

The Pluralism of the Complexity Era

Individualism, Institutionalism and Holism in Complexity Economics

By David Bassett

Supervisor: Jack Vromen

Advisor: Geoffrey Hodgson

Research Masters Thesis (30 ECTS) for Research Master in Philosophy and Economics, Faculty of
Philosophy, Erasmus Universiteit Rotterdam

July 2012

Word Count: 31,625

*“Complexity changes everything; well, maybe not everything,
but it does change quite a bit in economics.”*

David Colander¹

¹ Colander (2000a, pp. 31)

Contents

Chapter 1	4
1.1 The ‘emergence’ of the complexity era	5
1.2 Complexity economists puzzling views on methodology	8
1.3 Chapter 2: Neoclassical methodological individualism as causal individualism.....	12
1.4 Chapter 3: Individuals, institutions and social wholes in complexity economics.....	14
1.5 Chapter 4: Conclusions.....	15
Chapter 2	16
2.1 Methodological holism and institutionalism.....	18
2.2 Methodological individualism	23
2.3 Institutional individualism	25
2.4 Context dependent causality	29
2.5 Causal individualism, institutionalism and holism	35
2.6 Definitions	39
2.7 Are institutions properties of individuals or properties of social wholes?.....	40
2.8 How neoclassical reductionist views lead to causal individualism.....	42
2.9 Conclusions.....	50
Chapter 3	51
3.1 Complexity, agent-based modelling and emergence	52
3.2 Sugarscape	54
3.2.1 Individuals in Sugarscape	58
3.2.2 Institutions in Sugarscape	59
3.2.3 Wholes in Sugarscape: Emergent patterns and their complex explanations.....	59
3.2.4 Conclusions on Sugarscape	65
3.3 Kirman’s fish markets studies.....	65
3.3.1 Individuals in Kirman’s fish markets studies.....	69
3.3.2 Institutions in Kirman’s fish market studies	70
3.3.3. Emergent wholes in Kirman’s fish market studies	72
3.3.4 Conclusions on Kirman’s fish market studies	78
3.4 Conclusions.....	78
Chapter 4	80
4.1 The pluralism of the complexity era	81
References	82

Chapter 1

The Puzzling Nature of Complexity Era

“[E]conomics is currently undergoing a fundamental shift in its method, away from neoclassical economics and into something new. Although that something new has not been fully developed, it is beginning to take form and is centred on dynamics, recursive methods and complexity theory”

Colander, Holt & Barkley (2004), pp. 485

1.1 The ‘emergence’ of the complexity era

David Colander, Richard Holt & Barkley Rosser (2004, 2011)² have argued that mainstream economics is undergoing a radical transformation. At the heart of this transformation is what they call ‘the complexity approach’. These authors make bold claims about the prospects of complexity economics, going so far as to argue that a new complexity era is being birthed, replacing the neoclassical era of mainstream economics. Something like one of Thomas Kuhn’s ‘paradigm shifts’ (Kuhn 1970) is occurring according to these authors, although in their view this shift is occurring in quite a different fashion to how Kuhn outlined. More gradual in nature, evolutionary rather than revolutionary, they see the complexity era as being brought about through incremental change *within* the mainstream, rather than by a dramatic overhaul of prevailing ideas arising from challenges from without.

The authors lay out an alternative account of scientific change to Kuhn, based on a helpful understanding of the terms ‘mainstream’, ‘orthodoxy’, and ‘heterodoxy’. “Mainstream economics consists of the ideas that the elite of the profession find acceptable, where by ‘elite’ we mean the leading economists in the top schools” (CHR 2004, pp. 490). Mainstream economics is thus primarily a sociological, rather than intellectual category, concerning the *status* of ideas. It does not describe a particular school of thought and may in principle be quite heterogeneous if a wide variety of different views are found acceptable by elite economists. In contrast, “orthodoxy generally refers to what historians of economic have classified as the most recently dominant ‘school of thought’, which today is ‘neoclassical economics’” (*ibid.*). The orthodoxy is thus backward-looking and monolithic; it is identified after the fact as historians try to make sense of the prevailing views of a time and categorise them³. ‘Heterodoxy’ is defined negatively, pertaining to those who are against the prevailing orthodoxy. Today, Marxism, Post-Keynesianism, Feminism, Old Institutionalism, and Austrian economics are the main ‘heterodox’ schools.

According to these authors, the mainstream undergoes change in its beliefs through the work of economists at ‘the edge of economics’, “that part of mainstream economics that is critical of orthodoxy, and that part of heterodox economics that is taken seriously by the elite of the profession” (*ibid.*, p. 492). Mainstream economics and the orthodoxy are far from identical; it is perfectly possible to be a mainstream economist but hold heterodox views. Indeed, elite economists have often departed from orthodox ideas, sometimes of their own making. John Hicks renounced IS-LM late in his career for instance (Hicks 1980). It is this aspect that distinguishes CHR’s account of scientific change most

² Henceforth CHR.

³ At the time when the orthodoxy is being established it generally goes unnamed; it is cutting edge material. Often, the name for the school identified as the orthodoxy comes from ‘heterodox’ economists. It was Marx (1847) who labelled ‘classical’ economics, and Veblen (1900) who coined the term ‘neoclassical’.

clearly from Kuhn's. The mainstream is capable of being quite pluralistic and undergoing change from within. Change does not require an external invasion, a dramatic transfer or indeed, funerals.

Today's orthodoxy remains neoclassical. Neoclassical economics involves the 'whole trinity' of rationality, selfishness and equilibrium; its analysis "focuses on the optimising behavior of fully rational and well-informed individuals in a static context and the equilibria that result from that optimisation" (CHR 2004, pp. 490). Chicago-style Price Theory, Game Theory, Decision Theory, Social Choice Theory, New Institutional Economics and 'Modern' Macroeconomics, as many undergraduate and postgraduate courses now refer to the approach to macroeconomics that utilises a representative agent (the 'New Classical' and 'New Keynesian' schools) are the main branches of neoclassical economics.

But today's mainstream of economics is not completely neoclassical. Otherwise elite economists – Paul Krugman, Joseph Stiglitz, George Akerlof, Richard Thaler, Thomas Schelling, Kenneth Arrow, Robert Solow and Amartya Sen to name a few – work on material that is definitely unorthodox, yet their work is discussed by the mainstream. Several of these authors have engaged with complexity ideas, perhaps none more vigorously than Kenneth Arrow. For CHR, Arrow demonstrates clearly how mainstream economics is withdrawing from neoclassical ideas towards complexity. Initially a key figure in the development of neoclassical orthodoxy, more recently he has been heavily involved in complexity economics, as a close associate of the Santa Fe Institute, and co-editing the seminal *The Economy As An Evolving Complex System* (Anderson, Arrow, Pines 1988) which did so much to push the field forward in its infancy.

Just what is 'the complexity approach to economics'? CHR characterise the complexity era first and foremost by the belief that the economy is a complex system. Beyond this however it is difficult to be much more precise, as the notion of complexity is itself extremely equivocal⁴. More than merely involving the adoption of a particular technique or assumption, CHR argued that the complexity era involves a number of different branches of economics, just as neoclassical economics has a number of different core theoretical approaches. They argue that the complexity era encompasses many of the major developments in economics over the last twenty or thirty years or so. Everything from evolutionary game theory, ecological economics, behavioural economics, complexity theory, agent-based computational economics, approaches based on non-linear dynamics, experimental economics and some developments in econometrics are considered by these authors as participating in the new complexity paradigm shift.

There are a few reasons to doubt the narrative CHR put forward for the recent history of economics. First, it will probably come as a surprise to some economists involved in the fields they

⁴ The MIT physicist Seth Lloyd provided over 45 definitions, indicating just how much disagreement there is on what is meant by complexity (Horgan 1997, pp. 303, f11).

list that they are participating in a 'complexity era'. The link between agent-based computational economics and complexity is quite obvious, but one suspects that it would be news to behavioural and experimental economists in particular that they are involved in complexity economics. Is it really appropriate to draw in all these diverse fields together under the idea of 'complexity'?

Second, the scale and significance of the alleged shift in mainstream economics to complexity economics is questionable. While many are optimistic about the prospects of complexity theorizing being welcomed into the mainstream of economics, currently complexity economics is mainly regarded sceptically by the mainstream. If one looks narrowly at those fields that would definitely self-identify as being complexity economics, it is not altogether clear that they are making strong inroads into the mainstream. Articles using simulations – which are central to agent-based modelling, and are even understood by some to define complexity (Bedau 2003) - are occasionally published in mainstream journals, but mostly complexity economics is still to be found in its own fringe journals, such as *Journal of Economic Behavior and Organization* and the *Journal of Economic Dynamics and Control*. Aki Lehtinen & Jaaki Kuorikoski (2007) discussed “Why do economists shun simulations?”, and come up with some compelling evidence. Only 47 hits in JSTOR and 112 hits in the Web of Knowledge for the top five journals (American Economic Review, Journal of Political Economy, Econometrica, Quarterly Journal of Economics, and Review of Economic Studies) contained the word 'simulation'. “Furthermore, there is no visible trend towards its acceptance in these journals: on the contrary, many contributions were published in the 1960s when simulation was a new methodology.” (ibid., pp. 305).

High profile involvement by some elite mainstream economists does not seem to be providing the impetus expected by CHR then. They recognise the main limitation: “the current elite are relatively open minded when it comes to new ideas, but quite closed minded when it comes to alternative methodologies” (CHR 2004, pp. 493). Paul Krugman is quite typical here; while he has expressed sympathies with the view that the economy is a complex dynamic system (Krugman 1996a), he remains committed to simple models (Krugman 2000).

Complexity remains heterodox then, for the time being at least. We may more accurately say that currently mainstream economics has the *potential* to undergo a transition to the complexity era away from neoclassical economics. Should that come about, it would seem to be occurring in much the way CHR describe. The change is gradual, subtle and in some respects unconscious. In contrast to the Kuhnian view of scientific change, whereby the nature of the revolution is well understood by its participants as fundamental core beliefs change, on these authors account changes in science can be so gradual that “the profession often does not notice that the change has occurred” (Colander, Holt & Barkley 2004, pp. 489).

This provides a function for historians and philosophers of economics: to examine the methods, assumptions, and beliefs involved in innovations in economics and see how they compare to earlier orthodox methods. This thesis is intended to contribute to our understanding of contemporary

economics in this way. It will be up to future historians to fully describe the complexity era paradigm shift, should it occur. But we already have plenty of material to try and understand what is involved in this potential shift, methodologically and philosophically. There is plenty of complexity economics around after all. What are the fundamental methodological beliefs that the complexity era and neoclassical economists share, and – perhaps more importantly – what are the differences? Are there any principled reasons why neoclassical economists have not embraced complexity more wholeheartedly, or they simply being a bit old fashioned?

Several commentators have already undertaken methodological discussion of complexity economics, especially to examine how it relates to the mainstream. To date, most have focused on particular traits concerning modelling techniques and changes in fundamental concepts such as rationality⁵. There has not been, to my knowledge, a concerted take on the transition to the complexity era in terms of the more familiar philosophical thesis taken to be one of orthodox economics' most important commitments – methodological individualism. This is perhaps a surprising omission in the literature. My intention is to fill this gap: in short, my research question is: 'Is complexity economics methodologically individualist?'

This is not an easy question to address as methodological individualism can be understood in a variety of ways (Hodgson 2007, Udehn 2001, 2002), making this question altogether too vague as it stands. Since my intention is to understand any differences between complexity and neoclassical economics, the question can be better reformulated as: "Is complexity economics methodologically individualist in the manner of neoclassical economics?"

1.2 Complexity economists puzzling views on methodology

Complexity economists have fairly frequently made remarks that reflect some considerable ambiguity about whether their work is methodologically individualist, or whether it rather follows in the tradition of the opposing approaches of institutionalism or holism. For instance, two economists whose work has advanced the field as much as anyone, Joshua Epstein and Robert Axtell, briefly considered the matter in their seminal *Growing Artificial Societies: Social Science from the Bottom-up* (Epstein & Axtell 1996). On the one hand, the authors write:

"Our point of departure in agent-based modelling is the individual: we give agents rules of behavior and then spin the system forward in time and see what macroscopic social

⁵ Lehtinen & Kuorikoski (2007) examine the reluctance of the mainstream to accept simulations, Perona (2007), and Wible (2007) look at competing definitions of complexity, and Kirman (2011a) and Picton (2007) consider the refashioning of the meaning of rationality in complexity economics.

structures emerge... To that extent our work can be accurately characterized as 'methodologically individualist'.

Epstein & Axtell 1996, pp. 16

This is not the whole story however, as they also recognise the central role played by social institutions in their analysis, which they describe as 'collective structures'.

"However, we part company with certain members of the individualist camp insofar as we believe that the collective structures, or 'institutions' that emerge can have feedback effects in the agent population, altering the behaviour of individuals. Agent-based modelling allows us to study the interactions between individuals and institutions."

Epstein & Axtell 1996, pp. 16-17

Using social institutions as exogenous terms in explanations in this way is usually taken to be characteristic of methodological holism. Indeed, much of the intention behind methodological individualism was to endogenise social institutions. John Harsanyi (1968) put forward a common view of methodological individualism as meaning that "social norms should not be used as basic explanatory variables in analysing social behavior, but rather should be themselves explained in terms of people's individual objectives and interests" (Harsanyi 1968, pp. 321). Epstein & Axtell's remarks on this matter demonstrate a considerable degree of confusion then; at least according to some interpretations of methodological individualism, these two quotations contradict each other.

Even more puzzling, Mauro Gallegati and Matteo Richiardi (2010) regard complexity economics as achieving a kind of '*reconciliation*' between methodological individualism and holism (Gallegati & Richiardi, pp. 36), insofar as the complexity approach to economics seems to have some aspects which appear individualistic and others which appear holistic. For Gallegati and Richiardi, the 'individualism' is found as complexity economics proceeds with interacting individuals, while the 'holism' arises in the sense that the emergent properties examined in complexity economics are 'more than the sum of their parts' and also from the view that "the whole appears to act as if it followed a distinct logic, with its own goals and means" (*ibid.*). Individualism and holism are often thought irreconcilable, so it really would be quite something if complexity economics somehow reconciles them.

Other authors associated with complexity economics have suggested that in this emerging field the unit of analysis has escalated up from the individual to examine collective entities, principally markets. As Philip Mirowski put it in his recent proposal to direct the new paradigm to take as its central theme the idea that markets are computational algorithms, "the laws that are sought under the

new paradigm are laws of the markets, not laws of human nature” (Mirowski 2007, pp. 209). Mirowski’s particular method of treating markets as algorithms was met with some disquiet in the responses that followed in the issue of the *Journal of Behavior and Organization* that was dedicated to Mirowski’s typically flamboyant article, but there was also significant agreement with the broader proposition concerning the change in the unit of analysis. “I am in large agreement with the Mirowski position and think that markets provide very useful basic units of analysis” replied Alan Kirman (Kirman 2007, pp. 293).

Kirman has also written sympathetically of the view associated with Werner Hildenbrand (Kirman 2011b). Responding to the devastating results of Sonnenschein (1972), Mantel (1974) and Debreu (1974)⁶, which when taken seriously do serious damage to the general equilibrium project, Hildenbrand argued that we should simply abandon the individual in favour of observing collective behaviour and generating hypotheses about it to form higher level laws (Hildenbrand 1994). This statistical approach that looks for aggregate laws represents a kind of methodological holism that individualists have been critical of, questioning whether any such higher-level laws ever obtain (Popper 1944-5, Lucas 1976).

Kirman is not the only one to use ideas from complexity theory – most crucially the notion of emergence - to advocate holist themes. In a well-known article “The Financial Crisis and the Systemic Failure of the Economics Profession” (Colander et. al. 2009), a number of economists have castigated the economics profession for not only failing to forewarn of the financial crisis, but for positively stimulating it. The authors criticised the “extreme form of reductionism” of the representative agent, microfoundations approach.

“The dominance of the extreme form of conceptual reductionism of the representative agent has prevented economists from even attempting to model such all important phenomena [primarily financial crises]. It is the flawed methodology that is the ultimate reason for the lack of applicability of the standard macro framework to current events.”

Colander et al. 2009, pp. 258

⁶ These results showed that even with the usual strong assumptions on individuals preferences, the equilibria of economies are not necessarily stable under the usual *tatonnement* adjustment process which is taken to be the mechanism by which equilibria are maintained. If we cannot be sure that economies will arrive at equilibria, these states become of questionable interest; comparative statics becomes useless if those states are unlikely to be reached. Smale (1976) showed that stable equilibria could be achieved, but only at the expense of a vast increase in the amount of information required by individuals. The minimal amount of information that was thought to be required in the standard model – prices alone – was largely the whole point of the exercise. The efficient markets hypothesis, so essential in financial economics, becomes redundant, since it is essentially a thesis that says that all information is represented by prices.

The authors advocate ‘a new kind microfoundations’, which diverts emphasis towards the understanding of the effects of the *interaction* of agents, rather than individual’s particular tastes and endowments. “Economists’ micro foundations should allow for the interactions of economic agents, since economic activity is, essentially, interactive.” (*ibid.*). Although not explicitly articulated in the language of methodological individualism, which is by now somewhat outdated in favour of ‘microfoundations’, these critics are discussing essentially the same issues that were part of the older debates over methodological individualism, institutionalism and holism. Arguing for ‘microfoundations’ that emphasise interactions is essentially an argument for a focus on institutions, since institutions are the rules by which interactions occur. And echoing Kirman, they argue that the financial system should be viewed as “an involuntary emergent phenomenon” (*ibid.*), citing Philip Anderson’s article “More Is Different” (Anderson 1972) that argued for the distinctiveness of scientific domains on the basis of an emergentist view. This can be understood as a call for the legitimate return of macroeconomics as a domain distinct from microeconomics. These authors appear to be calling for a return of what would earlier have been called ‘holism’.

Similarly, the Post-Keynesian economist Steve Keen works with non-linear dynamic systems dynamics models of aggregates, justifying this holism on the basis of an emergentist view.

