting on Keynes's first book, *Indian Currency and Finance* (1913), written when Keynes was still, in terms of economic theory, entirely within the neoclassical school. The point is that even where, as in his work prior to World War I, Keynes obtained results formally consonant with the neoclassical and liberal traditions, such as the correctness of the *laissez-faire* approach to the economic policy framework in the nineteenth century, these results were obtained on the basis of a different more concrete and more historical methodology. It was this methodology which enabled him to develop a vision of *what* was wrong with *laissez-faire*, when his contemporaries could

only see that *something* was wrong (Pigou, for example, in the 1930s), and hence enabled him to develop a theoretical account of the economic problems of his time.

* * *

> "Whilst, therefore, the enlargement of the functions of government involved in the task of adjusting to one another the propensity to consume and the inducement to invest, would seem to a nineteenth- century publicist or to a contemporary American financier to be a terrific encroachment on individualism, I defend it, on the contrary, both as the only practicable means of avoiding the destruction of existing economic forms in their entirety and as the condition of the successful functioning of individual initiative." (**GT**: 380)

This passage, from the final chapter of the *General Theory*, is a concise statement of Keynes's rejection of *laissez-faire*. Yet taken out of context, it could be extremely misleading. The 'enlargement of the functions of government' does, it is true, include an expansion in the role of the existing state. This is concerned principally with the adjustment of the propensity to consume by manipulation of the rates of income tax and death duty, and by deciding how sharply progressive should be the former, channelling savings through a national savings bank, and a programme of emergency public works in severe recessions.

But this is not the main point for Keynes. His goal is not simply an 'enlargement of the functions' but a change in the *nature* of the state. What he wants is an extra- or non-parliamentary state consisting of a central bank and a national planning board linking together the enterprises (in the broadest sense) of the country into a single organisation. This organisation would, through discussion, draw out a consensus of the whole of the 'educated bourgeoisie'; there would be no call for compulsion. The cancellation of the artificial separation and atomisation of the 'entrepreneurs' (that is, the managerial class), by the principles of *laissez-faire*, would eliminate the uncertainty which gives rise both to damaging fluctuations in economic activity and to the under-employment equilibrium around which the economy oscillates.

This, then, is how a theory can be simultaneously 'revolutionary' (Keynes, 1935) and 'moderately conservative' (**GT**: 377) in its implications. The (revolutionary) introduction of central controls and planning to achieve full employment at the macro level is to provide the necessary environment in which the (conservative) micro-level 'Manchester system' comes into its own' (**GT**: 378-379). Keynes wants to combine

micro-level individualism with the macro-level planning required to preserve it. To put it another way, individual self-seeking behaviour at the micro level will generate desirable social outcomes at the macro level when the institutional framework ensures that the payoffs to individual actions are such as to avoid prisoners dilemmas.

The transition from the Era of Scarcity to the capitalist epoch required central controls on production and distribution to reduce uncertainty and the rate of interest, and raise the marginal efficiency of capital. This was the age of mercantilism and absolutism. Now, in the period of transition from capitalism to the 'economic paradise', similar problems call for similar solutions: a latter-day mercantilist policy (**GT**: Chapter 23), 'promoted by an authority unlikely to be superseded' (**GT**: 203).

Appendix: Bibliographical note

The works of Keynes consulted are as follows. Books – all in the *Collected Writings* series – are listed first, then articles and letters. Each is preceded by the abbreviation used in this thesis, where appropriate. Since all the shorter items are contained in the *Collected Writings,* only the books are included in the bibliography at the end of this thesis. An exception is Keynes's articles of early 1937 in *The Times,* where the versions reprinted in Hutchison (1977) are used here.

a *Books by Keynes*

John Maynard Keynes *The Collected Writings of John Maynard Keynes* (eds: Austin Robinson, Elizabeth Johnson and Donald Moggridge), London: Macmillan, for the Royal Economic Society:

TM Vol VI (1971) [1e: 1930] *A Treatise on Money* Vol II

GT Vol VII (1973a) [1e: 1936] *The General Theory of Employment, Interest and Money*

EP Vol IX (1972a) [1e: 1931] *Essays in Persuasion*

CWXIII Vol XIII (ed Donald Moggridge) (1973b) *The General Theory and After. Part I. Preparation*

CWXX Vol XX (ed Donald Moggridge) (1981) *Activities 1929-31. Rethinking Employment and Unemployment Policies*

CWXXI Vol XXI (ed Donald Moggridge) (1982) *Activities 1931-1939. World Crises and Policies in Britain and America*

CWXXVII Vol XXVII (ed Donald Moggridge) (1980) *Activities 1940-1946. Shaping the Post-War World. Employment and Commodities*

Sup Vol XXIX (ed Donald Moggridge) (1979) *The General Theory And After: A Supplement*

b *Shorter works by Keynes – articles, letters, reviews*

1923 'The Social Consequences of Changes in the Value of Money'. Reprinted in **EP**: 59-75

1926 'The End of *Laissez-faire*'. Reprinted in **EP**: 272-294

1925a 'A Short View of Russia'. Reprinted in **EP**: 253-271

1925b 'Am I a Liberal?'. Reprinted in **EP**: 295-306

1930b 'Sir Oswald Mosely's Manifesto' *Nation and Athenæum*, 13 December; reprinted in **CWXX**: 473-476

1930c 'Economic Possibilities for our Grandchildren'. Reprinted in **EP**: 321-332

1933 'National Self-Sufficiency' *Yale Review*, and *New Statesman and Nation*, 8 and 15 July. Reprinted in **CWXXI**: 233-246

1935 letter to GB Shaw. Reprinted in **CWXIII**: 492-493

Times 1937a 'How to Avoid a Slump' *The Times*, January; reprinted in Hutchison (1977)

Times 1937b 'Borrowing for Defence' *The Times*, March; reprinted in Hutchison (1977)

1944 letter to FA von Hayek reviewing the latter's *Road to Serfdom*. Reprinted in **CWXXVII**: 385-388

Chapter 7 Conclusion

1 Retrospective: Keynes and providentialism

In a holist view of the world, the individual agents composing an economic system are, and are *primarily*, components of a social totality: their life process is determined by their mutual relations, the totality of which is the economic system. Under capitalism, however, the individual agents are divorced from each other and their relations are refracted through their sole link with society: the money nexus. This gives them the appearance of independent, asocial, biological totalities, and hence gives the real social totality the appearance of a mere *congeries*.

It is in a sense immaterial where the economist commences his study of society, whether he 'starts' from the part and deduces therefrom the nature of the whole, or *vice versa*. Friedman, for example, correctly observes that both he and Keynes work 'from the top down', while many monetarists and Keynesians work in the opposite direction (Friedman, 1976: 316). That makes no difference: what matters is not where you 'start' but where you end up: do you understand the economy as a totality – with Keynes, Marx, Hayek, and Smith – or as a congeries – with Friedman, Lucas, and the neoclassical school.