“Macroeconomic phenomena – and even phenomena within one market – are emergent properties of the dynamic, disequilibrium interactions of individuals and social groups in a rich institutional environment, constrained by the physical, temporal and environmental realities of production. These phenomena will not be predictable from the behavior of isolated individuals. Instead, macroeconomics is a self-contained field of analysis, and must be reconstructed as such.”

Keen 2011, pp. 358

Today’s controversies about economic methodology ultimately amount to precisely the same issues that have vexed economists and philosopher of social science for decades. Are macroeconomic phenomena reducible to individuals? What role should institutions play in economic analysis? Complexity economics has added renewed impetus to these matters but the conflicting and puzzling remarks made by complexity economics are evidence of considerable confusion. Does the way in which the models of complexity economics are built out of individuals mean that complexity economics deserves the label ‘methodological individualism’, and is therefore consistent with the methodology of mainstream economics? Or does a ‘new kind of microfoundations’ that includes institutions as exogenous terms in explanations amount to a rejection of methodological individualism? Are Gallegati & Richiardi correct in viewing complexity economics as achieving a

kind of *reconciliation* between individualism and holism? Are Kirman, Mirowski and Keen justified in advocating methodological holism on the back of their emergentist view?

To address these matters this thesis is divided into two main chapters. The following chapter is occupied with issues surrounding the definition of methodological individualism, institutionalism and holism. Chapter 3 then examines some examples of complexity economics, and observes whether they involve individualism, institutionalism and holism. Before proceeding to those chapters, the remainder of this introductory chapter guides the reader through what they can expect by providing a précis of the overall argument of each, and summarises my conclusions.

1.3 Chapter 2: Neoclassical methodological individualism as causal individualism

The confusion found in complexity economists' remarks about their methodology primarily arises because of ambiguities about what precisely is involved in methodological individualism, institutionalism and holism; these ideas can all be understood in a variety of ways, hence the confusion. My intention in this chapter is not to discover a 'once and for all' definition of methodological individualism – this is probably impossible as different authors have very different understandings (Udehn 2001, 2002). Rather, my aim is to articulate the brand of methodological individualism that is particular to neoclassical economics, as it is this that is apposite to the research question.

To define the kind of methodological individualism involved in neoclassical economics it is necessary to understand the exact methodology that it involves, why, and what methods it regards as problematic. This is no easy matter as individualism and holism have been much discussed and there is no agreed upon definitions, despite all the ink. The first part of the chapter will play with the various crude and imprecise conceptions that one finds in the literature, note the key distinctions that these definitions are trying to articulate and then recount the main difficulties that these definitions face. I then work towards my own definitions, gradually overcoming these difficulties and becoming more precise. I come up with an innovative approach to defining the kind of methodological individualism that is involved in neoclassical economics, which will then allow me to compare the methodologies of complexity and neoclassical economics in the following chapters.

I begin by looking at the historical record of those branches of economics that have been branded individualist, institutional and holist, noting key similarities and differences in these schools, drawing out the main difficulties that surround the usual formulations of their definitions. Far from specifying a particular method 'holism' has come in different forms. One form of holism involves analysing the interplay of aggregates, with no regard for how these aggregates are made up out of individuals. This may involve statistical analysis (e.g. the Philips Curve, the Kuznets Curve) or macro modelling (e.g.

Goodwin's theory business cycles). Others treat groups as having functions and goals in their own right (e.g. Marxists, John Kenneth Galbraith). Others still focus on how institutions drive social change but ground their institutional analysis within the actions of individuals (e.g. Thorstein Veblen). Meanwhile, methodological individualism is commonly taken to be a description for 'explanations that *are in terms of individuals*', meaning that in individualist explanations the only permissible exogenous terms are properties of individuals. This is taken by some to be the definition of methodological individualism appropriate to neoclassical economics (Udehn 2001, 2002). I discuss how this definition confronts problems, especially concerning the status and treatment of institutions. Some have thought that institutions are social wholes: the individualist authors of New Institutional Economics attempted to endogenise institutions on the basis that they are social wholes requiring explanation in terms of individuals. But this appears to be a failed project, insofar as some of the main protagonists have admitted that this is impossible; all explanations appear to involve (often tacit) assumptions about institutions. Similarly, institutions have been recognised in the other most obviously avowed example of individualism in neoclassical economics – General Equilibrium Theory. Accordingly, several have argued that the *strong* form of individualism thought to describe neoclassical economics – which seeks to advance explanations *only* in terms of individuals – has never been committed. This has led to the weaker thesis of *institutional individualism* (Agassi 1960, 1975), which accepts institutions as exogenous terms in explanations along with individuals, seeming like the ultimate end point of all this discussion. Under institutional individualism, the picture is of institutions being on the same level as individuals, and so their involvement is not taken entail 'holism'.

Perhaps surprisingly then, there appears to be no substantive difference between Veblen's allegedly 'holist' position and the 'individualist' neoclassical authors'. Institutional individualism seems to be the end of the road somewhat, but ultimately leaves the kind of disappointment that can come with mediocre 'middle road' resolutions – as well as, for me at least, a lingering suspicion that institutional individualism does not capture the whole story. It is not that institutional individualism is entirely empty: it mandates that explanations are required to involve institutions *through* individuals, rather than groups, and so is distinguished from some holism. But is it really the case that there are no substantive differences in how Old Institutional authors like Veblen and the New Institutional economists involve institutions and individuals in their analysis?

I believe that there is in fact an important difference to be made here; that Veblen and Douglass North were not doing the same thing. In order to capture the difference it is necessary to examine more closely a part of the notion of methodological individualism that is sometimes overlooked: the notion of *explanation* that is involved. Hitherto, discussions of methodological individualism have mostly presumed that the relevant notion of explanation involved is theoretic reduction (Nagel 1961). I instead look at how these ideas are recast if one uses the more general category of *causal explanation*, drawing on some philosophy of causality to create some definitions for what I call *causal individualism, institutionalism and holism*. These are based on the idea that individualists,

institutionalists and holists can be distinguished not by what kinds of entities they make assumptions about, but by which entities they highlight as being *causal*.

For it is commonly agreed among philosophers of causality that causal claims rely on background assumptions that are often tacit; causal claims are not simple binary relations between cause and effect where all assumptions play an equal role, rather some parts of the antecedent in explanations are highlighted as being causes while others are mere background conditions. Background conditions specify a field within which a causal relation holds. Causal relations are then demonstrated through a quaternary structure; they involve showing how *changes* in the cause (c rather than c^*) bring about *changes* in effect (e rather than e^*). Accordingly, individualist explanations involve tacit assumptions about institutions, but these are only mere background conditions. I submit that individualism, institutionalism and holism can be distinguished as different approaches to social science that take different kinds of entities as being the ‘causes’. Individualists show how variation in the properties of individuals generates different social outcomes; institutionalists show how variation in institutions generates different social outcomes; holists show how variation in social aggregates generate different social outcomes. I demonstrate how this applies to an important example in neoclassical economics: I show how the Ramsey-Cass-Koopmans model is causally individualist.

Recasting individualism, institutionalism and holism in causal terms enables one to pick out an important difference between old and new institutional authors that is overlooked by the weak view of methodological (i.e. institutional) individualism. Causal individualism represents a stronger and narrower form of individualism than the idea of institutional individualism captures. More than this though, for neoclassicals methodological individualism has been a normative thesis instructing how explanation in the social sciences *should* proceed. In the final part of the chapter I show why neoclassical economists have been led to causal individualism, by showing how the reductionist ideas involved in their methodological individualism give it a normative force that leads to causal individualism.

1.4 Chapter 3: Individuals, institutions and social wholes in complexity economics

Having created a method for taxonomising individualist, institutional and holist social science and shown how neoclassical methodological individualism as a normative thesis leads to causal individualism, I turn to examine whether the complexity era involves causal individualism, institutionalism or holism. As already noted, CHR view the complexity era in broad terms such that it involves evolutionary game theory, ecological economics, behavioural economics, complexity theory, agent-based computational economics, non-linear systems dynamics, experimental economics and some developments in econometrics. Some of these fields might resist inclusion in ‘the complexity era’ so I select a field whose engagement with complexity is surely uncontroversial, namely agent-

based computational economics, and examine some case studies to draw out whether the causes cited in this modelling approach are individuals, institutions or social wholes.

I examine two case studies: Axtell & Epstein's Sugarscape models (Axtell & Epstein 1996), and Alan Kirman's studies of fish markets (Kirman 2011a, Kirman & Vriend (2001), Hardle & Kirman 1995, Gallegati et. al. 2011). The case studies demonstrate that the complexity era is pluralist, as the authors demonstrate individualist, institutional *and* holist causal claims at various times. Agent-based models are capable of demonstrating how variation in individual properties generates differences in social outcomes, although they primarily involve showing how differences in institutions generate differences. That is, agent-based modelling is primarily causally institutionalist, but may also engage in causal individualism. Further, holist causal claims are generated as irreducible emergent patterns are produced that allow one to show how variation in an aggregate brings about variation in other aggregates (e.g. the price-quantity relationship).

1.5 Chapter 4: Conclusions

By now I have created a system for taxonomising fields and explanations in social science according to whether they highlight individuals, institutions or social wholes as being the causes of social phenomena, shown how and why neoclassical economics exclusively looks at individualist causes because of the commitment to reduction involved in methodological individualism, and how in agent-based computational economics individuals, institutions and social wholes are at various points demonstrated as causes. The conclusion I can draw then is that the *transition to the complexity era from the neoclassical era involves the transformation of mainstream economics from being a science preoccupied exclusively with individualistic causes to being a science that examines a plurality of different kinds of causes*. In short, complexity economics rejects methodological individualism as a normative thesis.

This transformation arises out of the characteristic belief of the complexity era that the economy is a complex system; this belief entails that the economy will not submit to the reductive explanatory strategy that led to an exclusive focus on individualist causes in the neoclassical era.

Chapter 2

Neoclassical methodological individualism as causal individualism

“Economics is about what individuals do: not classes, not ‘correlations of forces’, but individual actors. This is not to deny the relevance of higher levels of analysis, but they must be grounded in individual behavior. Methodological individualism is of the essence.”

Paul Krugman (1996b)

The aim of this chapter is to articulate the kind of methodological individualism that is involved in neoclassical economics. Unfortunately methodological individualism has remained riddled with definitional imprecision, and this has meant that individualism can, and has been, understood in a variety of different ways. Likewise for individualism's alleged opposite, holism; here too a number of forms can be found in the literature, both in practice and as argued for explicitly as a methodological doctrine. Accordingly, it is exceedingly doubtful that any 'once and for all' definitions for methodological individualism and holism could ever be achieved. But as I will do in this chapter, it is possible to define the kind of methodological individualism that is involved in neoclassical economics precisely.

Methodological holism and methodological individualism can be loosely characterised by the primacy given to social wholes or individual actors in explaining social phenomena. "The holist focuses attention on how social 'forces' (institutions, social conventions, etc.) condition individual behavior" (Rutherford 1994, pp. 28), while individualists emphasise "how individual action gives rise to institutions and institutional change" (Rutherford 1994, pp. 27). Commonly, definitions of methodological individualism make the primacy just spoken of absolute. Jon Elster for instance defined it as "the doctrine that all social phenomena (their structure and change) are in principle explicable *only* in terms of individuals – their properties, goals, and beliefs" (Elster 1982, pp. 453, my emphasis). On this account, institutions are social wholes and are not permissible as exogenous terms in explanations; they require endogenisation. Lars Udehn (2001, 2002) claimed that neoclassical economics follows this form of 'strong' methodological individualism.

But this strong form of methodological individualism has faced rightful criticism. Several authors have pointed out that institutions play a necessary role in all individualistic explanations (Rutherford 1994, Arrow 1994, Hodgson 2007). Accordingly, it has been argued that *institutional individualism* (Agassi 1960, 1975) - a considerably weaker doctrine that allows institutions to be used as exogenous terms in explanations – is as strong as methodological individualism gets.

In this chapter I argue that the kind of methodological individualism involved in neoclassical economics is in fact stronger than institutional individualism. There are good *prima facie* reasons for believing so, as institutional individualism is a very broad category that brings together economists whose methodologies are sometimes thought to be opposed to one another. I argue that individualism and institutionalism can be distinguished not by which kinds of entities are taken as exogenous terms in explanations – as the discussions around methodological individualism are usually formulated – but by which of these entities are given a *causal* role in the explanation, as opposed to merely appearing in the *background conditions*. Drawing on some philosophy of causality, I create a taxonomy of the social sciences that distinguishes between *causal individualism*, *causal institutionalism* and *causal holism*. I then show how neoclassical economics involves causal individualism.

For neoclassical economists, methodological individualism is more than just a name for a particular approach to social science though; it is a thesis with normative force that instructs how

explanation *should* proceed. Having created my causal taxonomy of methodologies, I then discuss how the normative force of neoclassical methodological individualism comes about because of its commitment to *reductionism*. I then argue that it is this commitment to reduction that leads to causal individualism, not causal institutionalism. In short, neoclassical methodological individualism insists on causal individualism.

This chapter is structured as follows. I begin by looking at the various ways in which the categories of methodological individualism and holism have historically been understood. I describe how both ‘holism’ and ‘individualism’ have in fact taken various forms, and examine the flaws in the usual definitions of methodological individualism and holism accordingly. Section 2.1 examines methodological holism in its varying forms, noting three main themes: functionalism, institutionalism and theories of macro-level regularities. A few key protagonists of holism are discussed to draw out these various approaches within ‘holism’. Section 2.2 concerns methodological individualism and the attempts of theorists categorised as such to endogenise institutions. Section 2.3 then details the common criticism of the strongest form of methodological individualism that seeks to entirely eradicate all institutions from the exogenous part of explanations, by discussing how such strong individualism has never been attained. From section 2.4 I innovative toward my definitions: section 2.4 outlines the relevant ideas from the philosophy of causality which indicate the context dependence of causality; Section 2.5 demonstrates how these ideas apply to the debate around methodological individualism leading to section 2.6 where I articulate the definitions of causal individualism, institutionalism and holism. Section 2.7 deals with the problematic issue of how institutions should be treated. Finally section 2.8 shows how the reductionist views of neoclassical economics leads to causal individualism.

2.1 Methodological holism and institutionalism

Holism is commonly understood as involving the explanation of social outcomes as consequents of various social ‘forces’, which depending on one’s point of view may include institutions, norms, conventions, the functions of social groups, or alleged social ‘laws’ that influence how individuals behave. Various, interactions between cultures, society, and large scale social assemblages (governments, firms, social classes, etc.) are analysed as the key determinants of human experience. Beyond this focus on social wholes however, holism has been conducted in various fashions. Holism does not have a single face; the label ‘holism’ has been put on authors, schools and theories that vary greatly methodologically.

The most clear and incontrovertible form of holism involves theorizing directly on macro-level phenomena like aggregate demand, unemployment rates and inflation and so on; old style macroeconomics. Much economics from the mid part of the twentieth century was holist in this

fashion. The Keynesian revolution helped deliver macroeconomic enquiry that sought causal relationships between social wholes. Drawing on Bernard Mandeville's *Fable of the Bees* (1714), John Maynard Keynes argued that what was true for individuals did not always hold true for society as a whole and as such, that it was necessary to analyse the level of aggregate demand because the whole did not always respond in a way that was harmonious with the individual's response. The *General Theory of Employment, Interest and Money* (Keynes, 1936) appealed to aggregate properties as it evoked a mass psychology driven by 'animal spirits' that determine the level of effective, aggregate demand. Similarly, Richard Goodwin put forward a model of business cycles that had no individualistic basis (Goodwin 1951)⁷.

Much of this work relied heavily on statistics. Keynesian macroeconomics came to rely on the Phillips Curve – a measured relationship between inflation and unemployment, which was understood as a trade-off which policymakers could choose between. Developments in econometrics undoubtedly contributed to the impetus to theorise about and test relationships between aggregate variables based on historic time series. Wesley Mitchell's work on business cycles involved extensive use of statistical aggregates (Mitchell 1927) for instance and defended an analytical approach that did not require a basis in individual decision making, but operated solely on "mass phenomena" (Mitchell [1925] 1950, pp. 25). Likewise Simon Kuznets was led from his empirical investigations into national accounts to postulate the 'Kuznets cycle' of 15-20 years in economic prosperity (Kuznets 1930), as well as an inverse-U shaped relationship between economic development and inequality (Kuznets 1955).

This macroeconomic enquiry was defended by arguing that social aggregates were suitable for theorising about. Clarence Ayres argued against any need to ground analysis in the actions of individuals. For him, "the whole analysis must proceed on the level of generalisation of culture rather than of individuality" (Ayres 1962, pp. 112). Social wholes form a "causal nexus the analysis of which is the problem of the social sciences, and in this analysis of social causes and effects the acts of individual men are not at issue" (Ayres 1962, pp. 97). Similarly, in defending Mitchell's work on business cycles from Tjalling Koopmans's (1947) attack on his "Measurement without theory", Rutledge Vining argued that "the aggregate has an existence apart from its constituent particles and behavior characteristics of its own not deducible from the behavior characteristics of the particles" (Vining 1949, pp. 79).

⁷ In his capacity as advisor for this thesis, Geoffrey Hodgson has questioned whether it is appropriate to call Goodwin (and Kuznets) holists in this way, suggesting that they follow what would today be called a 'representative agent' approach. I cannot see how this is the case, with Goodwin's business cycle theory at least. Goodwin's theory involved differential equations describing the dynamics of aggregates of the economy (workers wages, employment, output, etc.). I suppose these could be interpreted as involving an economy with a single entrepreneur who pays himself profits (etc.), but as far as I am aware this was not Goodwin's intention, nor do the Post-Keynesian economists who continue to use the model interpret it in this way (e.g. Keen 2011). Perhaps most fundamentally, there is no maximisation by agents involved Goodwin's model: it simply proceeds according to identity conditions and assumptions about the dynamics and relationships between aggregate properties (including the Phillips Curve).

Sometimes the theories put forward treated social aggregates such that they have distinct aims and intentions of their own. John Kenneth Galbraith, notably in the series of books *The Affluent Society* (1958), *The New Industrial State* (1967) and *Economics and the Public Purpose* (1973), analysed the interaction of corporations and the state, giving a kind of consciousness to these elements of the “technostructure”. He argued “[a]s one proceeds from the smaller corporations to giants, the role of any single individual diminishes, the authority of organisation increases” (Galbraith 1973, pp. 83). Mostly his analysis relies on attributing functions, goals and interests to collective entities. For example, “[t]he process by which prices that serve the affirmative purposes of the technostructure are established ... derives from the frequently common interests of all firms and the shared sense of what serves those interests” (Galbraith 1973, pp. 116).

Others outside of economics often proceeded with their analysis likewise, evoking ideas of *collective consciousness*. Emile Durkheim’s ground-breaking *Suicide: A Study in Sociology* ([1897] 1951) is a notable early example of such holism. Durkheim amassed data that demonstrated a remarkable variety but also stability in the suicide rates of different societies. Based upon this empirical work, he argued that since the victims of suicide are in a very small minority in society, with each (or at least most) conducting the act separately and independently of others, the stability of different suicide rates indicated that “all these individual manifestations, however independent of one another they seem, must surely actually result from a single cause or a group of causes, which dominate individuals” (Durkheim ([1897] 1951), pp. 269), explicitly rejecting individualist explanations. From the individual’s perspective, one may report that a suicide results from a particular event, such as a divorce or poverty, but for Durkheim these factors do not report the real cause. They are merely a ‘triggering event’, the suicide actually being the result of some underlying social phenomena. The triggers merely serve to expose the source of the causes, but they are not the causes themselves, which are social in nature. Durkheim goes on to analyse the variations in suicide rate in terms of social phenomena, especially what he termed social cohesion.

Durkheim developed ideas involving collective consciousness, which includes both a consciousness about what is, but also a moral aspect concerning how things ought to be. Many theorists, from Georg Wilhelm Friedrich Hegel to Karl Marx, Auguste Comte and Durkheim have evoked some form of collective mind. Durkheim had defined collective consciousness in *On the Division of Labour* (1893) as “the totality of beliefs and sentiments common to the average members of a society that forms a determinate system with a life of its own” (Durkheim 1893, pp. 39). Insofar as collective consciousness has “a life of its own”, society itself has a reasoning that can be analysed.

Historians too have often evoked ideas of a collective conscious. The structuralist Fernand Braudel is one such author, for whom “In every period, a certain view of the world, a collective mentality, dominates the whole mass of society.... A society’s reactions to the events of the day ... are less a matter of logic or even self-interest than the response of an unexpressed and often inexpressible compulsion arising from the collective unconscious” (Braudel 1987, pp. 22). Braudel’s approach to

history is a reaction against the individualistic ‘Big Man’ version of historical narrative, which focuses on the role of individual’s decisions (heads of states and other ‘Big Men’ of history) in shaping events, to instead focus on economic, demographic, technological and ideological ‘forces’.

In the arguments that followed about methodological individualism, it was mostly this kind of holism that was attacked. Max Weber repudiated such talk of “social collectives, such as states, associates, business corporations, foundations, as if they were individual persons” (Weber 1922, pp. 13), favouring that “these collectives must be treated as solely the resultants and modes of organization of the particular acts of individual persons, since these alone can be treated as agents in a course of subjectively understandable action” (ibid.). In fact, the first use of the term ‘methodological individualism’ in English, by Joseph Schumpeter, came in an article in the Quarterly Journal of Economics entitled “On the Concept of Social Value” (1909) in which Schumpeter dismisses the notion of ‘social value’, recently introduced by John Bates Clark, on these grounds. Schumpeter argues that a concept such as ‘value’ is a psychological concept and so it is inappropriate to apply it to social categories. He dismisses the notion of social value by arguing that “only individuals can feel wants” (Schumpeter 1909, pp. 214).

Contrary to those sociologists, historians and Marxists who evoke ‘collective consciousness’, much of the theoretical work by economists did not go so far as to embody social wholes with human characteristics of consciousness and action however. Clarence Ayres argued that “the functions, factors and forces into which culture is resolved do not ‘act’ as men act” (Ayres 1962, pp. 92), but culture nonetheless proceeds according to its own internal dynamic. Macroeconomics did not usually involve imbuing aggregate entities with human-like intentionality; it simply involved theories of macro-level relationships that often relied on identity conditions and empirical testing. It is important to recognise that holists can be distinguished between those that do speak of aggregate entities in intentional terms (Marxists, Durkheim, Galbraith, etc.) and those that don’t (macroeconomists).

Another important difference to be made between holist authors is their attitude towards individuals. Although many within Old Institutionalism explicitly rejected the need to ground their explanations in the behavior of individuals, this is not a necessary feature of ‘Old Institutionalism’. Thorstein Veblen made more concerted attempts to ground his analysis in individuals than other authors from the Old Institutional tradition. Indeed, his methodological writing suggests that he saw such as task as essential.

“It is, of course, on individuals that the system of institutions imposes those conventional standards, ideals and canons of conduct that make up the community’s scheme of life. Scientific inquiry in this field, therefore, must deal with individual conduct and formulate its theoretical results in terms of individual conduct.”

To Veblen's mind individual ends are products of cultural habituation. The individual's "methods of life today are enforced upon him by his habits of life carried over from yesterday and by the circumstances left as the mechanical residue of the life of yesterday. What is true of the individual in this respect is true of the group in which he lives. All economic change is a change in the economic community,- a change in the community's methods of turning material things to account. The change is always in the last resort a change in habits of thought." (Veblen 1898, pp. 391). In his *Theory of the Leisure Class* (1994 [1899]), Veblen traced the evolution of cultural norms of consumption back through successive stages of human history, commencing his analysis in early stages of 'barbarism' to discover some original habits that undergo various changes as institutions of ownership develop as outgrowths of political and technological changes, ultimately ending up with the 'conspicuous consumption' of today's leisure class. As with other Old Institutionalists (such as Clarence Ayres) technology plays an important driving role in much of Veblen's analysis. Despite Veblen's methodological stance which professes that theories need to explain how institutional change becomes reflected in individual action, his actual theoretical work has been judged neglectful on this matter. "In some instances he fails to analyse properly exactly how technological changes alter the situation and goals of individuals or exactly how altered individual ideals actually bring about changes in social conventions and laws" (Rutherford 1994, pp. 39). Nonetheless, Veblen is distinct from others in Old Institutionalism for his efforts to ground his analysis in individual action. He does not proceed as if aggregates form a distinct 'causal nexus', as Ayres had argued.

Why then, if his analysis of institutions is supposed to operate through the actions of individuals (if not always wholeheartedly successfully) is Veblen a 'holist'? It is for his focus on institutions, which are – confusingly and debatably in my view – often considered as social wholes⁸. As we shall see in the following two sections also, there is ambiguity about whether institutions should be viewed as social wholes or whether they are on the level of individuals, and I take up this matter and clarify things in section 2.7. But for the time being it is sufficient to note the ambiguity. Geoffrey Hodgson defines institutions as "systems of established and prevalent rules that structure social relations" (Hodgson 2006). As rules, they generally take the form "in circumstances X, do Y". Since institutional rules are acted out by individuals they can be viewed as being on the 'micro' level. But they can also be viewed as being social rather than individual phenomena. They are 'established and persistent' within *societies*. Analysing them involves assuming that, as Veblen put it, "what is true of the individual ... is true of the group" (Veblen 1898, pp. 391). It is in this sense that the use of institutions

⁸ Rutherford (1994) chapter 2 for instance discusses individualism and holism in old and new institutional economics and understands 'holism' as concerning the use of institutions exogenously.

to explain social change has been accused of ‘holism’. Individualists demand to know *why* individuals adopt these institutions; absent such an explanation, it seems that institutions just fall from the sky.

Holism has come in a number of forms then. Many analyse relationships between aggregate categories as if they form a closed ‘causal nexus’, whether they utilise statistical analysis of aggregates to make causal claims about social phenomena (Mitchell, Kuznets), or put forward theoretical models of aggregate categories directly, often utilising identity conditions (Keynes, Goodwin). Others interpret changes in social phenomena as an interplay of a collective (un)conscious (Hegellians, Durkheim) or utilise functional analysis (Galbraith). Some seek to ground their analysis of social forces in the actions of individuals (Veblen), while others believe this is unnecessary (Ayres, Vining). Finally, explaining social change by reference to institutions is often viewed as form of holism.

2.2 Methodological individualism

Although methodological individualism was only articulated as an explicit philosophical doctrine in the twentieth century, examples of it have been recognised from considerably earlier times. Just as with methodological holism, a number of versions of methodological individualism can be distinguished (Udehn 2001, 2002). Most fundamentally, individualists vary in the degree to which they explicitly involve institutions, mirroring the way holists differed in the degree to which they involved individuals in their analysis. Again this reflects the ambiguity about the status and correct treatment of institutions. Some methodological individualists view institutions as social wholes and try and to entirely endogenise them. Others accept social institutions as exogenous conditions but are still considered as individualists. Lars Udehn (2002, 2003) has argued that there is thus a distinction to be made between *strong* and *weak* forms of methodological individualism. Strong methodological individualism involves explaining social phenomena solely in terms of individuals. Institutions are only ever endogenous. Weak methodological individualism involves explaining in terms of both individuals and institutions.

An early paradigmatic example of strong methodological individualism is social contract theory, which found its main representative in Thomas Hobbes, although the Greek Sophists and Epicurus are also credited with a version of the theory. Social contract theory explains civil society, with attendant laws and a state authority which enforces them, as the outcome of the rational choices of a collection of asocial individuals who are in an uncultured “state of nature”. This state is characterised by a lack of social norms, agriculture, industry or arts, leading to a state of perpetual war and a lived experience that is “solitary, poor, nasty, brutish and short” (Hobbes [1691] 1968, pp. 186). Rationally self-interested individuals are motivated to improve this condition, and so they enter into a contract with others to abide by laws. This is the original contract, but requires a second, that of an authority to

enforce them, as without such an authority the incentive to defect remains. Social contract theory is recognised as an early example of methodological individualism, as a rational choice explanation of social institutions.

Social contract theory diminished after the innovations of the Scottish Enlightenment, including David Hume and Adam Smith, for whom social phenomena were not the calculated products of rational individuals coordinating to improve their lot, but were instead seen as the unintended consequences of the actions of cultured individuals –as examples of spontaneous order. The “invisible hand” of capitalism is perhaps the most widely known example of spontaneous order.

“[E]very individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it... [H]e 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.”

Smith 1776, Book IV, Chapter II, pg.9

There are two elements that are distinctive from social contract theory in this approach: that coordination is unintended, and that the individuals involved are not perceived as being in an early “state of nature” but are already the symptoms society’s history and extant social institutions. Smith articulated the invisible hand memorably as that idea that “[i]t 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” (Smith 1776, Book I, Chapter II, pg.19); the butcher, brewer and baker are not the asocial individuals of social contract theory, but have defined social roles. It hence involves the weaker kind of methodological individualism, on Udehn’s understanding.

In both these instances individual choices are the driving force, although the treatment of individuals differs. Social contract theory sought to eliminate social institutions from the theory, while Smith did not.

The ‘neoclassical economics’ that began with the marginalist revolution of the 1870, ushered in by William Stanley Jevons's *Theory of Political Economy* (1879), Carl Menger's *Principles of Economics* (1876 [1871]), and Léon Walras's *Elements of Pure Economics* (1874–1877), is most explicitly dedicated to a strong form of methodological individualism (Udehn 2001,2002). Indeed it was in admiration of this work that Joseph Schumpeter first advanced the term methodological individualism (Schumpeter 1909). Thorstein Veblen, in a critical article, described the method advocated by the marginalists neatly.

“It is characteristic of the school that whatever an element of the cultural fabric, an institution or any institutional phenomenon, is involved in the facts with which the theory is occupied, such institutional facts are taken for granted, denied, or explained away. If it is a question of price, there is offered an explanation of how exchanges may take place with such effect as to leave money and price out of the account. If it is a question of credit, the effect of credit extension on business traffic is left on one side and there is an explanation of how the borrower and lender cooperate to smooth out their respective income streams of consumable goods or sensations of consumption.”

Veblen 1909, pp. 621-22

Walras is a particularly important figure as his ideas of pure economics, where isolated individuals respond to prices, remains a central plank of economics today, through General Equilibrium Theory and modern macroeconomics. Walras's ideas were clarified and purified into contemporary General Equilibrium Theory by Kenneth Arrow and Gerard Debreu. These authors lay out the foundations of this approach and proved the existence, if not the uniqueness or stability, of general equilibrium. Arrow is explicit about the intention to remove institutions from the analysis: "every relevant variable, except those classified as exogenous for the whole economic system, is the result of a decision on the part of some one individual unit of the economy" (Arrow 1959, p. 42).

Economists involved in the programme of New Institutional Economics were also explicitly committed to the strong kind of methodological individualism that sought to rid entirely endogenise institutions. John Harsanyi (1968) argued that "social norms should not be used as basic explanatory variables in analysing social behavior, but rather should be themselves explained in terms of people's individual objectives and interest" (Harsanyi 1968, pp. 321). Often this involved an approach which explained social institutions by their efficiency. Douglass North and Robert Thomas (1973) treat institutions as if "they were selected from a book of organisational blueprints, given endowments, technologies and preferences, but no exogenously specified rules.... The social problem, at each moment, is to choose the rule set which maximises output net of overhead costs." (Field 1981, pp. 184).

2.3 Institutional individualism

It has been noted how some holists like Veblen sought to show how institutions operate through the actions of individuals, although some feel this is unnecessary. Individualists, for their part, are often committed to strong methodological individualism and regard institutions as social wholes that require endogenisation in terms of individuals, although some do not go so far. Regardless, most

philosophical commentators, and some practitioners, have come to recognise that a strong form of methodological individualism that seeks to entirely endogenise institutions is not achievable. Geoffrey Hodgson (2007), reflecting a view already expressed by Joseph Agassi (1960, 1975), Alexander Field (1984), Douglas North (1990), Malcolm Rutherford (1994) and Kenneth Arrow (1994), put forward a ‘folk theorem’ on this matter: “when explanations are reduced to individuals, interactive relations [i.e. institutions] between individuals are also always involved” (ibid., 217). He explains...

“All versions of social contract theory and general equilibrium theory involve individuals communicating ... or ... adopting tacit presumptions of the intentions and stances of others. ... Exchange involves the transfer of property rights, with rules established through prior social interactions. ... Hence social contract theory and general equilibrium theory both presume structured relations between individuals, rather than individuals in isolation. ... These things may not be stated explicitly in the models, but they are presupposed.”

Hodgson 2007, pp. 217-8

For instance, although Menger sought to explain social institutions as the unintended consequences of otherwise rational actions by groups of individuals in a fashion similar to the classical economists, the individuals in his analysis remain acculturated. His explanation of the emergence of money is another paradigmatic example of methodological individualism, but it still assumes the existence of some institutions. On this account, money is an institution that enables exchange between goods by providing a way of storing value. Prior to the development of money, exchanges require a coincidence of the value of the goods to be exchanged in any given barter. Money develops out of ordinary consumer goods, as one particular good become recognised as a good unit of account and exchange, in the fashion in which cigarettes are often used as money in prisons. Eventually the good in question is typically replaced by the metal and paper we use today. This arises out of the choices made by rational individuals, but many social institutions are also presupposed, not least property rights.

Despite his early insistence that in General Equilibrium Theory all institutions were endogenised, Arrow came to recognise the social institutions that remained exogenous within the theory. While accepting that many of the factors involved in the theory – tastes, risk preferences – are certainly to be considered properties of individuals, he argued that “there is still one element not individual: namely, the prices faced by the firms and individuals. What individual has chosen prices? In the formal theory, at least, no one. They are determined on (not by) social institutions known as markets, which equate supply and demand” (Arrow 1994, pp. 4). He concludes that “individual behavior is always mediated

by social relations. These are as much a part of the description of reality as is individual behavior.” (ibid., pp. 5).

Moreover, a particular form of market is presupposed, and a highly elaborate and unimaginable marketplace it is too. A furiously busy individual dashes round to record everyone’s demand functions and endowments, and facilitates an orderly exchange. Although Udehn argues that the intention behind the creation of this omniscient figure was “clearly to rid the economic theory of general equilibrium of social relations and of all social institutions” (Udehn 2002, pp. 483), and he is correct in asserting so, this ambition was never fulfilled. The Walrasian auctioneer mediates individuals, acting as an institution.

Practitioners within New Institutional Economics have come to similar conclusions. The method of analysing institutions as if they were selected to maximise output net of costs ultimately faces a similar problem to that facing General Equilibrium Theory. Just as questions must be raised about the Walrasian auctioneer, one is left wondering exactly what brings about the maximisation process; how is it coordinated? Many New Institutional Economists turned to Game Theory to explain how institutions emerge out of an exogenous set up free of institutions. Andrew Schotter (1981) explained how cooperation might emerge even in a prisoner’s dilemma, where in a one-shot game individuals have an incentive to defect, if the game is played infinitely or indefinitely. Over this horizon agents can effectively punish defection by withdrawing their own cooperation. Robert Axelrod’s computational experiments later demonstrated how ‘tit for tat’ strategies could bring about cooperation in an otherwise uncooperative dilemma (Axelrod 1984). The difficulty here is that game theoretic approaches do not succeed in entirely endogenising institutions. In fact, the ‘rules of the game’ being played are imposed on the agents. Within a repeated prisoners dilemma, no party can exit the game or steal from the other agents. According to Field (1984), game theorists are prone to losing sight of what is presupposed – “the arena in which the players are to compete or cooperate” (Field 1984, pp. 699).