Reductionism is implicit in the 'classical' methodology that Keynes criticised, as well as the methodology of those neoclassical writers, such as Friedman, who re-assert the claims of pre-Keynesian economics post-Keynes. The agent is a rational, utility-maximising being; since society is merely a mass of like individuals, the results of the analysis of his behaviour can be applied directly to society as a whole. Thereby the latter is shown to be a rational, welfare-maximising aggregate of many individuals. Protracted, general, involuntary unemployment is not possible: no rational individual would under-utilise scarce resources, so humanity in the aggregate must necessarily be just as rational. On the other hand, the *appearance* of unemployment can be explained away as *false* appearance concealing the intrinsic rationality of the system: irrationality on the level of the system cannot be the fault of the system but only of the individuals

comprising it – so apparent unemployment must in fact be voluntary, caused, for example, by wage rigidity or other micro-irrationality.

Keynes, summarising his whole approach in a passage to which I have already drawn attention, goes straight to the heart of this question:

> "I have called my theory a *general* theory. I mean by this that I am chiefly concerned with the economic system as a whole … And I argue that important mistakes have been made by extending to the system as a whole conclusions which have been correctly arrived at in respect of a part taken in isolation." (**GT**: xxxii)

Keynes is saying that the principal *differentia* of his method from that of the 'classical economists' is that the system as a whole cannot be considered as a mere congeries of individuals 'taken in isolation'. This is so because those individuals *are not isolated* from each other: what one does affects others. An individual's decision to save or to invest, for instance, has consequences for other individuals who are not party to the relevant transaction and hence unable to affect its outcome.

In this clash between the private form and public consequences of the decisions to consume, and to save, to hold money and to invest, we see again the combination of independence in form and interdependence in content of those decisions, which lies at the heart of the prisoners dilemma. Keynes sees this clash between private action and public consequence as remediable only by the removal of the anachronistic private form of decision-making. Hence Keynes's opposition to *laissez-faire* and his demands for social control of the propensity to consume, for the 'comprehensive socialisation of investment', and for 'communal saving through the agency of the state' (**GT**: 378, 376).

Keynes's holism lies essentially in this: were the community as a whole, or some state agency representing it, to control saving and investment, there would need never be any discrepancy between the two. The desirability of the marginal unit of investment would be equal to the sacrifice involved in the marginal unit of saving, and with the accumulation of wealth, both would decline to zero.

The problem is the presence of an anachronistic institutional framework – *laissez-faire* – which fragments the decision-making process without mitigating the social consequences of the decisions made. The community can only do two things with its income: consume

it or invest it. The individual acting on the basis of self-interest, however, has third alternative: he can hoard part of his income as money. Indeed, if he foresees any slackening of aggregate demand, he would be unwise not to, even if he realises the damage he will inflict on the economy thereby: 'Every act of saving involves a 'forced' inevitable transfer of wealth to him who saves, though he in his turn may suffer from the saving of others.' (**GT**: 212) Hence hoarding, which is the cause of the economic disease, is the rational response of individuals to the fear of that disease: 'It may even be to the interest of individuals to aggravate the disease.' (**EP**: 318). Though in practice the matter might be highly complex, the solution is in principle simple: that individuals should act no longer as individuals but as a collectivity, in so far as *quantitative* investment decisions are concerned.

* * *

The assumption standing behind pre- and post-Keynesian mainstream economics is that the unintended consequences of individual actions are *essentially* benign. This providential assumption pervades the writings of Smith and Hayek, Friedman and Lucas. Keynes devoted his theoretical life to the demonstration that unintended consequences, just because they *are* unintended, are uncontrolled and liable to be thoroughly malign:

> "The world is *not* so governed from above that private and social interests always coincide … It is *not* a correct deduction form the principles of economics that enlightened self-interest always operates in the public interest." (**EP**: 287-288) "There is no design but our own … the invisible hand is merely our own bleeding feet moving through pain and loss to an uncertain … destination." (**CWXX**: 474)

Keynes's rejection, in these passages, of providentialism and the invisible hand bring us full circle. The episodes in the history of economic thought considered in this thesis have shown that the combination in decision-making of independence in form and interdependence in content is an issue which continually re-emerges in political economy. At every stage there is a clash between the scientific and the vulgar, the desire to understand and explain, on the one hand, and fear of the consequences of doing so, on the other. Providentialism plays a key role here.

A relatively unsophisticated strategy of simply ignoring the disparity between levels has been noted but given little explicit attention: it has been assumed that for present

purposes the reductionist methodology of the monetarist and new classical schools can be dismissed *a limine*. The bulk of the thesis has focused on two sophisticated attempts to underpin a reductionist *laissez-faire* policy prescription with a holistic methodology. Smith and Hayek, though separated by two centuries, have proposed very similar invisible hand mechanisms to mediate between the holistic nature of the world and the reductionist character of their desired policy framework. Consideration of Keynes has shone a light on their attempts: his account gives us an outstanding example of the fate of *laissez-faire* political economy if a holistic approach is not supplemented with the *deus ex machina* of an invisible hand.

The precise content of the two invisible hand mechanisms considered – the will of a deity in Smith and a group-evolutionary process in Hayek – was perhaps of less interest than the sheer fact of their existence. We were obliged to explore these propositions in detail to check their scientific status. Though from a systems perspective a default injunction, always to do nothing, is inherently implausible, there would be no justification for an *a limine* rejection. It might well have turned out that one or other of these proposed mechanisms grasped some unexpected aspect of the world. On the basis of that examination we may now see that that was not the case. Both turned out to be essentially ideological constructs, providential assertions which assumed what was to be demonstrated, namely the desirability of spontaneous outcomes.

Similarly, the precise content of Keynes's escape from the twin strategies of reductionism, on the one hand, and holism plus an invisible hand, on the other, is perhaps of less interest than its existence. Keynes had a particular view of the class of which he was part – he saw it as a universal class in a Hegelian sense, leading humanity from darkness into light. He was also, in my reading, a virulent racist with very strong, deeply ambiguous feelings about Jews.[138] I believe that all of this shaped and coloured his reading of writers such as Ricardo and Marx, his positive analysis, and his policy prescription. So from the perspective of this thesis, the details are less important than the fact that he showed that there was an escape route: the economy is to be studied as a system and not as a congeries, and our default is to act, not to do nothing. Against the

[138] This was mixed up in his mind with sexual questions – passages in his essays on Einstein and Dr Melchior being particularly remarkable expressions of this potent mixture of racial and sexual issues. It would be inappropriate, however, to develop this theme further here.

atomism of the 'classical' economists he argues for a clear, holistic, systems view, and against the providentialism of the invisible hand theorists he simply and clear-sightedly denies that any such providential mechanism exists, and shows in detail the implications, positive and normative, of that denial.