Douglass North recognised this too, and came to stress the importance of ideology in driving social change. North found it impossible to provide a rational choice explanation in those cases where individuals act against their own self-interest in order to change institutions. For instance, he argues he knows of “no way to explain the demise of slavery in the nineteenth century that does not take into account the changing perception of the legitimacy of one person owning another” (North 1990, pp. 24). North ultimately believes the attempt to entirely endogenise social institutions is futile: “change and stability in history require a theory of ideology to account for these deviations from the individualistic rational calculus of neoclassical theory” (North 1981, pp. 12).

Hodgson’s ‘folk theorem’ – that institutions are necessarily exogenously part of any individualistic explanation – appears to be the case then. Methodological individualism seems to be an illusion and a misnomer. Many commentators have concluded as such, and argued that the actual

method followed by most in the social sciences is in fact institutional individualism (Agassi 1960, 1975), a considerably weaker doctrine under which both individuals and institutions are involved.

We may define institutional individualism as follows.

***Institutional Individualism:** The explanation of social outcomes by reference to individuals and social relations (in institutions) or their properties. i.e. I(individuals) + R(elations) -> O(utcome)*

Institutional individualism is a broad church, accommodating authors whose ‘individualism’ and ‘holism’ are often thought to be in conflict. Nonetheless, institutional individualism does impose some limitations. Holist approaches such as that conducted by Kuznets, Goodwin, Ayres and Vining, where macro-level phenomena are analysed as forming a complete ‘causal nexus’ ungrounded from individual actions, are assuredly outside of institutional individualism. It is also assuredly opposed to those holists who give agency, aims or interests to social wholes, unless these have been decided upon by those who make up the group. Accordingly, one can define methodological holism as follows. This definition excludes authors such as Veblen (at least in his methodological writing), who might otherwise be called ‘holist’.

***Methodological holism:** The explanation of social phenomena by reference to social wholes (aggregates of various kinds) or their properties. Holism encompasses functionalism, and macro-causal theorising that often includes the construction of identities.*

Institutional individualism seems to be accepted by many as being the end of the story for the ‘strong’ methodological individualism that neoclassical economists have argued justifies their method against institutionalist alternatives. Institutionalists have supported it insofar as it directs one to recognise the importance of institutions in explanations, even ‘individualistic’ ones (Hodgson 2007), and so suggests that there is no principled basis on which their analysis is deficient, provided it is demonstrated how institutions act through individuals.

I want to argue that institutional individualism is not the end of the story however. In fact, while it is true to say that strong methodological individualism has never been carried out, institutional individualism does mask some important differences in methodology between institutionalism and individualism. For even though institutionalism and individualism may be united under the label institutional individualism, it remains true to say that individualists focus on the role of individuals in driving social change while institutionalists focus on the role of institutions. As I will shortly show, this focus is a *causal* focus. To demonstrate this, it is helpful to draw on some philosophy of causality.

2.4 Context dependent causality

The common sense view of causality is that it is a binary relation between cause and effect: ‘c causes e’. With this binary view of causality, anything that it is necessary to assume in order to explain a phenomena is a ‘cause’. The binary view of causality is widely rejected by most recent philosophers of science however, because it gives rise to a number of paradoxes that do not sit comfortably with our intuitions about causal claims. The rough nature of these paradoxes can be indicated by an example I take from Asbjorn Stelich-Petersen (2012) and slightly modify. It highlights how changes in the emphasis put on parts of a causal claim give rise to different meanings. Consider the following causal claims:

- (1) Susan *stealing* the bicycle caused her arrest.
- (2) Susan stealing the *bicycle* caused her arrest.

The stresses put upon different parts of the sentence direct our attention to what is felt to be causally relevant. In (1), the stress on *stealing* indicates that it was the fact that Susan stole something that she should have paid for that caused her arrest. That it happened to be a bicycle appears to us somewhat irrelevant. In (2), the stress on *bicycle* indicates that it was the fact that Susan stole a *bicycle* rather than something else that led to her arrest. Perhaps Susan is an accomplished and successful thief of other things – clothes for instance – and it was the fact that she tried to steal something else that she was not practiced at that led to her downfall.

The lesson from examples such as these is that statements such as (1) and (2), which posit a binary relation between cause (Susan stealing the bicycle) and effect (Susan’s arrest), fail to unambiguously pick out causes. Subtle differences in emphasis direct us toward aspects of the ‘cause’ which are of greater significance somehow.

A related problem is that causal judgements seem to vary depending on the enquirer. Absent any clues in emphasis that direct us to a pick out a particular aspect as being the most causally significant part of ‘the cause’, it seems that different enquirers bring with them their own expectations and interests. A well-known example from H. L. A. Hart and Tony Honoré (1985) demonstrates the issue well. They discuss the possible causes of a great famine in India, and note how causal judgements are different for Indian peasants and official authorities. Indian peasants, who will be keenly aware of how weather conditions affect crops, typically locate drought as the cause of famines. The World Food Authority (WFA), on the other hand, takes a different perspective, more long-term in nature, and more focused on the possible policy responses to drought conditions. From the FWA’s perspective, droughts are a phenomenon that governments should expect and manage. The WFA is accordingly led

to criticise the government for failing to build up food reserves, locating their failure to do so as the ‘cause’. So, which is *the* cause, the famine or the government? The answer seems to depend on who is asking the question.

These examples give just a basic indication of one of the many paradoxes that the binary account of causation gives rise to, which we may call the *paradox of selection*. Many other paradoxes arise from a binary account, which expose troubling and counterintuitive consequences of a binary account⁹. The paradox of selection is troubling to many philosophers because it apparently makes causality context and mind-dependent - dependent upon the observer and causal enquirer. This threatens a causal realist view, under which causality is objective matter and causal claims have context-independent truth-values.

One popular response to this issue is to argue that while causal explanations are context sensitive, and depend on particular causal questions, the actual causal relations involved in any given situation remain objective. This is to say that there is network of causes which lead up to any event, but that when asked to supply an answer to the question “what is the cause of E” different answers may be supplied, even though we can agree on the objective causal relations. In the famine example, peasants and the WFA may well ultimately agree on the causal chain that led to the famine, even though their initial thinking when asked to isolate ‘*the* cause’ varies. One way of exposing how different judgements arise in this case is to distinguish between whether one is asking “Why did the famine occur at this time rather than another?” or if the question is instead “Why did the famine occur in India rather than another country that similarly suffered drought?”. In answering the first question, the fact that the government had not stock piled reserves is a permanent condition and so is not treated as causally relevant. Rather it is the change in the weather conditions that is noted as the cause. In response to the second question, India’s susceptibility to drought is already considered and a contrast is made to other countries. In doing so, government policy becomes the distinguishing factor.

David Lewis noted the selection problem, but rather dismissed the issue as simply a matter resolved by *pragmatic* considerations of the causal enquiry at hand.

‘We sometimes single out one among all the causes of some event and call it ‘the’ cause, as if there were no others. Or we single out a few as the ‘causes’, calling the rest mere ‘causal factors’ or ‘causal conditions’.... We may select the abnormal or extraordinary causes, or those under human control, or those we deem good or bad, or just those we

⁹ Jonathan Schaffer (2005) provides detail of several of these paradoxes, include the paradox of absences (do absences count as causes?), of fragility (how modally fragile are causal events?), extensionality (can causal claims change truth-value when event descriptions are substituted?), transitivity (is causality transitive?) as well as selection (is selection of the ‘cause’ objective?).

want to talk about. I have nothing to say about these principles of invidious discrimination.'

Lewis 1973, pp. 559

Lewis (1973) then went on to advance a counterfactual analysis of causation, with an understanding of causality as binary relation. The counterfactuals are provided by the pragmatics of the causal enquiry, with causality ascertained thereupon, but their selection is not analysed any further back. But this would seem rather too easy. It may be claimed that the causal network about which causal claims are made is objective, but how can we know this? If causal talk is unavoidably tied to a particular perspective, how can we speak of an objective causal network?

Another response is to argue that examples such as these only serve to illustrate ambiguity in the *semantics* of causality. It may be argued that there is no single concept of causality and that in each different context different meanings apply to situations that nonetheless have context-invariant truth conditions. Christopher Hitchcock (2003) has put forward this position, arguing that philosophers have been on the wrong track when attempting to uncover a single meaning of "Do events C and E stand in *the* causal relation" (*ibid.*, pp. 21). Instead, he suggests that there are several possible causal relations: "C belongs to a causal chain of events leading up to E; C has a component effect on E along some particular causal route; C has a net effect on E when all causal routes from C to E are taken into consideration; C is a cause of E on average in some contingently constituted population; C is a cause of E as a matter of causal law; C is a cause of E relative to some particular range of alternatives of domain of variation" (*ibid.*, pp. 21-22). This is to say that the concept of causality itself is *ambiguous*, but that different unambiguous varieties of causality can be ascertained. The paradox of selection is averted by arguing that asking after *the* cause is an improper question. There is not a single cause, but there may be many depending on the kind of causality one has in mind.

Another response is to transfer the context-sensitivity of causal enquiries into the semantics of causality itself, and offer an account of this. The account offered may or may not be claimed as covering all causal talk, or simply a kind of causality (as Hitchcock argued). Jonathan Schaffer (2005), for instance, has argued that causality is fundamentally *contrastive* in nature. As usual, the causal relatedness of events is evidenced as a matter of counterfactual dependence. According to Schaffer, causality is not simply a binary relation between cause and effect, but rather a *quaternary* relation that utilises contrasts at both ends. The general form is: "c rather than c* causes e rather than e*". The structure of the causal enquiry involves three parts that serve to fix the relevant contrasts: (i) the background circumstances, (ii) the causal options, $c_1 \dots c_j$ where $c_i \in C$ and C is the set of possible events that might occur, and (iii) the effectual options, $e_1 \dots e_j$ where $e_i \in E$ and E is the set of possible alternative events (Schaffer (2005, pp. 348). This is a view in keeping with John Mackie's, whereby a

“causal statement will be the answer to a causal question... where both causes and effects are seen as differences within a field” (Mackie 1974, pp. 34-35).

This account can be readily applied to the preceding examples. The subtleties of emphasis in (1) and (2) can be picked apart by supplying the relevant contrasts. In (1) the emphasis directs us to contrast between a situation where Susan steals the bicycle and a situation where Susan pays for it. In (2) the emphasis directs us to contrast between Susan stealing a bicycle and Susan stealing something else. (1) and (2) are made unambiguous by supplying the relevant causal contrasts at both ends.

(1*) Susan stealing the bicycle rather than purchasing it caused her arrest.

(2*) Susan stealing the bicycle rather than stealing clothes caused her arrest.

(1*) and (2*) express (1) and (2) precisely and without ambiguity by making clear each element of the quaternary relation of causality. Somehow, these contrasts arise out of the background contexts of each case. The background context creates the scenario in which these possibilities are in place.

Likewise, the causal claims about the Indian famine can be distinguished by articulating them in this quaternary structure. In the peasants evaluation, the drought is the cause of the famine and the failure of the government to stockpile food reserves is a mere background condition. This is because the relevant contrast the peasant has in mind is not whether or not the government had food reserves, but whether or not the weather has been favourable. The World Food Authority considers a different contrast however; from this perspective drought is a background condition and the relevant causal factor is whether or not the government has built up food stocks. In both cases the same regularities, counterfactuals and probabilistic dependencies hold, but different causal judgements arise because of differences in the context considered.

These responses to the paradox of selection all take it as accepted that causal judgements are context sensitive. The only difference between these accounts is where the context sensitivity occurs, either: (i) as a pragmatic consideration of little philosophical interest, or (ii) as fully part of the semantics of causal talk. Either way, it is accepted that the relevant contrasts are supplied by contextual considerations.

At this point it would be agreeable to be able to flesh out some principles for the ‘invidious selection’ of contexts and causes. *Prima facie*, there are some good reasons for thinking that some principles might be forthcoming; we often tend agree on what counts as ‘the cause’ and what are the ‘background conditions’ in a way that seems like more than mere coincidence. For example, if James does not stub his cigarette correctly before going to bed, and there is a fire in his flat that begins with this heat source, most people will agree that James’s negligence was ‘the cause’ of the fire, and understand the host of other conditions (oxygen, fuel, etc.) as being mere ‘background’.

Peter Menzies (2007) argues that in determining causes, we first try and conceptualise situations as having a normal course of evolution, which the cause interrupts. “Default worlds” are imagined that generate a model for the system under study. Similarly, Tim Maudlin (2004) writes:

“In judging causes, we try to carve up the situation into systems that can be assigned inertial behavior (behavior that can be expected if nothing interferes) along with at least a partial specification of the sort of things that can disturb the inertial behavior, analogous to the Newtonian forces that disturb inertial motion. Let us call the things that can disturb inertial behavior ‘threats’: they are objects or events that have the power – they interact in the right way – to deflect the system from its inertial trajectory. We then think about how the situation will evolve by expecting inertial behavior unless there is interaction with a threat, in which case we see how the threat will change behavior.”

Maudlin 2004, pp. 436

Beyond this rough characterisation however, it is very difficult to articulate positive characteristics for how default worlds and ‘threats’ are identified. Menzies (2007) put forward a few suggestions, although he admits that they are only ‘rough and ready’ generalisations. He observes how “known laws or regularities clearly influence the expectations of what is the normal course of evolution for a system of the given kind” (Menzies 2007, pp. 222). This applies just as much in everyday as scientific situations: “if every car passing down the street in front of my house has been reasonably quiet, then the expectation based on that regularity will determine what I think calls for explanation. A car backfiring, for example, will be something that is anomalous with respect to that expectation” (Menzies 2007, pp. 222). So prior beliefs and expectations are likely to play an important role; deep seated convictions about the ‘normal course of events’ that may either come from everyday common sense or sophisticated scientific knowledge influence what is examined as causal and what is treated as mere background. Beyond this rather basic characterisation however, there is little in the philosophical literature that is much guidance here.

The philosophical accounts of causality that I know of can all be seen as involving ‘default worlds’ and ‘threats’ in one way or another. One obvious way in which threats can come about is through human intervention. Peter Menzies and Huw Price (1993) developed just such an ‘agency’ theory of causation: “an event A is a cause of a distinct event B just in case bringing about the occurrence of A would be an effective means by which a free agent could bring about the occurrence

of B” (*ibid.*, pp. 187). This provides an obvious sense in which ‘threats’ arise, i.e. through human interventions on systems that otherwise proceed naturally¹⁰.

James Woodward’s *interventionist* account of causality (Woodward 2003, etc.) operates with similar manipulationist ideas, except his account is intended to apply to situations where causal dependencies hold without human intervention (the understanding offered by examples involving human intervention are only heuristic). Under his account, an intervention is an “exogenous causal process that brings about the antecedent of the counterfactual in question” (Woodward 2003, pp. 9), and may not be the result of human agency. Other, objective, criteria for interventions are provided. One of the crucial aspects of an intervention is that it is ‘surgical’; interventions must alter only the cause directly while not interfering with other causal mechanisms so as not to alter the effect except through the intervention on the cause. The notion of an intervention is thus relativized to a background context; its definition requires specifying a normal course of events.

Neither an agential or interventionist view provide any guidance on how threats and default worlds (or ‘causes’ and ‘background conditions’, ‘interventions’ and a ‘causal model’) are selected. Principles for the selection of causes and background conditions are hard to come by.

To conclude this excursion into the philosophy of causality then, the essential point that is pertinent to this thesis is that causal claims are necessarily tied to background contexts. On this point there is widespread agreement among philosophers. Quite how this is dealt with philosophically is not especially important for my purposes¹¹. Whether one tries to accommodate the context sensitivity of causality within a binary account of causality, and account for context sensitivity as purely a pragmatic consideration (Lewis), or whether one factors context sensitivity into the semantics of causality itself (Schaffer, Maudlin, Menzies), whether one believes that this issue makes causality a fundamentally subjective matter or if one is a causal realist, the essential point that causes are identified with respect to some background considerations is widely accepted and plays an important role in contemporary views of causality. Quite how context and causes are selected remains quite mysterious, although Menzies and Maudlin’s suggestions that selection has something to do with our expectations of the normal course of events seems like a reasonable starting point.

¹⁰ Such agency theories of causality are questionable in how they can deal with cases where there is no obvious means for human intervention (e.g. planetary motion, the eruption of volcanoes, the weather), but where causal claims seem nonetheless well founded. Price and Menzies accommodate these examples by arguing that provided that there are other situations that are sufficiently similar in which human interventions are possible – in artificial models, for example – then causality can be ascribed based upon the potential for intervention that the similarity affords. “We can make such causal claims [about earthquakes] because we believe that there is another situation that models the circumstances surrounding the earthquake in the essential respects and does support a means-end relation between appropriate pairs of events. The paradigm example would be that created by seismologists in their artificial simulations of the movements of continental plates” (*ibid.*, pp. 197). There are difficulties in this view however, insofar as the similarity between actual and model worlds that Price and Menzies rely on requires specifying in non-causal terms in order for this understanding of causality to be non-circular, but that seems unlikely.

¹¹ Schaffer concluded: “It does not matter here whether the contrasts are to be factored into the semantics or shunted into the pragmatics. Either way, the contrasts must be generated from the context. But I don’t know how.” (Schaffer 2005, pp. 350).

2.5 Causal individualism, institutionalism and holism

This is all of utmost relevance for the debate over the kind of methodological individualism involved in neoclassical economics. For the realisation that causal claims are always made with respect to some background conditions makes the observations by Arrow, Rutherford, Hodgson and others that individualist explanations always also involve assumptions about institutions a less surprising critique. The criticism of strong methodological individualism stems from a binary understanding of causality; it is argued that since institutions are also exogenous terms, they must be causes too. But causal claims are more than simply binary relations between cause and effect; they involve counterfactual comparisons relative to a background context. While it is absolutely the case that individualists also always make assumptions about institutions, these assumptions form part of the background context that is being assumed, and are not singled out as being ‘the cause(s)’. That role is preserved for properties of individuals. This is really the crux of the matter. *Individualist and institutionalist explanations can be distinguished by whether or not the causal claims they make take individuals or institutions as the causes respectively. Individualist (institutionalist) explanations’ assumptions about institutions (individuals) are only part of the background circumstances of their analyses.* This simple idea can be fleshed out more clearly with the help of some of the preceding philosophy of causality and applying it to a paradigmatic example of neoclassical economics.

Those philosophical accounts that transfer the context dependence of causality directly in to the semantics of causality (rather than simply assuming contextual clues are delivered by ‘invidious’ pragmatics) are more assistance as they provide frameworks to discuss these issues directly (either in terms of default worlds and threats or in terms of an understanding of causality as a quaternary relation). Admittedly, this assistance provides only relatively scant pickings as there is little in the literature that has satisfactorily articulated how causes and background conditions are selected. I shall use Schaffer (2005)’s account of causality as being a *quaternary, contrastive* relation, as this at least allows one to see the work done by the truly causal aspects, even if it remains unclear exactly how the causes are selected.

According to Schaffer, causality is indicated by demonstrating counterfactual dependence between phenomena by showing how ‘c rather than c* causes e rather than e*’. Accordingly *causal individualism* can be defined as being a description for explanations that involve demonstrating how different values for the properties of individuals – preferences, endowments of wealth, capital, income, technology, educational ability, etc. - generate different social outcomes.

Examples abound within neoclassical economics. General choice theory provides a framework for explaining observed differences in the slopes of market demand curves as being determined by differences in consumer preference orderings, assuming an institutional set up that aggregates individual demand curves unproblematically and brings markets into equilibrium. In game theory, the familiar 2x2 matrix encodes background conditions of interaction between individuals but yields

different outcomes depending on the payoffs received by individuals, or the utility functions that individuals have to evaluate those payoffs. In expected utility theory, background assumptions about interactions are made but the different outcomes examined ultimately arise from changes in individuals' attitudes toward risk.

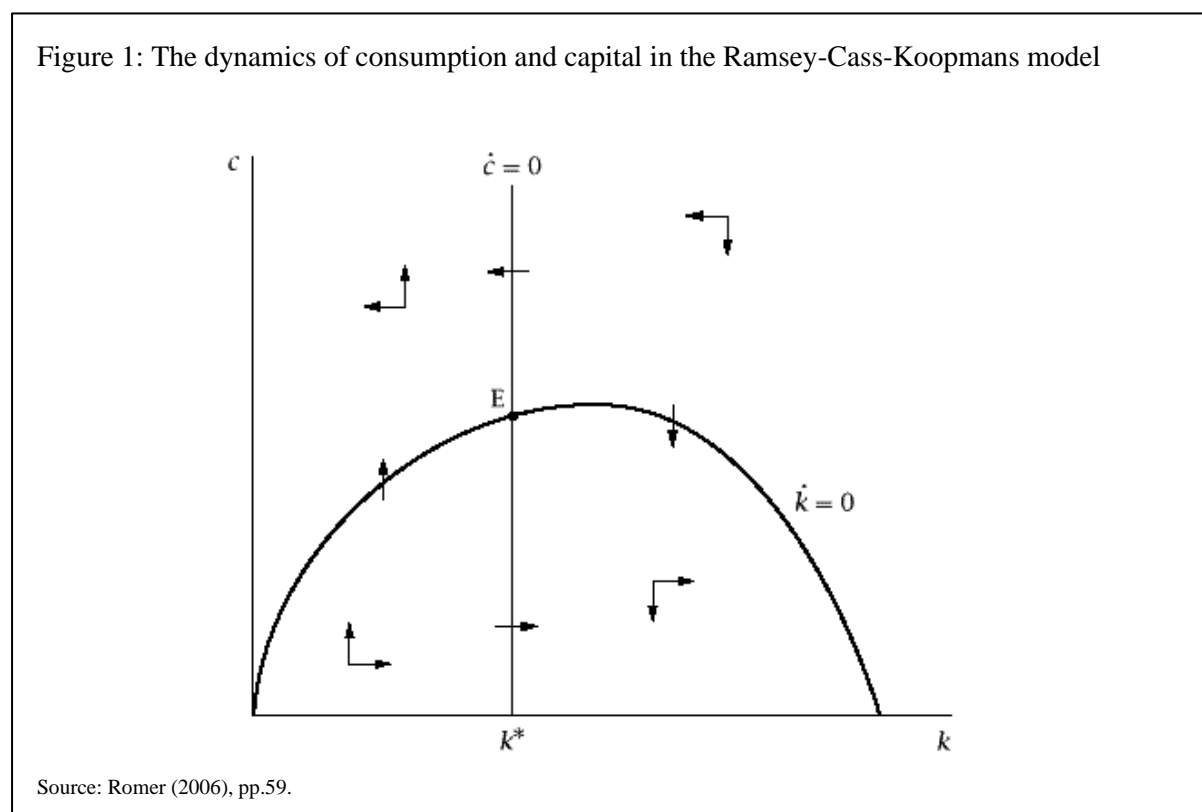
Let me demonstrate this for an example in some detail. The Ramsey-Cass-Koopmans overlapping generations model of economic growth is something of a corner stone in neoclassical growth theory. Initially developed by Ramsey (1928) and extended by Cass (1965) and Koopmans (1965), this model analyses the performance of the economy as the outcome of rational behaviour by utility maximising individuals. Its use in modern economics is widespread because it is a very flexible, simple model that can be readily adapted to a wide range of different circumstance. The model has some similarities to the Solow growth model in that the growth rates of labour and technology are exogenous, but departs from the Solow model as the saving rate is endogenised as the outcome of the interaction between maximising households and firms operating in competitive markets.

In the basic Ramsey-Cass-Koopmans model there are firms and households that take an infinite horizon. There are a large number of identical firms, each with a production function of $Y = f(K, AL)$, where Y is output, K capital, L labour, and A the efficiency of labour. They hire workers and capital in competitive markets. A grows at a constant rate, g , driven exogenously by technological progress. There are also a large number of identical households, with the size of each household growing at rate n (population growth), with each member providing one unit of labour in each time period. Each household has a utility function as follows: $U = \int_{t=0}^{\infty} e^{-\rho t} u(C(t)) \frac{L(t)}{H} dt$, where $C(t)$ is consumption of each household member, $u()$ is the instantaneous utility function, $L(t)$ is the total population, H is the number of households and ρ is the discount rate. It is assumed that the instantaneous utility function displays constant-relative-risk aversion (CRRA) and is of the form $u(C(t)) = \frac{C(t)^{1-\theta}}{1-\theta}$, $\theta > 0$, $\rho - n - (1 - \theta)g > 0$. This assumption assures that the economy converges on a balanced growth path. Households maximise this utility function. Firms are in perfectly competitive markets and employ labour and capital at their marginal products. It is assumed that there is no depreciation so the return on capital is the interest rate, $r = f'(k(t))$, where $f'(x)$ is the first derivative of $f(x)$. A final crucial assumption is that there is not a possibility to engage in Ponzi schemes, whereby debt is rolled over for ever. This is captured by assuming that $\lim_{s \rightarrow \infty} e^{-R(s)} \frac{K(s)}{H} \geq 0$.

The most convenient way to represent the evolution of the model is to examine the dynamics of consumption and capital in a phase diagram that demonstrates how the economy moves towards equilibrium. The dynamics of consumption are captured by the Euler equation resulting from the households' maximisation problem, $\frac{\dot{C}(t)}{C(t)} = \frac{r(t) - \rho}{\theta}$, into which is substituted the return on capital, which is equal to the interest rate in perfectly competitive markets, i.e. $\frac{\dot{C}(t)}{C(t)} = \frac{f'(k(t)) - \rho}{\theta}$, where $\frac{\dot{C}(t)}{C(t)}$ is

the growth rate of consumption. Since all households are assumed to be identical, this equation also represents the level of growth of consumption per capita, c - which is really what is of interest. Capital stock growth per capita, k , is produced from investment (output less consumption) less the break-even level of investment (that required to provide a stable level of capital to each worker), i.e. $\dot{k}(t) = f(k(t)) - c(t) - (n + g)k(t)$ (as per the Solow growth model without depreciation).

The phase diagram captures the interaction of firms and households and indicates how the economy moves from various starting values of consumption per capita and capital per capita (see figure 1 below)¹².

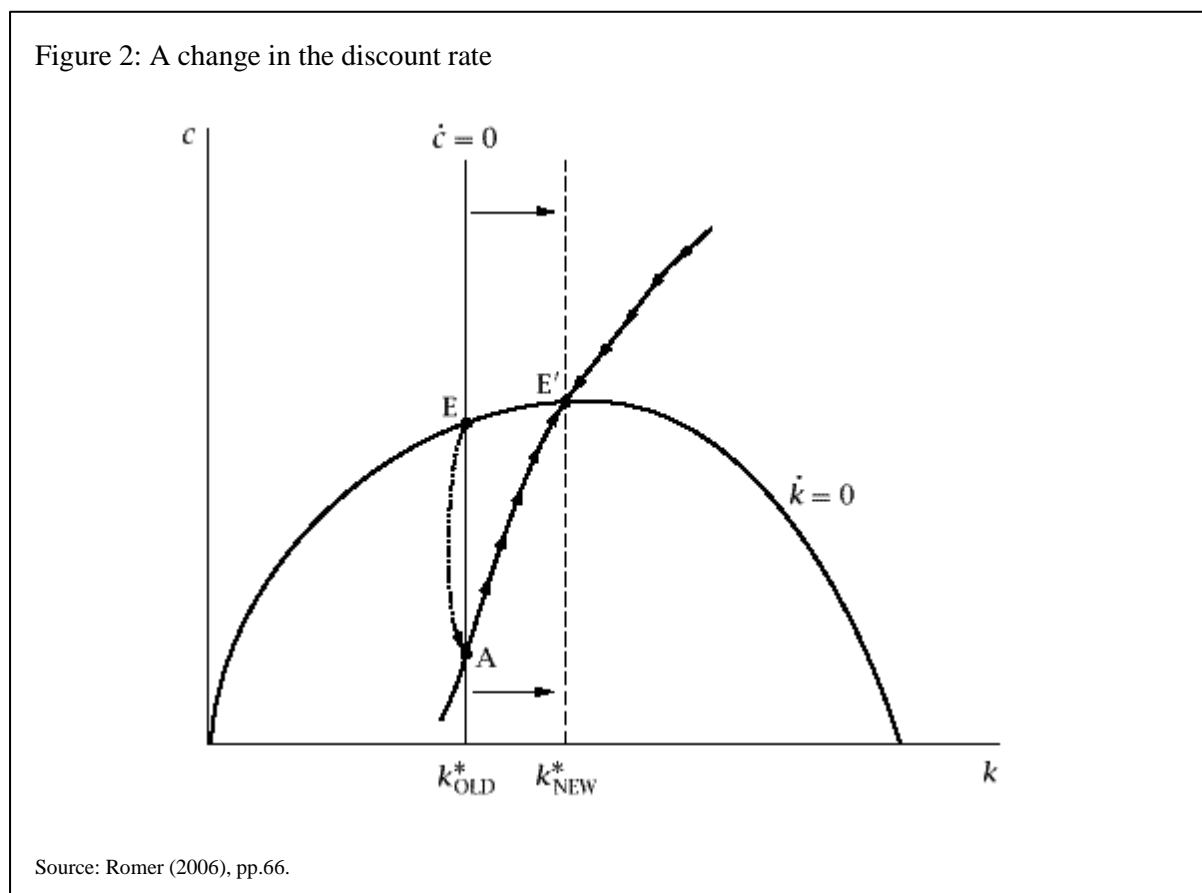


This model forms the backbone of much contemporary macroeconomic theory and new classical macroeconomics, and is viewed as having microfoundations via its use of representative agents¹³. The model is *causally individualistic*, insofar as the exogenous model *parameters* all pertain to individuals – their time preference represented by rate at which they discount future consumption, ρ , their number (the population, n) and their ‘efficiency’ growth, g . The optimal savings rate for the economy, and the

¹² This diagram is taken from Romer (2006) chapter 2, where a more complete exposition of the model with full derivation of the Euler equation can be found.

¹³ Interestingly, this was not Ramsey’s original interpretation. He understood the utility and production functions as pertaining to aggregates and instead addressed the problem from the perspective of a social planner (see Gaspard 2003). Regardless, the modern interpretation is one of representative agents, and so the model is intended to capture the evolution of the economy as the outcome of decisions by a large number of identical agents.

evolution of consumption and investment that follows are part of the endogenous solution of the model. Accordingly, it is commonplace to examine how differences in the exogenous parameters influence the savings rate the economy's equilibrium growth path. For example, the effect of individual's discount rate is demonstrated below (figure 2).



It is on this basis that individualistic causal claims are made. Various assumptions are made about institutions too: property rights are in place, as are numerous social institutions concerning firms, labour and money. These are not the focus on the analysis however. They are merely background conditions. The properties of individuals do the causal work; as per Schaffer's contrastive account of causality, differences in the values of the properties of individuals are causally linked to different social outcomes. The Ramsey-Cass-Koopmans model is an exemplary piece of causal individualism: individuals perform the properly causal work, while institutions are mere background conditions. It is in this sense that individuals are given primacy under causal individualism, even if this primacy is not of an absolute kind where by institutions are entirely endogenised as per the intention under strong methodological individualism.

2.6 Definitions

I am now in a position to define *causal individualism*, *institutionalism* and *holism*, based upon the idea that these different approaches cite different types of causes.

Causal individualism: *The explanation of social phenomena using causal claims that appeal to counterfactual differences in individual properties (preferences, utility, wealth, income, capital assets, etc.). Causal claims are made of the form, $d(I|R) \rightarrow d(O)$ (i.e. difference in individuals, given some assumptions about social relations, lead to differences in social outcomes).*

Causal institutionalism: *The explanation of social phenomena using causal claims that appeal to counterfactual differences in institutions (conditions for the interactions of individuals that instruct ‘in circumstance X, do Y’). Causal claims are made of the form, $d(R|I) \rightarrow d(O)$ (i.e. differences in institutions, given some assumptions about individuals, lead to differences in social outcomes).*

Causal holism: *The explanation of social phenomena using causal claims that appeal to counterfactual differences in the properties of social wholes (meaning groups and aggregates of various kinds, included classes, groups and organisations).*

One important point to clarify about these definitions concerns whether or not they require that explanations *exclusively* appeal to the relevant category of counterfactual differences, i.e. whether causal individualism (institutionalism/holism) means that any given explanation must *only* involve counterfactual differences in the properties of individuals (institutions/social wholes). I only wish to give these new categories descriptive content; on their own they do not enforce any such restrictions. In principle it is possible to be causally individualist and causally institutionalist at the same time, if the causal claim made involves both changes in individuals and institutions in the antecedent of the causal claim at the same time. One can expect though that this will be quite rare, as this kind of causal claim will be quite ambiguous as to which of the causes is really doing the causal work. One can expect that the authors of causal claims will try to avoid any such ambiguity, perhaps by varying each component separately and observing the resultant outcome (either through the application of a theory, or empirically by cutting the data in the relevant way or controlling for each effect separately in an econometric model, say).

2.7 Are institutions properties of individuals or properties of social wholes?

A further point of clarification is required concerning the status of institutions; are they properties of social wholes or properties of individuals? The answer here has to be nuanced. As discussed, they have been treated by institutional economists in both ways; some like Veblen sought to show how they act through individuals, while others such as Galbraith examined the interactions of groups such that they follow institutional imperatives collectively. On the one hand, institutions can be considered as a property of groups that dominate individuals; when their effect is examined, it is assumed that some particular group of individuals collectively follow some institutional imperative. Here institutional rules seem to dominate and dictate the action of groups of individuals; the ‘whole explains the individual’. It is in this sense that neoclassical economists have understood their aim of explaining the ‘social whole’ that is any given institution. On the other hand, institutions may be considered as simply being the rules that a collection of individuals have in common; one doesn’t need to see them as being ‘imposed’ by the whole on individuals at all, but as imperatives that individuals have determined. It seems that institutionalism can be either individualist or holist, in that institutions can be considered as being possessed by individuals or social groups.

What is required here is a better understanding of what distinguishes ‘a property of a social whole’ from ‘a property of individuals’. One can see more clearly what is involved with holism by looking outside of institutions to other basic economic holistic concepts like inequality. The important point about something like inequality is that even though it is calculated from the properties of individuals, it is unavoidably a property of a group. An individual cannot be more or less equal with himself. In philosophical parlance properties like inequality are *emergent* (Bedau 2003). This is the most uncontroversial kind of emergence, nominal emergence, which denotes cases where a macro property is the kind of property that cannot be a micro property. For example, while water can have the property of being liquid, this property cannot be meaningfully applied to its individual components of H₂O. So it is with several macroeconomic concepts.

We can think about institutions in the same way. In arguing that “[a]s one proceeds from the smaller corporations to giants, the role of any single individual diminishes, the authority of organisation increases” (Galbraith 1973, pp. 83), Galbraith is arguing that the behavior of organisations is emergent. It is not that all individuals within an organisation all act the same way; we could not observe all within the organisation doing the same thing, and then treat the macro property as simply an expression of all the micro properties. Rather, it is that organisations can be treated as a single entity which follows its own institutional imperatives. Often organisations do things that are not the kind of thing that an individual can do; facetiously, pay corporation tax.

Institutions may not be thought of as being emergent though; they can be thought of as properties that each member a group of individuals have. In this case, institutions are assumed such that every single individual in the group follows the *same* institutional rule.