So for Keynes, the invisible hand ensuring that desirable social consequences flow from self-seeking individual behaviour is a myth: but the job it was supposed to do, the reconciliation of partially conflicting and partially overlapping interests, still needs to be done. This reconciliation is to be achieved in Keynes's view by the universal class, the educated bourgeoisie, and, in particular, by the extra-parliamentary state which it will build, based around a board of national planning linking all the enterprises of the country to the central bank. For Keynes it is precisely the educated bourgeoisie which will take the place of the invisible hand.

2 Results and prospects

This thesis has thus used a review of some episodes in the history of economic thought to illuminate the ways writers have viewed the relationship between micro and macro levels in economics. The episodes considered are the prisoners' dilemma in game theory, the libertarian response to Arrow's impossibility theorem, Adam Smith's invisible hand, Friedrich Hayek's evolutionary theory, and the case made for planning by John Maynard Keynes.

A number of themes have emerged. Firstly, we saw that in a world of agents with partially overlapping and partially conflicting interests, prisoners' dilemmas may arise and lead optimising agents spontaneously to suboptimal social outcomes. The appropriate policy prescription, the case for state action or otherwise in the economy, depends on the importance we attach to such suboptimalities. We have seen that the extent to which writers adopt a holistic or reductionistic approach is of great importance in appraising their overall standpoint and the policy prescription they promote. Nevertheless, we have also learnt that there is no strict relationship such that reductionists and only reductionists will support *laissez-faire*, and holists intervention. On the contrary, we have seen two towering examples of holists advocating *laissez-faire*: Smith and Hayek. The argument has been made that *laissez-faire* is essentially a reductionist policy

prescription; that those advocating it face severe difficulties in what is in fact a holistic world; and that invisible hand mechanisms constitute an attempt to reconcile the two. We examined two examples: the invisible hand of an omnipotent, omniscient and beneficent deity in Adam Smith, and the group selection theory of the evolution of institutions in Friedrich Hayek. By contrast, in the case of John Maynard Keynes, we saw that when an explicitly holist standpoint is adopted, along with the explicit rejection of invisible hand mechanisms, then the logical result is a call for far-reaching reform and state activity in the economy.

The research begun here has opened many doors to further investigation. One example is the struggle with providentialism to be observed in the various editions of works by Thomas Malthus and his disciple, Charles Darwin (Poovey, 1998: 283; Darwin, 1928: 19-20, 462-3). Further instances include Walras's struggle with the issue of coordination shown in the evolution and hypostatisation of the auctioneer through the first four editions of his *Elements* (Mirowski, 1989: 252). Lastly, there is the current resurgence of evolutionary ideas in political economy shown by Hodgson (1988, 1993) and Vromen (1995), to give just three examples of a burgeoning literature. These will constitute the starting points of future research in this area. Perhaps the most important underlying theme not so far explicitly addressed, but highly relevant for these proposed future studies as well as this thesis, is the continuing and fundamental importance of the Darwinian evolutionary theory of Richard Dawkins for an understanding of the key issues in political economy.

Glossary

Agent. An agent is an entity which carries out some action of interest, such as surviving, transmitting an impulse, buying a good, processing information or allocating resources between its functions. An agent can in principle exist at any level, micro or macro. A macro-level agent, then, is the aggregate or complex of all the micro-level agents composing it.

Aggregate. See under *Micro and macro*.

Apologetic. Of the nature of a formal defence or vindication. In the chapter on Smith, I identify an apologetic aspect to his work. Essentially, Smith is trying to defend two potentially incompatible things: the existing system of ranks and orders of society in politics, and the 'simple system of natural liberty' in economics.

Borda counts. A system of ascribing numbers to preferences such that in addition to the ranking of alternatives, some information on the intensity of preferences is impounded. On might, for example, give a weight of 3 to first choices, 2 to second, and 1 to one's third choice. Then, instead of majority voting, one would arrive at a social choice by adding up the points scored by each alternative. Such a procedure defeats both the paradox of voting (*qv*) and Arrow's impossibility theorem since we are no longer dealing with a purely ordinal ranking of individual preferences.

Cardinal. 'The primitive numbers *one*, *two*, *three*, etc' (Onions, 1973: 285) – but note that cardinals can take any value, positive, negative, fractional, etc, while ordinals can only have positive integer values. The older, cardinal approach to utility assumed not only that the consumer ranked consumption bundles, but that it was meaningful to speak of the magnitude of the utility yielded by these bundles, in terms of the (cardinal) number of 'utils' associated with each bundle. See also *ordinal*.

Classical. The term 'classical' is a source of much confusion; it is much used in economics and, in particular, in the history of economic thought, and indeed in highly inconsistent senses. It seems to me that there are at least three meanings of the term in

general use. Marx originally divided political economists into 'classical' and 'vulgar' classes, with Smith and Ricardo representing the pinnacle of the scientific or classical group and, roughly, everyone after Ricardo being consigned to the apologetic, or vulgar trend (Marx, 1972: 501). Keynes then borrowed the designation 'classical economists' from Marx – and proceeded to use it in a completely different – almost opposite – sense. For Marx, 'classical' political economy referred to scientific economics – that is, economics which, he felt, tried to explain, rather than to explain away, the nature of capitalistic production – from Petty in the late 17th century on, and culminating in Smith and Ricardo. Subsequent mainstream economists Marx designated 'vulgar', and considered to be only interested in explaining away the undesirable features of capitalism. For Keynes, on the contrary, 'classical' economists are those mainstream economists, since Ricardo, who, like Ricardo, adopt Say's Law: he names JS Mill, Marshall, Edgeworth and Pigou as examples (**GT**: 3). Hence, for Keynes, the labour theory of value of Smith and Ricardo is 'pre-classical' (**GT**: 213), though logically Ricardo and Say should be considered classical in Keynes's schema since they both adopted Say's law. Marx, incidentally, while designating Ricardo a classical economist, identified the adoption of Say's law as a vulgar element in Ricardo, unworthy of the rest of his contribution (Marx, 1968: 468, 502). Finally, in standard History of Economic Thought contexts, 'classical' is generally taken to refer to the principal economists – principally Smith, Ricardo, Malthus, the Mills, and Marx – up to the marginal revolution of the 1870s, and 'neoclassical' to those since.

Compatibilism. The view that causal determinism and moral responsibility are not mutually exclusive.