This then seems to threaten to fold some institutionalism into individualism and the rest into holism, if we are then calling institutions ‘properties of individuals’ or ‘properties of social wholes’. Certainly we can distinguish between institutionalists who are individualistic in the sense that they show how institutions act *through* individuals and those that are holistic in the sense that they show how institutions act *through* organisations. But it is not correct to then describe institutions as either ‘properties of individuals’ or ‘properties of social wholes’. Institutions are above all social *relations*; they are conditions for the *interactions* of individuals or groups in *particular situations*, not properties of the individuals or groups themselves. Whenever they are spoken of one necessarily also conjures the existence of a social situation. They are not the sort of things that individuals can have in isolation. To talk of a property of something is to talk of something that a thing possesses regardless of the other entities surrounding it. If I have red hair, I have it regardless of others around me. If I have a loaf of bread, I can eat it and derive utility from it and it has nothing to do with anybody else. Individual properties are the kind of things that can be expressed as single variables; I can have more or less wealth, income, different utility values for consuming things and so on. Causal individualism involves showing how the variation in these properties influences social outcomes. But this cannot be done with institutions; they cannot be squashed into a single variable, ‘*x*’, but are unavoidably binary, *social relations*, placing individual action in particular social settings: ‘in circumstance *X*, do *Y*’. Analysis of them does have something unavoidably ‘holistic’ about them, as even if in the individualistic form when one looks to examine their effect one assumes that everyone in a *group* follows them.

Hence institutionalism can operate at either the level of individuals or the level of social groupings, but this does not make institutionalism identical with individualism and holism respectively. To recognise the distinction we can refer to *individualistic institutionalism* or *holistic institutionalism*.

Individualistic institutionalism: *The explanation of social phenomena in terms of individuals and the social relations between them, using causal claims that appeal to counterfactual differences in institutions. Causal claims are made of the form, $d(R/I) \rightarrow d(O)$.*

Holistic institutionalism: *The explanation of social phenomena in terms of groups of individuals and the social relations between them, using causal claims that appeal to counterfactual differences in institutions. Causal claims are made of the form, $d(R/I) \rightarrow d(O)$.*

Notice that individualistic institutionalism is not identical with institutional individualism as defined by Joseph Agassi (otherwise known as weak methodological individualism).

2.8 How neoclassical reductionist views lead to causal individualism

It remains to understand how causal and methodological individualism are related to each other. Since methodological individualism has been understood and argued for in a number of different ways over the years, I will not be able to provide a completely thorough account of how my new categories relate to all the different understandings of methodological individualism out there. Rather my focus will here be on understanding the kind of methodological individualism found in neoclassical economics in terms of my causal categories. I shall ignore other views of methodological individualism (e.g. Schumpeter, von Mises, von Hayek).

Let me recap the ideas at hand. Causal individualism and individualistic causal institutionalism may be considered as sub-categories of institutional individualism. Both causal individualism and individualistic causal institutionalism take both individuals and institutions as exogenous terms – the difference is that in causal individualism the individuals are highlighted as causes while the institutions go to the background, and *mutatis mutandis* for individualistic causal institutionalism. The impossibility of strong methodological individualism discussed earlier suggests that institutional individualism is as strong as methodological individualism gets, but is notable for failing to pick out the differences in methodology between old institutionalist and neoclassical economists. The question then arises: if institutional individualism is all there is, why has neoclassical economics gone down the road of causal individualism and not causal institutionalism? Is this ‘invidious selection’ just a matter of preference?

Recall that in section 2.4 I noted that philosophers have not much of a handle on how causes and background conditions are selected. One suggestion was that causal enquirers bring with them expectations about a ‘normal course of events’ which supplies a ‘default world’ of background conditions. Something similar could be going on here; it could be that neoclassical economists have a tendency to view the world as being one in which market institutions work reasonably well, and so set up a default world with a highly effective Walrasian auctioneer, and then look at the effects of different choices made by individuals – who they allow autonomy and idiosyncrasy. On the other side, it is plausible to think that institutionalists believe that institutions are rather more unstable and liable to change, and the ultimate drivers of social change among individuals who are rather mechanical actors. What guides this choice? Some might well be tempted to say that this kind of invidious selection is precisely what is meant by ideology. This may well have something to it, but I believe there is more to the way neoclassical economists have settled on causal individualism than mere ideology.

The major difference between my causal categories and neoclassical methodological individualism is that my categories are simply descriptive, allowing one to taxonomise social scientists by which type of phenomena they highlight as being causes, whereas methodological individualism is usually understood by neoclassical economists to have normative force, instructing

that explanations of social phenomena *should* be in terms of individuals (and institutions)¹⁴. Another major difference is that my definitions involve a different view on scientific explanation than is usually assumed by methodological individualists, which is best understood as a “species of reductionism” (Udehn 2001, pp. 329). It is here that the normative element is introduced: understanding how methodological individualism has been intimately related to the idea that science proceeds by reduction holds the key to understanding its normative force for neoclassical economists. It also reveals why the kind of methodological individualism involved in neoclassical economics leads to causal individualism rather than causal institutionalism. So let me now examine the normative aspects of methodological individualism.

There are two possible interpretations of the normative force of methodological individualism. The strongest interpretation entails that *all* explanations must be in terms of individuals (and institutions). But methodological individualism has also sometimes been understood more weakly as the idea that explanations must be ‘in principle’ expressible in terms of individuals. This position accepts that regularities between social wholes might be forthcoming, and even that they have some explanatory import, but that they are only ‘half-way’ explanations, as John Watkins put it (Watkins 1957, pp. 107). In this weaker view, holist science provides some knowledge, but it is only partial. Providing an account of how it comes about in terms of its components adds to our knowledge to make a ‘full’ explanation.

But I do not think that this normatively weak version of methodological individualism says very much and fails to capture what methodological individualism is really all about for the neoclassical orthodoxy. This notion of methodological individualism is so weak as to be without much real significance, and is probably impossible to argue against. As Harold Kincaid put it, “[p]resumably for every entity with composing elements, there is always something that can in principle be said about its micro-structure that adds somewhat to our comprehension” (Kincaid 1986, pp. 510-511). And more to the point, one need only consider how ideas about methodological individualism have served to demarcate orthodoxy from heterodoxy to realise that methodological individualism has been stronger than this. Such a reading of methodological individualism would mean that traditional macroeconomics that sought no microfoundations would be considered with greater respect than it has been over the last thirty years or so. On this reading, macroeconomics that did not supply microfoundations would not be seen as deficient or fatally flawed, but rather as providing useful ‘half way’ contributions, that might be added to with microfoundations. But, mainly since the Lucas critique, this has not been the attitude of mainstream economics. If it were, a figure like Dean Baker – who forewarned of the financial crisis on the basis of deviations of key macroeconomics indicators

¹⁴ Recall the quotation opening this chapter from Paul Krugman, who believes individualism is “of the essence” (Krugman 1996b).

from historical long-run norms – would not be considered a heterodox economist and his warnings might have been listened to rather than ignored. Some might regard methodological individualism as only having this weaker normative force, but this is clearly not the understanding of orthodox neoclassical economists. Methodological individualism has been a stronger thesis than this weak interpretation, revealed by the attitude the orthodoxy has had towards macroeconomics that did not provide microfoundations.

To understand the kind of normative force that neoclassical methodological individualism involves, it is necessary to look at the understanding of scientific explanation that is assumed. Many advocates of methodological individualism have viewed it as a reductionist thesis and it is this that generates the normativity. Jon Elster was explicit about his reductionist inclination in defending methodological individualism, defining it as “the doctrine that all social phenomena - their structure and their change - are in principle explicable in ways that only involve individuals - their properties, their goals, their beliefs and their actions. Methodological individualism thus conceived is a form of reductionism” (Elster, 1985:5). Similarly, Frank Hahn proudly declared “I am a reductionist in that I attempt to locate explanations in the actions of individual agents (Hahn 1984, pp. 1). David Mellor made the argument even more explicit: group laws “relate attributes supervenient on its members’ actions and attitudes. The [group] law ... must therefore be derivable from some true explanatory psychological theory” (Mellor 1982, pp. 70). On this reductionist view, social phenomena are only ‘epiphenomenal’, while the ‘real’ entities are the individuals. To this day, some economists deny the possibility of macroeconomics on the basis that ‘it does not exist’¹⁵.

But reductionism has been understood by economists in a rather strange way. Reductionism is typically understood as the relation between two theories (Nagel 1961)¹⁶; as Mellor put it, it concerns the reduction of ‘group laws’ to ‘psychological theory’. Again this would imply that there was utility to be had in searching for and using group laws. But this is not how economists have usually interpreted reductionism. In fact, many have argued that group laws do not exist, giving rise to a strange kind of reductionism that involves the reduction of *phenomena* rather than theories. Karl Popper put forward this strange version of reductionism in his *Poverty of Historicism* (Popper 1944a/b, 1945). First he explains why macro-level laws do not exist.

¹⁵ See <http://www.youtube.com/watch?v=F5chnCkIguk&feature=related> for a video by an unknown microeconomist under the pseudonym maximizutility, who offers a particularly dogmatic instance of this view: “Economics is the greatest of all sciences. It has two fields: macroeconomics and microeconomics. Macro’s what you see in the press all the time: stories about the central bank lowering rates to stimulate the economy, to stimulate demand, get more GDP growth. This is all nonsense. There is no such thing as a macroeconomy. There is no such thing as GDP”.

¹⁶ Nagellian theoretic reduction involves the deduction of higher level laws from lower level laws, and often involves providing so called ‘bridge laws’ that allow one to define the properties found at the higher level in terms of those found at the lower level. The classic example is the reduction of the gas laws to the laws of statistical thermodynamics.

“[A] combined diagram [showing collective categories] cannot be said to represent the path of the movement of society; it does not tell us more than do the single ones together; it does not represent the movement of ‘the whole society’, but only changes of selected aspects. The idea of the movement of society itself—the idea that society, like a physical body, can move as a whole along a certain path, and in a certain direction is merely a holistic confusion. The hope, more especially, that we may some day find the ‘laws of motion of society’, just as Newton found the laws of motion of physical bodies, is nothing but the result of these misunderstandings. Since there is no motion of society in any sense similar or analogous to the motion of physical bodies, there can be no such laws.”

Popper 1945, pp. 72

Popper then elaborates on this argument against the use of social laws in a more nuanced fashion by distinguishing between laws from mere *trends*. The observation of empirical data to establish historic trends is a perfectly respectable business, but according to Popper patterns discovered in such data should not be mistaken as demonstrating the existence of universal laws. The important point is that trends cannot be relied upon to make predictions since they are liable to break down at any moment.

“Are ... trends not something comparable with Newton's law of inertia? The answer is: trends exist, or more precisely, the assumption of trends is often a useful statistical device. But trends are not laws. A statement asserting the existence of a trend is existential, not universal. And a statement asserting the existence of a trend at a certain time and place would be a singular historical statement, not a universal law. The practical significance of this logical situation is considerable: while we may base scientific predictions on laws, we cannot (as every cautious statistician knows) base them merely on the existence of trends. A trend ... which has persisted for hundreds or even thousands of years may change within a decade, or even more rapidly than that.”

Popper 1945, pp. 72-3

Popper recognises that social level laws may exist but that they will always be particular, derived laws rather than fundamental, universal ones. Regularities might be discernible in historical series, but, following John Stuart Mill, Popper criticises the view that these regularities be seen as laws, *unless they have been deduced from more fundamental laws* (i.e. those of behaviour).

“[H]e [Mill] had been careful to emphasise that such a uniformity could “only be an empirical law” (the term is somewhat misleading) and that it should not be considered secure before it was reduced, “by the consilience of deduction a priori with historical evidence” to true laws of nature; and that he had even announced the “imperative rule never to introduce any generalisation from history into the social science unless sufficient grounds can be pointed out for it”, viz. deducing it from true natural laws (the laws he had in mind were those of ‘human nature’, i.e. psychology)”

Popper 1945, pp. 75

This is a more accurate vision of neoclassical methodological individualism. It *insists* that macro-level generalisation *must be reduced* else they be considered unreliable. And ultimately, since macro-level laws are only ever derivative of more fundamental laws of behaviour, social science can proceed with an a priori, deductive method which builds up consequences from laws of behaviour. Lower-level behavioural laws are privileged because they are more fundamental (meaning, universal). Indeed, it is presumed that ‘laws of human nature’ are available and universal; they are thought to be ripe to provide a scientific basis for the explanation of social phenomena because of their universality. The problem with historical laws is precisely that they are not universal, but are the consequent of particular conditions. The central mistake of historicism, Popper claims, is to confuse particular, derived laws with universal ones. “It’s ‘laws of development’ turn out to be absolute trends, trends which like laws do not depend on initial conditions” (*ibid.*, pp. 77). The problem Popper has with ‘historicism’ is that it too easily overlooks the particular conditions that give rise to any observable regularities. “There is, for example, a trend towards ‘an accumulation of means of production’ (as Marx puts it). But we should hardly expect it to persist if the population decreases rapidly... There are, indeed, countless possible conditions... [The historicist] firmly believes in his favourite trend, and conditions under which it would disappear seem to him utterly unreal.” (*ibid.*, pp. 78).

Although he may not have been aware of it, Robert Lucas’s critique of econometric methods follows Popper’s argument and can be understood as an application of Popper’s critique of historicism to econometrics. Lucas (1976) attacked the use of econometric methods in policy evaluations on the grounds that the models of the economy that they seek to establish are reliant upon particular conditions that are liable to change, namely the policy-orientation of the time. The effect of this argument was profound and forms the basis for the view that macroeconomics requires microfoundations. The econometric policy tradition of the 1960s and 1970s involved putting forth an holistic view of macroeconomic policy making. The economy was characterised by a single equation, along the following lines:

$$y_{t+1} = F(y_t, x_t, \theta, \varepsilon_t)$$

Where y_t is a vector of endogenous state variables, x_t a vector of exogenous variables, ε_t of independent, identically distributed shocks, $F()$ a functional form and θ a vector of fixed parameters to be estimated. The exogenous and endogenous variables are macro-level, with the model characterised by (F, θ) . F is specified by the modeller, and θ estimated. In order to make policy evaluations, values of x_t are inputted to generate simulations of the endogenous variables, and these are evaluated according to some normative criteria. Based upon these sorts of techniques, the Phillips Curve (Phillips 1958) posited a non-linear relationship between inflation and unemployment and wage changes, based on evidence that this relationship had remained stable for over a century in the United Kingdom. This relationship was absorbed into policy decisions; it was understood that there was a ‘trade-off’ between inflation and unemployment, and that policy-makers could therefore use this trade off to guide policy. Lucas criticised this methodology on theoretical grounds. Such methods are fundamentally based on the stability of (F, θ) . Lucas probes at the micro-behavior that must underpin any such empirical regularity.

“[T]o obtain the decision rules underlying (F, θ) then, we have to attribute to individuals some view of the behaviour of future values of concern to them. This view, in conjunction with other factors, determines their optimum decision rules. To assume stability of (F, θ) under alternative policy rules is thus to assume that agents’ views about the behaviour to the system are invariant under changes in the true behaviour of these shocks. Without this extreme assumption, the kinds of policy simulations called for by the theory of economic policy are meaningless.”

Lucas 1976, pp. 25

Hence Lucas’s Popperian conclusion: econometrically measured historical trends cannot be relied upon since the conditions that bring them about are particular to the time in question. Any historical trends are never the reflection of universal laws that can be relied upon. Since individuals rationalise about the behaviour of the system, any trends observed in macro-data are unreliable unless it can be shown how they are the outcome of individual decisions. Lucas rehearses Popper’s argument that inductively inferred empirical relationships are mere trends that cannot be relied upon to form laws.

“[G]iven that the structure of an econometric model consists of optimal decision rules of economic agents, and that optimal decision rules vary systematically with changes in the

structure of series relevant to the decision maker, it follows that any changes in policy will systematically alter the structure of econometric models.”

Lucas, 1976, pp. 41

The upshot of Popper and Lucas’s argument, as was followed after the New Classical revolution, is that all knowledge about macroeconomic phenomena can be derived from microeconomic theory. Any methods that use social-level laws in their explanations are suspect unless they have been derived from microeconomics.

Udehn (2001) provides a perspective on the view of reduction that is involved here that rings true. Reductionism as understood by methodological individualists is actually less an epistemological thesis than an ontological one; it more concerns a reductionist view on the structure of the world, upon which a presumption to turn towards individuals and a rejection of social wholes is based. Helmut Spinner, in a critique of a psychologistic reductionist view of sociology put forward by the German sociologists Hummel and Opp (1968), put it as follows:

“[T]he predominant logico-epistemological level-picture [Nagel style reductionism] has been supplemented by, and based on, an 'ontological' doctrine. This additional ontological assumption postulates, in parallel with the logico-epistemological hierarchy of theories, a corresponding ontological hierarchy of reductive levels, based on a part-whole relation, in the order of the 'things' ... Each of these levels constitutes a characteristic proper domain of some specific empirical science.”

Spinner 1973, pp. 56

That is, the reductionist inclinations of some methodological individualists have been based on a micro-reductionist ontological view (Oppenheim & Putnam 1958). Here the essential feature is that the objects dealt with by the reducing theory are parts of the objects dealt with by the reduced theory. The ontological picture is one where there is a hierarchy of ontological layers, each conforming to laws, with each law derivable from the laws of the lower level. There is an implicit sense in which tackling the lower level is more ‘fundamental’, in light of the reductionist’s view that the lower-levels are more general than the higher-level laws. Udehn concludes that this worldview motivates methodological individualists. They ignore the problem that many of them deny the possibility of anything other than individualistic laws, and so instead...

“...they use the term 'reduction' in a less demanding sense. Probably they intend something like this: whatever can be truly said about social phenomena can, in principle,

be derived from statements about individuals and their actions. Most methodological individualists avoid this roundabout way, however, and go directly to social phenomena. For them methodological individualism is not a principle about reduction of social theories, but about the explanation of social phenomena.”

Udehn, 2001 pp. 330

For the micro-reductionist perhaps it is not so much that methodological individualism must be followed, as that it might as well be. Since all truth about social phenomena can be discovered by following it, holistic methods are simply a waste of time, and can be safely ignored. Social theories cannot provide any knowledge that cannot be built from individualistic theory.