Connation. Sharing the same route into a subsequent generation. All a biological individual's genes share a route into the next generation via the organism's gametes – they thus share an interest in cooperating to ensure the success of those gametes. Parasite genes do not share a route to subsequent generations and thus are at liberty to damage host interests if it aids their own survival and reproduction. Memes and meme complexes do not share a route into the next generation with their human hosts and hence may have interests quite antagonistic to those hosts.

Congeries. A collection of things merely heaped together.

Consequentialism. The view that what gives our actions their moral character is their consequences. The claim that the deliberate killing of one innocent person in order to save the lives of more than one other person is a morally justified act would be a consequentialist claim. The end can justify the means. The opposite of *deontology* (*qv*).

Coordination. In an ensemble of interacting purposive elements coordination arises when the actions of the individual elements are consistent rather than chaotic and mutually defeating.

Cyclical preferences. Suppose three individuals, *A*, *B*, and *C*, and three possible policies or states of the world, *x*, *y* and *z*. Suppose also that the pattern of preferences is

$$A: x > y > z$$
$$B: z > x > y$$
$$C: y > z > x.$$

Simple majority voting on each pair of alternatives will elicit the social preference that x is preferred to *y*, *y* to *z*, and *z* to *x*: we are left going round in circles and each policy is preferred to both of the others by a 2-to-1 majority:

$$S: x > y > z > x \ldots$$

See also, *transitivity* and *transitive closure*.

Decisiveness. In social choice theory, the preference of a decisive individual or group on some specific alternative about states of the world that we can bring about, is automatically the society's preference, whatever anyone else's preference may be. Decisiveness refers only to one alternative while *dictatorship* (*qv*) refers to all alternatives on which the society must choose.

Deontology. The view that some actions are right and others wrong by virtue of their intrinsic nature, regardless of the consequences of those actions. The end cannot justify

the means. So, to refuse to tell a lie, even to save someone's life, would be a deontological standpoint. The opposite of *consequentialism* (*qv*).

Dictatorship. An individual, or a group of individuals with identical preferences, is a dictator if its preferences over every possible state of the world that the society can bring about are automatically the society's preferences, regardless of the preferences of other members of the society.

Dominant strategy. Player's *strategy* (*qv*) in which the rationally optimal moves are independent of the moves made by the other player(s).

Emergence. Emergence is the idea that features of an entity at a particular level may 'emerge' at that level, that is, are not reducible to features of the underlying substrate of the entity. Emergence is therefore a characteristic of holist modes of thought. See also *holism* and *reductionism*.

Evolutionarily stable strategy (ESS). An ESS is a strategy such that, if all members of a population adopt it, then it is immune from invasion by any mutant strategy (Smith, 1982: 10). ESS is a close biological analogue of the concept of *Nash equilibrium* (*qv*).

Fallacy of composition. The fallacy of composition asserts that what is true of the parts taken in isolation is true of the whole. Thus Plott claims that the concept of social preferences is based on the fallacy of composition since 'the concept of social preference involves an illegitimate transfer of the properties of an individual to the properties of a collection of individuals' (Barry and Hardin, 1982: 242).

Genotype. 'The genetic constitution of an organism at a particular locus or set of loci. Sometimes used more loosely as the whole genetic counterpart to phenotype' (Dawkins, 1989b: 287). See also *phenotype*.

Idealism. A philosophical standpoint in which the phenomenological world which we observe is seen as the product of a reality standing behind it, and the substance of that reality is thought or ideas or abstract forms. So, for example, Plato's Theory of Forms is held to be a species of idealism: we see a variety of forms of appearance, but they are all

products of and reducible to a few abstract geometrical forms. A materialist would say the forms are obtained in our minds by abstraction from the things we see, an idealist that the things we see are obtained by an obscure process of materialisation of the pre-existing forms or ideas. See also *materialism*.

Holism. The view that phenomena at one level can be understood as emergent at that level, that a higher level entity can be understood as a product of the interrelationships between its component parts. The opposite of *reductionism* (*qv*). See also *micro and macro*, and *micro-macro dichotomy*.

Materialism. The philosophical standpoint in which the fundamental nature of the world we see is taken to be matter in motion. Thoughts, ideas, forms, are all held to be reflections of the material world, by a material, physiological process of abstraction, in brains which are themselves just one particular part of that physical world. See also *idealism*.

Meme. The units of selection in biological evolution are genes, the corresponding units of selection in cultural evolution are cultural traits or features with the capacity to be adopted, consciously or unconsciously, by human beings (Dawkins, 1989a: 192, 1989b: 290).

Micro and **macro**. Micro and macro just mean small or lower-level and big or higher-level. They are relative concepts. An atom is a macro level phenomenon as far as electrons, protons, quarks, neutrinos and so on are concerned; it is a micro concept as far as molecules, cells, organisms, etc, are concerned. Similarly, in economics, a market is a macro level concept when discussing the behaviour of agents within it, and a micro concept when looking at the whole economy. I also use the terms substrate and aggregate to refer to micro and macro level phenomena, respectively.

Micro-macro dichotomy. The putative failure of rationality at the micro level, the level of the individual agent, to guarantee rationality at the macro level, the level of the whole society.

Monism. The view that ultimate reality consists of only one kind of stuff. Consistent idealism (*qv*) and materialism (*qv*) are monist, since they hold that reality consists exclusively of ideas, or matter, respectively. The opposite of pluralism, which holds that there are many kinds of stuff in the world. Cartesian dualism, in which there are material bodies but also souls, is a species of pluralism.

Nash equilibrium. An equilibrium in which each agent is doing the best it can, *given* what all the other agents are doing. Closely related to the concept of an *evolutionarily stable strategy* (*qv*).

Natural law. The conception that there are certain divinely appointed principles of human conduct, awaiting discovery by human reason, with which human law must conform if it is to be valid (Hart, 1961: 182, 152).

Neoclassical. Mainstream economists since the marginal revolution of the 1870s.

Ontogeny. The process of development of the individual organism from foetus to sexually mature adult. See also *phylogeny*.

Ordinal. 'Marking position in an order or series, as *first*, *second*, *third*, etc' (Onions, 1973: 1460). Opposite of *cardinal* (*qv*). In the ordinal approach to utility, we only assume information on whether a bundle of commodities yields more, less, or the same utility to the consumer than (as) an alternative bundle.

Paradox. A statement or condition that in some sense seems self-defeating. In the strong or logical sense, a paradox is the bringing together of two (not necessarily distinct) incompatible but apparently irrefutable statements. 'This statement is false' is a version of the Epimenides or Liar Paradox, used in the Gödel indecidability theorem. In the weak sense, a *paradox* is a statement merely contrary to *orthodox* belief. The parable of the Good Samaritan is paradoxical, not because it was logically incoherent for Jesus to posit such an entity, but because he knew his audience would have strong negative defaults about people from Samaria. Keynes's paradox of thrift is in this latter category, as is the paradox of voting. In the latter, individuals with cyclical preferences (*qv*) between a number of alternative actions are unable to reach a coherent decision by

simple majority voting on each pair of alternatives. This result is unexpected, rather than logically self-contradictory.