I believe therefore that the most accurate view of methodological individualism involved in neoclassical economics is the idea that *social phenomena should be explained by reduction to a theory of individual behavior.*

From this understanding of methodological individualism it is but a short step to causal individualism. The causal explanatory power involved in reductions of this type comes about through the application of laws of behavior: applying such laws allows one to indicate how variation in individual properties produces variation in individual actions: causal individualism. The theory involved has typically been rational choice theory, as an allegedly universal theory of individual behaviour. Rational choice theory relates the actions of individuals to their beliefs and desires. Thus applying it involves showing how changes in the desires and beliefs of individuals leads to changes in their actions. This is then linked to the behavior of the aggregate. We saw this in the example above: here the actions of individuals are assumed to add up unproblematically, so that one can make causal claims based on how one's theory of individual behavior dictates that changes in the properties of individuals – like the discount rate - effect their actions.

Taking this micro-reductionist attitudes precludes causal institutionalism because institutions are not the components of social wholes, individuals are. As explained in section 2.5, institutions are about the *inter-relations* of components, take the form ‘In circumstance X, do Y’, and so cannot be set as values within a theory of individual behavior through which one relates cause and effect. One cannot reduce to institutions.

Causal holism is discarded because according to this view regularities between social aggregates are highly unstable: they are tied to particular micro-conditions that are liable to change. According to methodological individualism they are only to be trusted once they have been given reductive microfoundations. Any holist causal patterns are just reflections of individual causes, on this view.

In short, the reductionist worldview involved in neoclassical methodological individualism leads to an insistence upon causal individualism. *Neoclassical methodological individualism is defined by its insistence on explanation by causal individualism exclusively.*

2.9 Conclusions

In conclusion then, there is more to neoclassical methodological individualism than institutional individualism. Firstly, causal individualism and institutionalism can be distinguished as descriptive sub-categories of institutional individualism by noting which kinds of entity are highlighted as causal and which are mere background. Using these definitions I have shown an example of how neoclassical economics is causally individualistic. I have also shown why the normative dimension to neoclassicism's methodological individualism leads to causal individualism rather than causal institutionalism - the micro-reductionist ontology of entails causal individualism exclusively.

Having clarified neoclassical methodological individualism in this way, I can now proceed to examine some complexity economics to see whether it is individualist, institutionalist or holist. Is the complexity era exclusively causally individualist and following in mainstream economics' methodological individualism? Or does it cite other kinds of causes, thereby rejecting neoclassical methodological individualism?

Chapter 3

Individuals, institutions and social wholes in complexity economics

CHR characterise the complexity era first and foremost by the belief that the economy is a complex system. They argue that there are many recently developments in economics that can be thought of as stemming from this belief, drawing in fields as diverse as evolutionary game theory, ecological economics, behavioural economics, complexity theory, agent-based computational economics, non-linear systems dynamics, experimental economics and some developments in econometrics. As noted in my introductory chapter, some participants of these fields might question whether they are part of a wider ‘complexity era’. To tackle the present issue then, I shall focus on a field that I am sure all would agree it is appropriate to call complexity economics: agent-based computational economics. I shall take some case studies from this field and examine them for how they treat individuals, institutions and social wholes. I shall examine whether in this work economists cite individuals, institutions or social wholes as being the causes of social phenomena, and in doing so determine whether the complexity era is committed to methodological individualism.

I begin with a brief exposition of the components of agent-based modelling (section 2.1). The following sections then present some examples of agent-based models, starting with Joshua Epstein and Robert Axtell’s *Sugarscape* (section 2.2), followed by Alan Kirman’s models of fish markets (section 2.3). In each case I examine the roles played by individuals, institutions and social wholes in the explanations these models produce.

I shall show how complexity economics is *methodologically pluralist*: in these examples we shall see examples of causal individualism, causal institutionalism and causal holism. Of these, causal individualism appears to be the least common; in fact, very often the explanatory strategy involves demonstrating that an institutional cause is robust to changes in the background assumptions about individuals, i.e. that individuals properties are not causally significant. The causal holism of complexity economics arises because agent-based models are predicated on the idea that there are *emergent* patterns in the aggregate. Through these examples I shall discuss what is meant by emergence – principally the idea that such aggregate patterns are irreducible to individual laws - and show its association with causal holism.

3.1 Complexity, agent-based modelling and emergence

Agent-based techniques have found applications in almost all areas of science and engineering. The more general term for this methodology is the multi-agent systems approach. The agents that populate these systems can really be many different kinds of things; human individuals yes, but also aeroplanes or ants. But social scientists have taken up the multi-agent systems approach to develop models of societies built out of agents that are intended to represent human individuals. Joshua Epstein and Robert Axtell are two social scientists who have contributed much to the development of the field, who co-authored the seminal *Growing Artificial Societies: Social Science from the Bottom-*

up (1996). They describe how they view their work as the involving the specification and examination of artificial societies, the intention being to provide explanations for social phenomena.

“In this approach fundamental social structures and group behaviors emerge from the interaction of individuals operating in artificial environments under rules that place only bounded demands on each agent's information and computational capacity. We view artificial societies as laboratories, where we attempt to ‘grow’ certain social structures in the computer- or in silico- the aim being to discover fundamental local or micro mechanisms that are sufficient to generate the macroscopic social structures and collective behaviors of interest.”

Epstein & Axtell (1996, pp.4)

Specifying an artificial society involves three aspects; *agents*, an *environment* or space and *rules*. Agents are the ‘people’ of the artificial society that is developed. These agents are defined by a number of characteristics, some of which may remain constant, but many are in constant flux as they interact with those around them and their environment. Some characteristics, like sex or race, might well be fixed for the agent’s lifetime, but others, such as wealth, income or preferences might well be modelled such that they change. Agents are also characterised as employing various behavioural rules that determine their actions as they go about their artificial lives.

The agents of these artificial societies live in some environment. This might simply be a flat grid, with each individual taking up one square from which they interact with those around them. Alternatively, the agents may be allowed to explore a more elaborate environment, such that each position of the grid is characterised by certain characteristics. For instance, it could be fertile land from which our agents’ food grows. The environment may also be a more abstract structure, such as a communication network, or it could be a market with various selling outlets, which buyers are able to explore.

Finally there are the rules that govern how the artificial society develops over time. Firstly there are rules that govern how an agent interacts with his environment: how he is able to move about it, if at all, or how he collects resources. There may also be rules about how the environment develops by itself: one area might be more fertile than another, say. Thirdly, there may be rules that determine the interaction between agents: how they communicate with each other, or compete, cooperate, trade or otherwise go about their social lives.

The agents, environment and rules are then usually programmed on to a computer, although this is not strictly necessary. Indeed, one of the first agent-based models in social science, Schelling’s model of social segregation (Schelling 1978) is simple enough to be played on a chess board. In any case, whether worked out laboriously by hand or with the aid of a computer, the interactions between

agents and with their environment are played out step by step. An initial state for the world must first of all be specified, and then the outcome in each successive state is calculated according to the rules. Insofar as the artificial world that is specified is supposed to in some way correspond to the real world, these are referred to as *simulations*.

Data that is collected from simulations of how the model world develops is then examined for *emergent properties*: stable features that seem to ‘emerge’ out of all of the interactions. “[T]he defining feature of an artificial society model is precisely that fundamental social structures and group behaviours emerge from the interaction of individual agents operating on artificial environments under rules that place only bounded demands on each agent’s information and computational capacity” (Epstein & Axtell 1996, pp.6).

Let me now turn to some examples of agent-based models, and begin how they use individuals, institutions and aggregates in their explanations.

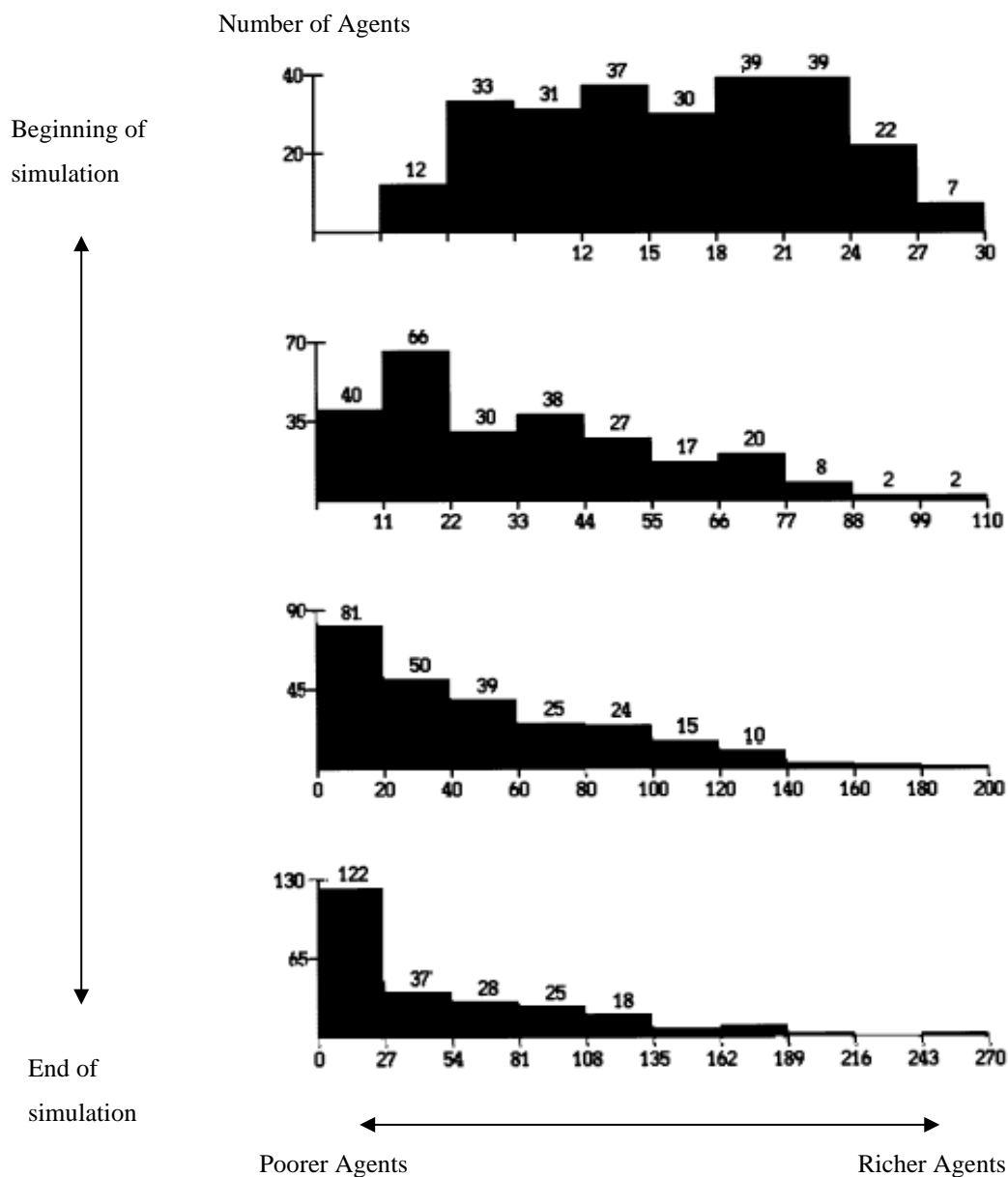
3.2 Sugarscape

In *Growing Artificial Societies*, Epstein & Axtell developed an artificial world called Sugarscape, in which emerges a wide variety of features commonly observed in the real world, including population density, inequality, migration, trade, and even conflict and banking. Sugarscape is a landscape populated by agents that like to ‘eat’. The landscape is divided into a grid, with each square containing a certain amount of our agent’s food. The authors begin with a simple set up for Sugarscape and gradually introduce more complicated features. In the basic set up sugar is the only food, which is initially distributed across the landscape such that some regions have lots of sugar while others have little. Sugarscape has two ‘sugar mountains’, in the upper left and bottom right corners of the landscape. These are fertile areas that grow more sugar than the barren land in the middle of the Sugarscape. Once sugar has been eaten, it begins to grow again on the square, up to the amount that its fertility can manage. Agents are initially randomly distributed. They are endowed with characteristics that define their vision and metabolism, with random variation of these properties built into the population of agents. The agents then proceed according to a simple rule that says ‘Look around as far as your vision permits, find the spot with the most sugar, go there and eat the sugar’. The metabolic rate of the agent determines how much of the agent’s sugar he uses up with each move he makes. Agents die if and when they have used up all their sugar.

On simulating this artificial society it becomes apparent that the agents quickly learn where the sugar mountains are, and begin to coalesce around them. They are very efficient grazers, quickly moving to eat sugar wherever it appears. The agents effectively organise themselves into two concentrated ‘tribes’ on top of each mountain, tribes which harvest the available crop well.

Another aspect that becomes very apparent once one looks at these simulations is that Sugarscape develops to become quite an unequal society. Sugarscape starts out very egalitarian, as each agent is randomly assigned some sugar. But as the history of Sugarscape unfolds the distribution of wealth changes and becomes skewed: a few agents become very rich, there is a middle class, and a big underclass of very poor agents. The development of the income distribution is shown below, in figure 3.

Figure 3: The distribution of wealth in Sugarscape



Source: Axtell & Epstein (1996), Animation II-3, Evolution of wealth histogram, pp. 36.

Sugarscape develops to be quite an unequal society with a small number of agents holding most of the wealth. This is of interest not because we are concerned about the welfare of electrons, but because highly skewed distributions of wealth are a commonly observed feature in human societies too. The distribution that develops – an emergent property - is in fact a Pareto distribution. This is a distribution that Vilfredo Pareto first observed in 1895 after collecting data on income distribution (Pareto 1965[1895]¹⁷). From this distribution the famous ‘80-20 rule’ is derived: the idea that 80% of the wealth goes to 20% of the population. “Thus we find the first instance of a qualitative similarity between extant human societies and artificial society on the Sugarscape” (Axtell & Epstein 1996, pp. 7). This similarity between the model world and the actual world is important. The intention behind the model is to provide a plausible account of how inequitable distributions of wealth occur¹⁸.

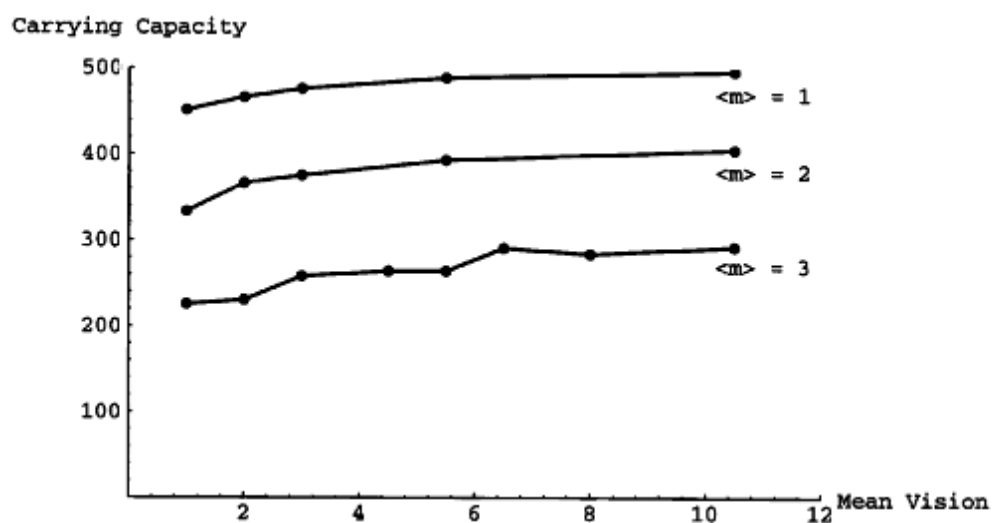
The authors then go on to make some changes in Sugarscape’s design and observe how Sugarscape’s emergent properties are affected. The way that the environment updates can be altered by changing the rule that determines how the sugar grows back after it has been eaten. Under one specification it grows back instantly, in another it grows back slowly over time and in yet another setting seasonal features are added. These different ‘growback rules’ have implications for the amount of life that can be sustained on Sugarscape - what is known in ecology and environmental studies as *carrying capacity*.

When sugar grows back quickly agents are able to acquire more food and so the carrying capacity increases. The carrying capacity also depends on the agents. Varying the initial mean vision or the initial mean metabolism of agents changes how efficiently the agents harvest the sugar, and so results in different carrying capacities. The result is shown in figure 4 below. The graph shows how carrying capacity is related to initial mean agent vision under three different assumptions about the mean metabolism of the agents ($\langle m \rangle = \{1, 2, 3\}$). The carrying capacity increases with higher initial values for the agents’ vision. As agent vision improves each agent can see further and becomes more efficient at collecting sugar, increasing the carrying capacity. Similarly, a lower metabolic rate means that agents find it easier to survive, and so again the carrying capacity improves. The growback rule is fixed in all the simulations.

¹⁷ See Persky (1992) for an historical account of Pareto’s law.