Pareto efficiency. An outcome is Pareto-efficient if it is impossible to make an agent better-off without making any other agent worse-off. A change in behaviour causing a change in outcome which improves the welfare of at least one agent without worsening that of any other agent is a Pareto improvement. Pareto efficiency is a minimal requirement of social welfare that most observers can agree on – although a normative rather than a positive statement about the world, it is a relatively robust one. It might not be possible to agree what would constitute maximising social welfare, but still be possible to agree that Pareto-inefficient outcomes show that social welfare is not maximised.

Peripatetic. Aristotelian.

Phenotype. 'The manifest attributes of an organism, the joint product of its genes and their environment during ontogeny' (Dawkins 1989b: 292). See also *genotype*.

Phylogeny. 'Ancestral history on the evolutionary timescale' (Dawkins, 1989b: 292) – the sequence of forms taken by the species over long periods of time. See also *ontogeny*.

Providentialism. Belief in 'the foreknowing and beneficent care and government of God (or of nature, etc); divine direction, control or guidance' (Onions, 1973: 1696).

Quago. Quasi-autonomous governmental organisation. See also *quango*.

Quango. Quasi-autonomous non-governmental organisation. See also *quago*.

Rationality. Given an objective function to maximise, rational behaviour comprises just those actions which do in fact systematically maximise that function, within the constraints the agent is subject to. For example, an agent who fails to derive the maximum satisfaction from consumption because of systematic errors in forecasting the satisfaction to be gained from a particular class of commodities is acting irrationally. One mode of procedure would be to regard the problem – the agent's habit of

underestimating the enjoyability of a product, for example – as a constraint. The agent can then be defined to be acting rationally, *given* the constraint. This would be to define the problem away, and hence is a fundamentally uninteresting approach.

Reciprocity. The ability of players in a game to influence the behaviour of other players by the moves they themselves make.

Reductionism. The view that an entity at one level can be understood as an aggregate of entities at a lower, substrate level, that the properties and behaviour of higher level entities can be understood in terms of the properties and behaviour of its constituent lower level parts, taken in isolation, taken, that is, as a *congeries* (*qv*). The opposite of *holism* (*qv*). See also *micro and macro*, and *micro-macro dichotomy*.

Say's Law. 'Say's Law' is the view that 'supply creates its own demand' in the form of the revenues to the factors participating in production, hence we can only have offsetting over- and under-production in different sectors, but not generalised over-production. Suppose national income is Y. Now new output of ΔY is produced. All the factors combining to produce ΔY receive some compensation for their contribution. With this new revenue they demand additional products – perhaps some of the new product, perhaps some of the previously produced output. Now it may be that the demand for ΔY is greater than or less than this value, but that is an allocative matter: it just means that resources should be redirected towards this product and away from others, or *vice versa*. It has no meaning in this view to say that total demand is too high or too low.

Social welfare function (SWF). A measure of social happiness; a putative algorithm for reconciling and aggregating the preferences of members of society so that states of the world that we can bring about by our actions may be unambiguously ranked in terms of better and worse.

Solipsism. The view that only oneself exists.

Strategy. List of (or algorithm generating) all the moves that a player will make in all the different circumstances that can arise in the course of a game.

Substrate. See under *micro and macro*.

Teleology. 'From the Greek word for goal, task, completion, or perfection. Teleological explanations attempt to account for things and features by appeal to their contribution to optimal states, or the normal functioning, or the attainment of goals, of wholes or systems they belong to' (James Bogen in Honderich, 1995: 868).

Theodicy. A vindication of God's divine character in the face of the challenge of the existence of evil. *The Theodicy* of 1710 was the only book on philosophy Leibniz published in his lifetime. He argued that we necessarily live in the best of all possible worlds, since it was the one world, of all logically possible ones, which a necessarily morally perfect god had chosen to actualise. Hence any apparent imperfections in the world must be logically necessary ones.

Transitive closure. Cyclical preferences (*qv*) generate the preference ranking

$$S\text{: } x > y > z > x \ldots$$

This violates transitivity (*qv*): x is preferred to y and y to z, but z is preferred to x. Hence the ranking is not an ordering. 'Taking the transitive closure' just means replacing such a preference ranking with the statement that

$$S\text{: } x = y = z = x$$

Which retains the salient fact that society is indifferent between the three alternatives, but restores transitivity.

Transitivity. Some relationships are transitive, some aren't. 'Older than' is transitive: that Jane is older than Anne, and Anne is older than Jill, implies that Jane is older than Jill. 'Is the mother of' is intransitive: that Jane is the mother of Anne, and Anne is the mother of Jill, does not imply (indeed, in this case it precludes), that Jane is the mother of Gill.

Utilitarianism. A family of consequentialist views in which it is held that aggregate social welfare is an operational concept, a measurable entity which it is our duty and interest to maximise. The utilities of individual members of society feed into that aggregate, but it may be possible to trade off the utility of one member against that of another. A view fiercely opposed by the deontological standpoint, which argues that utilitarianism is an ideological cover for unjustified and immoral incursions into the freedom of members of society. See also *social welfare function*, *consequentialism* and *deontology*.

References and bibliography

A Anikin (1975) *A Science in its Youth (Pre-Marxian Political Economy)* Moscow: Progress Publishers (trans KM Cook)

Kenneth J Arrow (1951, 2e: 1963) *Social Choice and Individual Values* New Haven, Conn: Yale University Press

Robert Axelrod (1984; Penguin edition with Foreword by Richard Dawkins 1990) *The Evolution of Co-operation* New York: Basic Books

John D Barrow and Frank J Tipler (1986, pbk ed with corrections 1988) *The Anthropic Cosmological Principle* Oxford: Oxford University Press

Brian Barry and Russell Hardin (eds) (1982) *Rational Man and Irrational Society? An introduction and sourcebook* (London: Sage)

Norman P Barry (1988) *The Invisible Hand in Economics and Politics. A Study in the Two Conflicting Explanations of Society: End-States and Processes* London: Institute of Economic Affairs (Series title: *Hobart Papers* ed Arthur Seldon, No 111)

Carl L Becker (1932) *The Heavenly City of the Eighteenth-Century Philosophers* London: Yale University Press

David Begg, Stanley Fischer and Rudiger Dornbusch (3e: 1991) *Economics* London: McGraw-Hill

Ken Binmore (1994) *Playing Fair. Game Theory and the Social Contract* Vol I London: MIT Press

Susan Blackmore (1999) *The Meme Machine* Oxford: Oxford University Press

Mark Blaug (3e: 1978) *Economic Theory in Retrospect* Cambridge: Cambridge University Press

Mario Bunge (2000) 'Systemism: the alternative to individualism and holism' *The Journal of Socio-Economics* 29: 147-157

Eamonn Butler 'Foreword', in **KES**: 7-8

RH Campbell and AS Skinner (1976) 'General Introduction' in Adam Smith (1976-1980) Volume II *An Inquiry into the Nature and Causes of the Wealth of Nations* (in two volumes), ed RH Campbell, AS Skinner and WB Todd

Anna Carabelli (1988) *On Keynes' Method* London: Macmillan

Nancy Cartwright (1995) '*Ceteris Paribus* Laws and Socio-Economic Machines' *The Monist* 78, 3, July, 276-294

PH Clarke (1996, 1998) 'Adam Smith and the Stoics: The Influence of Marcus Aurelius', University of the West of England, Faculty of Economics and Social Science, Working Papers in Economics No 18, April 1996; revised version January 1998, mimeo, School of Economics and International Studies, University of Lincolnshire and Humberside

B Corry (1978) 'Keynes in the history of economic thought'. In AP Thirlwall (ed) (1978) *Keynes and Laissez-Faire* London: Macmillan

Rod Cross and Douglas Strachan (1997) 'On George Soros and Economic Analysis' *Kyklos* 50, 4, 561-574

Charles Darwin (1928) [1e: 1859] *The Origin of Species* [*by means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*] London: J.M. Dent & Sons, Everyman's Library

Richard Dawkins (2e: 1989a) [1e: 1976] *The Selfish Gene* Oxford: Oxford University Press

Richard Dawkins (1989b) *The Extended Phenotype: the Long Reach of the Gene* Oxford: Oxford University Press. Second edition of Richard Dawkins (1982) *The Extended Phenotype: the Gene as the Unit of Selection* London: Freeman

Richard Dawkins (1995) *River out of Eden – A Darwinian View of Life* London: Phoenix/Orion

Andy Denis (1996a) 'Collective and Individual Rationality in Economics: the Prisoners' Dilemma', City University, London, Department of Economics and Applied Econometrics Research Unit, Discussion Paper Series, 53

Andy Denis (1996b) 'Collective and Individual Rationality in Economics: Arrow's Impossibility Theorem', City University, London, Department of Economics and Applied Econometrics Research Unit, Discussion Paper Series, 54

Andy Denis (1997) 'Collective and Individual Rationality in Economics: The Invisible Hand of God in Adam Smith', City University, London, Department of Economics and Applied Econometrics Research Unit, Discussion Paper Series, 62

Andy Denis (1999a) 'Was Adam Smith an Individualist?' *History of the Human Sciences* 12, 3, August, 71-86

Andy Denis (1999b) 'Friedrich Hayek: a Panglossian evolutionary theorist' Interdisciplinary Seminar Discussion Paper, City University, London, School of Social and Human Sciences

Andy Denis (2000) 'Epistemology, observed particulars and providentialist assumptions: the *Fact* in the history of political economy' *Studies in the History and Philosophy of Science* 31, 2, 353-361

Andy Denis (2001a) 'Methodology and policy prescription in economic thought: a response to Mario Bunge'. Submitted to *The Journal of Socio-Economics*.

Andy Denis (2001b) 'Was Hayek a Panglossian evolutionary theorist? A reply to Whitman'. Submitted to *Constitutional Political Economy*.

Daniel C Dennett (1993) *Consciousness Explained* London: Penguin Books

Daniel C Dennett (1995) *Darwin's Dangerous Idea. Evolution and the Meanings of Life* London: Penguin Books

Arthur De Vany (1996) 'Information, chance, and evolution: Alchian and the economics of self-organisation' *Economic Inquiry* XXXIV, 3, June, 427-443

Mohammed Dore (1997) 'On Playing Fair: Professor Binmore on Game Theory and the Social Contract' *Theory and Decision* 43, 3, November, 219-239

John Eatwell, Murray Milgate and Peter Newman (eds) (1989) *'The New Palgrave' The Invisible Hand* London: Macmillan. Essays on 'The Invisible Hand' from John Eatwell, Murray Milgate and Peter Newman (eds) (1987) *The New Palgrave: A Dictionary of Economics* London: Macmillan

Antony GN Flew (1967) *Evolutionary Ethics* London: Macmillan (Series title: *New Studies in Ethics* ed WD Hudson)

Milton Friedman (1962) *Capitalism and Freedom* Chicago: Chicago University Press

Milton Friedman (1976) 'Comment' in J Stein (ed) (1976) *Monetarism* Oxford: North-Holland Publishing Company

Roger W Garrison and Israel M Kirzner 'Friedrich August von Hayek' in John Eatwell, Murray Milgate and Peter Newman (1989) 119-130

Peter Gay (1964) *The Party of Humanity. Studies in the French Enlightenment* London: Weidenfeld and Nicolson

Peter Gay (1969) *The Enlightenment: An Interpretation* Vol II *The Science of Freedom* London: Weidenfeld and Nicolson

Alexander Gray (1931) *The Development of Economic Doctrine. An Introductory Survey* London: Longmans

Richard L Gregory (ed) (1987) *The Oxford Companion to the Mind* Oxford: Oxford University Press

Frank Hahn (1982) 'Reflections on the Invisible Hand' *Lloyds Bank Review* 144, April, 1-21 (The Fred Hirsch Memorial Lecture, Warwick University, 5 November 1981)

HLA Hart (1961) *The Concept of Law* Oxford: Clarendon Press/OUP

Alan Haworth (1994) *Anti-libertarianism. Markets, philosophy and myth* London: Routledge

For further discussion of works by Friedrich Hayek, please refer to the bibliographical note appended to Chapter 5 of this thesis.

Friedrich Hayek (1944) *The Road to Serfdom* London: George Routledge & Sons

Friedrich Hayek (1948) *Individualism and Economic Order* Chicago: University of Chicago Press; (reprinted, 1949, London: Routledge)

Friedrich Hayek (1952) *The Sensory Order. An Inquiry into the Foundations of Theoretical Psychology* London: Routledge & Kegan Paul

Friedrich Hayek (1e: 1952) *The Counter-Revolution of Science. Studies on the Abuse of Reason* Glencoe, Illinois (second edition, 1979, Indianapolis: Liberty*Press*)

Friedrich Hayek (1960) *The Constitution of Liberty* London: Routledge & Kegan Paul

Friedrich Hayek (1967) *Studies in Philosophy, Politics and Economics* London: Routledge & Kegan Paul

Friedrich Hayek (1978) *New Studies in Philosophy, Politics, Economics and the History of Ideas* London: Routledge & Kegan Paul

Friedrich Hayek (2e: 1978) (ed SR Shenoy) *A Tiger by the Tail* London: Institute for Economic Affairs

Friedrich Hayek (1982) *Law, Legislation and Liberty. A new statement of the liberal principles of justice and political economy* London: Routledge:

Volume 1 (1973) *Rules and Order*

Volume 2 (1976) *The Mirage of Social Justice*

Volume 3 (1979) *The Political Order of a Free People*.

Friedrich Hayek (1983) *Knowledge, Evolution and Society* London: Adam Smith Institute

Georg Wilhelm Friedrich Hegel (1952/1820) *Hegel's Philosophy of Right* (ed, trans TM Knox) Oxford: Clarendon Press/OUP

Georg Wilhelm Friedrich Hegel (1975/1830) *Logic* (Part One of the *Encyclopaedia of the Philosophical Sciences*) (trans William Wallace, forward by FN Findlay) Oxford: Clarendon Press/OUP

Robert L Heilbroner (ed) (with the assistance of Laurence J Malone) (1986) *The Essential Adam Smith* London: Norton

Geoffrey M. Hodgson (1988) *Economics and Institutions. A Manifesto for a Modern Institutional Economics* Cambridge: Polity Press

Geoffrey M. Hodgson (1993) *Economics and Evolution. Bringing Life Back into Economics* Cambridge: Polity Press

Douglas R Hofstadter (1981) 'Reflections [on Searle, 1980]' in Hofstadter and Dennett (1982) 373-382

Douglas R Hofstadter (1985) *Metamagical Themas* Basic Books

Douglas R Hofstadter and Daniel C Dennett (eds) (1981, Penguin ed 1982) *The Mind's I. Fantasies and Reflections on Self and Soul* London: Penguin

Ted Honderich (ed) (1995) *The Oxford Companion to Philosophy* Oxford: Oxford University Press

Kevin D Hoover (1995) 'Is Macroeconomics for Real?' *The Monist* 78, 3, July, 235-257

TW Hutchison (1977) *Keynes vs the Keynesians ...?* London: Institute of Economic Affairs

TW Hutchison (1981) *The Politics and Philosophy of Economics. Marxians, Keynesians and Austrians* Oxford: Basil Blackwell

Immanuel Kant (1934) (trans JMD Meikljohn) [1st German ed: 1781, 2e: 1787] *Critique of Pure Reason* London: Dent Everyman's Library

Immanuel Kant (1950) (trans Lewis White Beck) [1st German ed: 1783] *Prolegomena to Any Future Metaphysics Which Will Be Able to Come Forth as Science* New York: Bobbs-Merrill

John Kemp and Bruce Philp (1996) *Epitaph for a Legless Centipede? A Paradox of Backward Induction* Manchester Metropolitan University Discussion Papers in Economics and Economic History, Series No: 96-03, January 1996

For further discussion of works by John Maynard Keynes, please refer to the bibliographical note appended to Chapter 6 of this thesis.

John Maynard Keynes *The Collected Writings of John Maynard Keynes* eds: Richard Kahn, Roy Harrod, Austin Robinson et al, London: Macmillan/Royal Economic Society:

Volume VI (1971) [1e: 1930] *A Treatise on Money* Vol II

Volume VII (1973) [1e: 1936] *The General Theory of Employment, Interest and Money*

Volume IX (1972) [1e: 1931] *Essays in Persuasion*

Volume XIII (ed Donald Moggridge) (1973) *The General Theory and After. Part I. Preparation*

Volume XX (ed Donald Moggridge) (1981) *Activities 1929-31. Rethinking Employment and Unemployment Policies*

Volume XXI (ed Donald Moggridge) (1982) *Activities 1931-1939. World Crises and Policies in Britain and America*

Volume XXVII (ed Donald Moggridge) (1980) *Activities 1940-1946. Shaping the Post-War World. Employment and Commodities*

Volume XXIX (ed Donald Moggridge) (1979) *The General Theory And After: A Supplement*

P Lambert (1963) 'The social philosophy of JM Keynes', in John Cunningham Wood (ed) (1983) *John Maynard Keynes: Critical Assessments* London: Routledge

Harry Landreth and David C. Colander (3e: 1994) *History of Economic Thought* Boston: Houghton Mifflin

Titus Lucretius Carus (55 BC) *De Rerum Natura*. (i) trans, ed RE Latham (1951) Harmondsworth: Penguin, Series title: Penguin Classics, ed Betty Radice; (ii) trans, ed Martin Ferguson Smith (1969) London: Sphere

Robert E Lucas, Jr (1987) *Models of Business Cycles* Oxford: Blackwell

AL Macfie (1959) 'Adam Smith's *Moral Sentiments* as Foundation for his *Wealth of Nations*' *Oxford Economic Papers* 2 (new series), October, 209-228

AL Macfie (1961) 'Adam Smith's *Theory of Moral Sentiments*' *Scottish Journal of Political Economy* 8: 12-27

AL Macfie (1967) *The Individual in Society. Papers on Adam Smith* London: George Allen & Unwin (Series title: *University of Glasgow Social and Economic Studies* (New Series) ed Professor D.J. Robertson, Vol 11)

Alec Macfie (1971) 'The Invisible Hand of Jupiter' *Journal of the History of Ideas* 32: 595-99

Edward McCurdy (ed, trans) (1938) *The Notebooks of Leonardo da Vinci* London: Jonathan Cape

Fritz Machlup (ed) (1977a) *Essays on Hayek* London: Routledge & Kegan Paul

Fritz Machlup (1977b) 'Hayek's Contribution to Economics' in Machlup (1977a) 13-59

Bryan Magee (1973) *Popper* London: Collins (series title: *Fontana Modern Masters* ed Frank Kermode)

Arthur Marwick (1970) *The Nature of History* London: Macmillan

Karl Marx (1843) 'Critique of Hegel's Doctrine of the State' (trans Rodney Livingstone) in Marx (1975: 57-198)

Karl Marx (1968) *Theories of Surplus Value* Volume II, Moscow: Progress Publishers

Karl Marx (1972) *Theories of Surplus Value* Volume III, Moscow: Progress Publishers

Karl Marx (1974) *Capital. A Critique of Political Economy* Volume II *The Process of Circulation of Capital* London: Lawrence & Wishart

Karl Marx (1975) *Early Writings* Harmondsworth: Penguin/London: *New Left Review* (Series title: *The Pelican Marx Library* General Editor: Quintin Hoare)

Karl Marx and Friedrich Engels (1848) *Manifesto of the Communist Party*. In Karl Marx and Friedrich Engels (1976) pp 477-519

Karl Marx and Friedrich Engels (1976) *Collected Works* Vol 6 *Marx and Engels: 1845-48* London: Lawrence & Wishart (for the Institute of Marxism–Leninism, Moscow)

Hyman P Minsky (1975) *John Maynard Keynes* London: Macmillan

Marvin Minsky (1985) *The Society of Mind* London: Picador/Heinemann

Philip Mirowski (1989) *More Heat than Light. Economics as Social Physics: Physics as Nature's Economics* Cambridge: Cambridge University Press (Series title: *Historical Perspectives on Modern Economics* ed Craufurd D Goodwin)

James A Mirrlees (1997) *1996 Nobel Prize Lecture*: 'Information and Incentives: the Economics of Carrots and Sticks' *Economic Journal* 107, 444, September, 1311-1329

Alan Nelson (1984) 'Some issues surrounding the reduction of macroeconomics to microeconomics' *Philosophy of Science* Vol 51: 573-594

Anthony O'Hear (1995) 'Hayek, Friedrich August von' (1899-1992) in Ted Honderich (ed) *The Oxford Companion to Philosophy* Oxford: OUP, 336

CT Onions (ed) (3e: 1973) *The Shorter Oxford English Dictionary on Historical Principles* (2 Vols) Oxford: Clarendon Press

Gordon Pask (1961) *An Approach to Cybernetics* London: Radius Books/Hutchinson

Geoffrey Pilling (1986) *The Crisis of Keynesian Economics*: *A Marxist View* London & Sydney: Croom Helm

Mary Poovey (1998) *A History of the Modern Fact. Problems of Knowledge in the Sciences of Wealth and Society* Chicago: Chicago University Press

Karl R Popper (1957) *The Poverty of Historicism* London: Routledge & Kegan Paul. First published as articles in *Economica* XI (new series), 42 and 43 (1944) and XII, 46 (1945)

William Poundstone (1988) *Labyrinths of Reason. Paradox, puzzles and the frailty of knowledge* London: Penguin

DD Raphael (1985) 'Smith' in Keith Thomas (ed) (1997) *Three Great Economists* Oxford: Oxford University Press (Series title: *Past Masters* ed Keith Thomas)

DD Raphael and AL Macfie (1976) 'Introduction' in Adam Smith (1976-1980) Volume I *The Theory of Moral Sentiments*, ed DD Raphael and AL Macfie

DD Raphael and AS Skinner (1980) 'General Introduction' in Adam Smith (1976-1980) Volume III *Essays on Philosophical Subjects*, ed WPD Wightman and JC Bryce

Matt Ridley (1996) *The Origins of Virtue* London: Viking

Joan Robinson (1962) *Economic Philosophy*, Watts, reprinted 1964, 1966, Harmondsworth: Penguin/Pelican

George C Roche III (1977) 'The Relevance of Friedrich A. Hayek' in Fritz Machlup (ed) *Essays on Hayek* London: Routledge & Kegan Paul

Nathan Rosenberg (1989) 'Bernard Mandeville' in Eatwell et al (1989: 196-198)

Nathan Rosenberg (1990) 'Adam Smith as Social Critic' *The Royal Bank of Scotland Review* 166 'Special Adam Smith Edition', June, 17-33

Rudy Rucker [Rudolf vB Rucker] (1e: 1982; 2e: 1995; Penguin ed: 1997) *Infinity and the Mind. The Science and Philosophy of the Infinite* London: Penguin Books

George H Sabine (3e: 1951) *A History of Political Theory* London: Harrap

Andrew Schotter (1985) *Free Market Economics. A Critical Appraisal* New York: St Martin's Press

John R Searle (1980) 'Minds, Brains and Programs' in Hofstadter and Dennett (1982) Ch 22

John Searle (1984) *Minds, Brains and Science. The 1984 Reith Lectures* London: BBC

Amartya Sen (1970) *Collective Choice and Social Welfare* London: Holden-Day (Series title: *Mathematical Economics Texts* eds K Arrow et al)

Amartya Sen (1983) 'The Profit Motive' *Lloyds Bank Review* No 147, January 1983: 1-20

Amartya Sen and Bernard Williams (eds) (1982) *Utilitarianism and beyond* Cambridge: Cambridge University Press

Arthur Shenfield (1977) 'Scientism and the Study of Society' in Machlup (1977a) 61-72

Robert Skidelsky (1975) *Oswald Mosley* London: Macmillan

For further discussion of works by Adam Smith, please refer to the bibliographical note inserted as Section 2 of Chapter 4 of this thesis.

Adam Smith (1976-1980) *The Glasgow Edition of the Works and Correspondence of Adam Smith* Oxford: Clarendon Press/OUP (reprinted (1981-82) Indianapolis: Liberty Fund):

Volume I (1976/1776), *The Theory of Moral Sentiments*, ed DD Raphael and AL Macfie

Volume II (1976/1759), *An Inquiry into the Nature and Causes of the Wealth of Nations* (in two volumes), ed RH Campbell, AS Skinner and WB Todd

Volume III (1980) *Essays on Philosophical Subjects*, ed WPD Wightman and JC Bryce

John Maynard Smith (1982) *Evolution and the Theory of Games* Cambridge: Cambridge University Press

Jan Christiaan Smuts (1996/1926) *Holism and Evolution* Highland, NY: Gestalt Journal Press

Elliott Sober and David Sloan Wilson (1998) *Unto Others. The evolution and psychology of unselfish behavior* Cambridge, Massachusetts and London: Harvard University Press

Jim Tomlinson (1990) *Hayek and the Market* London: Pluto Press (Series title: *Pluto Perspectives* ed Robert Moore)

John Toye (1997) 'Keynes on population and economic growth' *Cambridge Economic Journal* 21, 1, January, 1-26

Arnold Toynbee (1972) *A Study of History* (new edition revised and abridged from the 12-volume edition (Toynbee, 1920-61) by Arnold Toynbee and Jane Caplan) Oxford University Press/Thames and Hudson

Jacob Viner (1927) ‘Adam Smith and Laissez Faire’ *Journal of Political Economy* Vol XXXV; a lecture given in 1926 and reprinted in Jacob Viner (1958), 213-245

Jacob Viner (1958) *The Long View and the Short* Glencoe, Ill: The Free Press

Jack J Vromen (1995) *Economic Evolution. An Enquiry into the foundations of New Institutional Economics* London: Routledge (Series title: *Economics as Social Theory* ed Tony Lawson)

William Wallace (ed, trans) (1975) [1st English ed: 1873, 1st German ed: 1830] *Hegel's Logic* Oxford: Clarendon Press/OUP

AMC Waterman (1997) ‘Economics as Theology: Adam Smith’s *Wealth of Nations*’, paper presented to the Southern Economic Association.

WPD Wightman (1980) ‘Introduction’ in Adam Smith (1976-1980) Volume III *Essays on Philosophical Subjects*, ed WPD Wightman and JC Bryce.