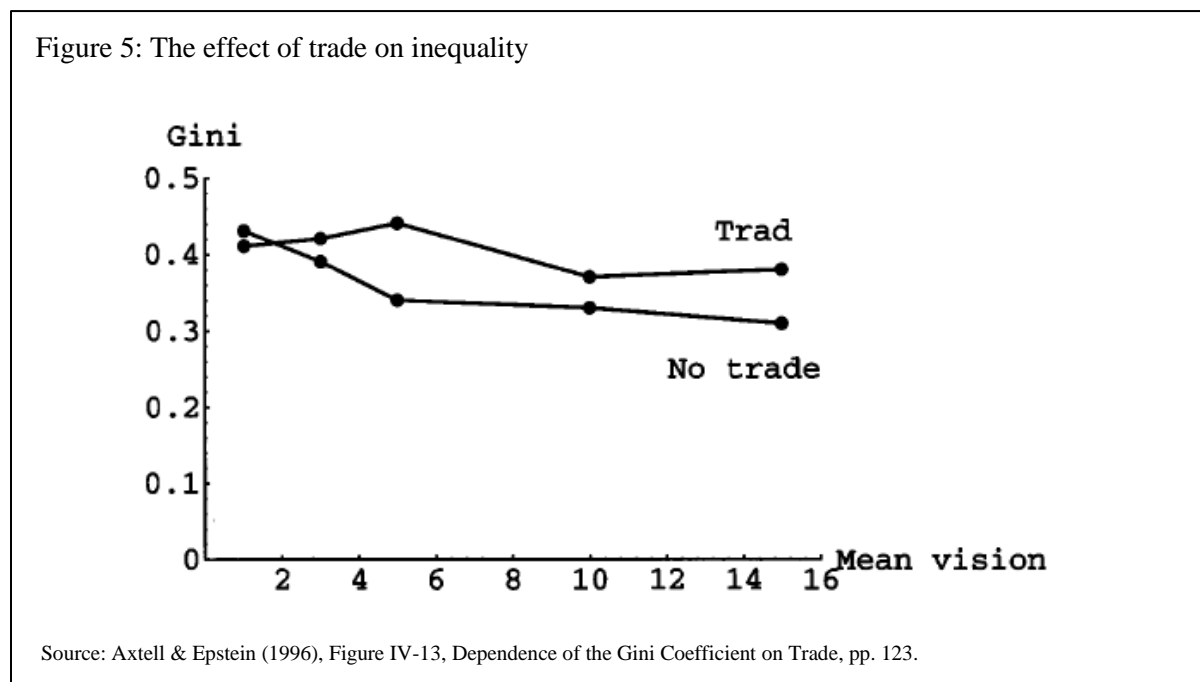
¹⁸ In fact the Pareto distribution has been noted to be extremely common in society, not only in income distribution, but also for things like city size (Auerbach (1913), Zipf (1949)). It is a distribution that many agent-based models reproduce also. Analysis of city size using complexity theory is the subject of much recent research (e.g. Krugman (1996a), Batty (2005)).

Figure 4: Carrying capacity as a function of mean agent vision



Source: Axtell & Epstein (1996), Figure II-5, Carrying Capacities as a Function of Mean Agent Vision, pp. 31.

As well as varying the initial conditions of the agents and the environment's updating rule, the authors also introduced novel ways in which the agents interact with each other and examined the effect of these modifications on Sugarscape's emergent properties. Once reproduction and genetic inheritance were introduced the inequality observed increased. The introduction of inheritance increased the Gini coefficient from 0.503 to 0.743, under identical assumptions about the initial conditions of individual properties and the other updating rules. The authors also added another good, 'Spice'. This introduced trade, allowing the authors to examine whether or not the prices that are observed coincide with the equilibrium price that neoclassical economics predicts (in general, it does not). They were also able to examine the effect of trade on carrying capacity and inequality. Trade increased the carrying capacity of society, but also made it unequal (see figure 5). Another innovation to the kinds of interaction the authors made was the introduction of the possibility of borrowing and lending between the agents of Sugarscape. A capital market developed, involving a hierarchy of super-rich money-men, who lent to middle men who made loans to borrowers. In other words, Sugarscape developed its own versions of HSBC and Citibank.



With this brief description of Sugarscape complete, let me now turn to examine how the Sugarscape models - and the explanations that are grounded in them - use individuals, interactions and aggregate properties. Is the Sugarscape model an example of causal individualism, institutionalism or holism?

3.2.1 Individuals in Sugarscape

Evidently Sugarscape is populated by agents that are intended to represent human individuals. These individuals are characterised by various properties. At the beginning of each simulation the agents are endowed with a range of properties, which are typically randomly distributed across some range of values. These properties were things like wealth, metabolic rate, vision and sex. Some new kinds of properties were introduced to make Sugarscape a more complicated world.

Are the properties of individuals demonstrated as causes, background conditions or effects in Axtell & Epstein's Sugarscape models? In keeping with the quaternary structure of causal claims discussed in chapter 2, there are several instances where the authors show how variation in the initial conditions of individual properties produce different emergent outcomes. Different carrying capacities result from different initial conditions on the agents vision and metabolism; ergo, agent vision and metabolism are demonstrated as being causes of carrying capacity. A set of background conditions about the kinds of interactions that can take place in Sugarscape are specified – some institutions are assumed, but held constant across the comparison. The authors then show how variation in individual properties produces variation in social phenomena. This is clearly an example of causal individualism.

At other times the individual properties are mere background conditions. In several instances, the authors examine the effects of variations in interaction rules, holding individual properties constant. They did this for instance when observing the effects of reproduction and genetic inheritance on inequality; to make this comparison robust they demonstrate how the change arises in simulations with identical initial conditions for the properties of individuals.

3.2.2 Institutions in Sugarscape

The different versions of Sugarscape that the authors develop mainly come about as they introduce different institutions. These institutions are the rules of interaction between the agents; the various updating rules of the form ‘in circumstances X, do Y’. In the most basic setup, agents behave according to the ‘look as far as you can and go to the place with the most sugar’ rule. As they move and eat, their activity has knock on effects to the other agents as they are consuming the same resources. These interactions come about as agents apply the institutional rules that characterise how agents interact with themselves and their environment.

What role do these institutions play in the explanations that Axtell and Epstein put forward: are they causes, background conditions or effects?

I mentioned briefly earlier how the emergent income distribution changed when Epstein and Axtell introduced trade to Sugarscape. Sugarscape becomes more unequal once agents are allowed to trade with one another. If one asks what explains the *change* in the income distribution between these two cases, then the answer is that the way that individuals *interact* has changed. The introduction of a new kind of interaction – bilateral trade – brings about a change in the emergent income distribution. In fact most of *Growing Artificial Societies* is about studying how introducing different institutions – I have highlighted inheritance, trade and credit - brings about differences in the outcomes observed.

Once again, what is produced from these models are causal claims in a quaternary structure: different institutions, different social outcomes. When these causal claims are made, background assumptions about the properties of individuals are made – their initial conditions on vision, metabolism and so on - but these are held constant across the comparisons. *Growing Artificial Societies* is predominantly a work of causal institutionalism.

3.2.3 Wholes in Sugarscape: Emergent patterns and their complex explanations

Having designed Sugarscape by specifying rules of interaction and some initial conditions on the properties of individuals, multiple simulations are run to generate data on individual outcomes. This data is then examined for statistical regularities and patterns. For instance, the authors plot histograms

of income and the data is tested for conformity to well-known distributions. As statistical properties of the entire artificial society being simulated, the properties that are searched for once the simulation has been run – such as the income distribution – are *macro* properties. Complexity science treats macro properties as being *emergent*. What is meant by this and does it relate to methodological holism?

Emergence has been a philosophically controversial idea. The first wave of emergentist thinking, dubbed as ‘British Emergentism’¹⁹, was especially concerned with the scientific gaps that persisted between physics, chemistry and biology. At the time of writing there were no plausible micro-explanations of chemical bonding available, and it seemed that the laws that chemists had discerned about chemical bonding had to be taken as mere ‘brute facts’, which emergentists felt had to be accepted with a certain ‘natural piety’. Around this time, as a result of the lack of micro-explanations for chemical processes, scientists spoke of chemical *forces* (McLaughlin 1992, p. 24). Alongside the concern for the gap between physics and chemistry, the British Emergentists were also interested in the gulf between biology and chemistry. Emergentists sought a path between reduction-minded ‘mechanists’, who took it that life is governed purely by physics and chemistry, and the anti-reductionist vitalists, who endowed life with a mysterious ‘entelechy’, a primitive, life-giving substance. A middle way was to found by eschewing vitalist substance, but maintaining the explanatory irreducibility of biological properties. Emergentists argued that a certain class of higher-level properties exists that cannot be explained in terms of lower-level properties scientifically, despite owing their existence to lower level properties.

“The higher-quality emerges from the lower level of existence and has its roots therein, but it emerges therefrom, and it does not belong to that lower level, but constitutes in its possessor a new order of existent with its special laws of behaviour. The existence of emergent qualities thus described is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the “natural piety” of the investigator. It admits no explanation.”

Alexander 1920, p. 46-47

The British Emergentist position brought with it a certain amount of humility with regards the possibilities of science: it claimed that certain phenomena were simply not going to be explained in terms of lower-level phenomena and had to be accepted as ‘brute facts’. Emergentism did not fare well

¹⁹ Brian McLaughlin (1992) named this first wave ‘British emergentism’, commencing in parts of J. S. Mill’s *System of Logic* (1843) and continuing through Alexander Bains’ *Logic* (1870), George Henry Lewes’s *Problems of Life and Mind* (1875), Samuel Alexander’s *Space, Time and Deity* (1920), Lloyd Morgan’s *Emergent Evolution* (1923), C. D. Broad’s *The Mind and Its Place in Nature* (1925) and later traces can be found in authors such the eminent neurophysiologist Roger Sperry (e.g. Sperry 1964).

during the immediate postwar period however. Brian McLaughlin (1992) argues, sensibly it seems as certainly the timing matches well, that the British Emergentist tradition was brought to a rather abrupt close as a consequence of developments in quantum mechanics. Niels Bohr's quantum mechanical explanation of chemical properties meant that there was no need to invoke any configurational chemical forces and to accept chemistry with 'natural piety'. This was a great boon for the value of reduction and cast considerable doubt over the status of emergent properties. Hempel & Oppenheim (1965, p. 64) castigated the doctrine of emergence for "encouraging an attitude of resignation which is stifling for scientific research" (p. 65). Claims about emergent properties were taken to be nothing other than admissions of ignorance, waiting to be filled by satisfactory explanations.

Nonetheless, even though complexity science speaks of emergence ubiquitously, the emergent macro phenomena complexity science studies – like the emergent distributions of income that Axtell and Epstein produce in their Sugarscape simulations – *are* modelled to be produced out of the micro-level. Mark Bedau (2003) has articulated the notion of *weak emergence* to capture the kind of emergence at hand in complexity science. In contrast to *strong emergence*, whereby the causal powers of macro phenomena completely evade micro-level explanation, weakly emergent phenomena can be derived from the micro-facts.

Bedau defines weakly emergent phenomena as being distinctive from other macro-level phenomena in that they can only be derived by *simulating* all the interactions between micro-level phenomena, rather than by an analytical deduction from micro-level laws using theoretic reduction. In a normal scientific reduction, macro-level laws are reduced to micro-level laws by providing bridge laws that link micro-level and macro-level phenomena²⁰. For example, the ideal gas law has been reduced to the kinetic theory of gases by providing a bridge law that relates higher-level properties pertaining to whole gases (temperature, pressure, volume) to the lower-level properties of their component particles and coupling this to the laws known about the behavior of component particles. A bridge law indicates that temperature is proportionately related to the mean velocity of component particles, which allows one to analytically reduce the ideal gas law to laws of the motion of a typical particle.

A 'short-cut' is provided here: it is not necessary to examine in detail the properties and interactions of all component particles to provide an explanation of the gas law relating temperature, pressure and volume: it is sufficient to understand the behavior of the mean particle. Hence weak emergence is about explanatory technique: weakly emergent phenomena are not reducible, but rather require a full simulation of all the interactions between particles.

How does this apply to the present example? Consider the emergent phenomena of the Pareto income distribution that Axtell & Epstein produce. A reductionist explanation of this distribution

²⁰ Nagel (1961) is the seminal account of theoretic reduction.

would seek to analytically deduce this distribution from the properties and laws pertaining to individuals. There are many things about the individuals in Sugarscape that could explain the observed income distribution. For one, some agents have better vision than others and since this will affect how much food an agent is likely to find, it could be that the income distribution is caused by the differences in agents' vision. Similarly agents with a lower metabolic rate might well do better as they are more likely to survive and find their way to the lucrative sugar mountains. And of course some agents are lucky enough to be born in those fertile mountains, and so Sugarscape's landscape might explain the emergent income distribution.

It is common to divide these sorts of properties between those that involve 'nature' (genetic characteristics) and those that are based on 'nurture' (the environment one happens to be born in). In the Sugarscape world, agents are born with certain 'natural' characteristics such as their genetic dispositions of vision and metabolism. There are also 'nurture' factors, as some agents happen to be born on top of sugar mountains while others are born in the barren valley. One can then ask whether the distribution of income that is observed arises because of either of these 'nature' or 'nurture' factors, or perhaps some combination of them. Traditionally one tends to think of these factors as being separate, meaning that if the contributions of 'nature' and 'nurture' could be discerned separately, their effects could be added together and this should give us what actually happens.

So, first one might ask whether the observed Pareto distribution of income arises because of our agents 'nature' - their genetic endowments of vision and metabolism. The observed distribution cannot be reduced to these properties however, because these properties are initially given to agents according to a uniform distribution. If wealth corresponded with genetic fortunes, then Sugarscape would be a much more egalitarian society. Secondly, one can consider whether the differences in wealth that come about arise because of the way that agents are initially distributed - 'nurture'. It could be, for instance, that the way that some agents happen to be 'born' in the fertile mountains and some in the barren valley explains the differences in income. But this does not account for the income distribution either; agents are placed entirely randomly on Sugarscape, and as shown in the preceding figure, the initial income distribution is a more egalitarian bell-curve. Moreover, one can trace the fortunes of agents that are placed very close to each other on Sugarscape, and observe a wide divergence in their eventual wealth. Early moves by an agent can make a big difference to their fortune, as the agents in Sugarscape are very quick to discover the Sugar mountains, and so the moves an agent makes in the first few minutes of his life can make a big difference. Two agents born in similar locations can have very different life outcomes depending on whether in their early life they happen to move toward a mountain or away from it.

So what explains the emergent income distribution, if it is not the agent's genetic endowments or their fortunes at birth? The answer is, as Eric Beinhocker (2007) puts it, "everything" (*ibid.*, pp. 86).

“It is a macro behaviour that emerges out of the collective micro behaviour of the population of agents. The combinations of the shape of the physical landscape, the genetic endowments of the agents, where they were born, the rules that they follow, the dynamics of their interactions with each other and with distribution.”

Ibid.

The crucial essence of the notion of emergent properties is that they cannot be reduced to the laws and properties of components. Instead they need to be derived by a simulation which details all the interactions of components as we have seen in the Sugarscape agent-based models. One cannot separate the factors that are involved and combine them neatly.

As a result, the macro-level emergent patterns that are produced are not simply reflections of laws pertaining to the behavior of individual. In going through a full simulation of all the interactions, a complete account of everything that makes up the whole is given. To say that the patterns and regularities discovered at the macro level are emergent means that cannot be reduced or simplified in any way; they remain features of the entire system, rather than scaled up features of components. It is in this sense that Kirman and Mirowski remark that the laws of complexity economics are laws of societies, not laws of individual behavior. The complexity of economic systems entails that to examine how the micro and macro level relate the operation of the entire system needs to be accounted for in a simulation. If macro features are emergent, they need to be produced *in toto*.

The complexity vision – the belief in emergent patterns – can be associated with holist attitudes. The early British Emergentists believed that the irreducibility of certain properties generates the hierarchy of sciences, each of which is concerned with studying the irreducible causal powers of the objects at its level. In economics, methodological holists were emergentists who argued in favour of a domain of irreducible macroeconomics. When Clarence Ayres argued that social wholes form a “causal nexus the analysis of which is the problem of the social sciences, and in this analysis of social causes and effects the acts of individual men are not at issue” (Ayres 1962, pp. 97), he is stating an emergentist position.

A complexity outlook is fully consistent with the idea that “the aggregate has an existence apart from its constituent particles and behavior characteristics of its own not deducible from the behavior characteristics of the particles” (Vining 1949, pp. 79), if one understands that deduction as involving theoretic reduction, as Vining must surely have done. Complexity economists believe, just as methodological holist authors have, that economies, markets and other social wholes display statistical patterns that will recur and have a stability to them. Contrary to the reduction minded methodological individualists like Mill, Popper, and Lucas, complexity economics involves the holist belief that there are regularities, patterns or laws of entire economies.

One such holist law that has been speculated upon is the ‘Pareto law’ that Epstein & Axtell simulated in Sugarscape; repeated observation of the Pareto distribution in income statistics has given rise to the idea that such a distribution occurs as a social level law.

“[S]ome have argued, especially in the context of the so-called Pareto ‘law’, that highly-skewed distributions of income and wealth represent some sort of ‘natural order’, a kind of immutable ‘law of nature’.”

Epstein & Axtell 1996, pp. 36

The holism of complexity economics attenuated however. Complexity economics involves *weak* emergence, and so can be said to involve a kind of *weak* holism. One cannot go so far as to say that “the acts of individual men are not at issue”, as some holists have argued. It is not that individuals do not matter at all. I have already demonstrated an instance of causal individualism in complexity economics; different initial conditions of metabolism and vision make a difference to the inequality that emerges. The complexity view does not entail that patterns observed at the macro-level are hard-wired laws that forever obtain. Rather, they are sensitive to the individuals and institutions involved in the system.

“Artificial social systems let us explore just how immutable such distributions are. We can adjust local rules – like those concerning inheritance and taxation – and see if the same global pattern in fact emerges.”

Epstein & Axtell 1996, pp. 36

So the holism of complexity economics is of a weaker kind than some earlier holist economics, but it does remain in the spirit of holist thinkers. The crucial difference between holist and individualist thinkers really lies in their attitude towards the stability of social level regularities. Where individualist authors have been highly sceptical of the notion that there could even be higher level regularities (see section 2.8), complexity economists instead believe that stable regularities will be forthcoming in statistically aggregated data. The principle of emergence, even in its weak kind, suggests that economies will settle down to produce macro-level regularities, provided that certain micro-level conditions persist. The same could be said of our individualists, to a degree, except that the proviso is much weaker for our complexity economists. Whereas individualists believe that changes to micro-level conditions – such as a change in individual preferences – are very likely to lead to a break-down of the macro-level law, complexity economists view emergent patterns as being robust to certain changes in the micro-conditions. Indeed, the idea is that an economy ‘settles down’

into producing and re-producing such emergent patterns, and only very fundamental changes – like changes to institutions – produce change in the observable macro-pattern.

3.2.4 Conclusions on Sugarscape

All this serves to demonstrate great plurality about complexity economics. I have shown how through their simulations of Sugarscape Axtell & Epstein have committed both methodological individualism and institutionalism and how complexity's commitment to weakly emergent phenomena renders them holistic in the sense that it involves the belief in stable social level laws that are not simply scaled up individual laws.

3.3 Kirman's fish markets studies

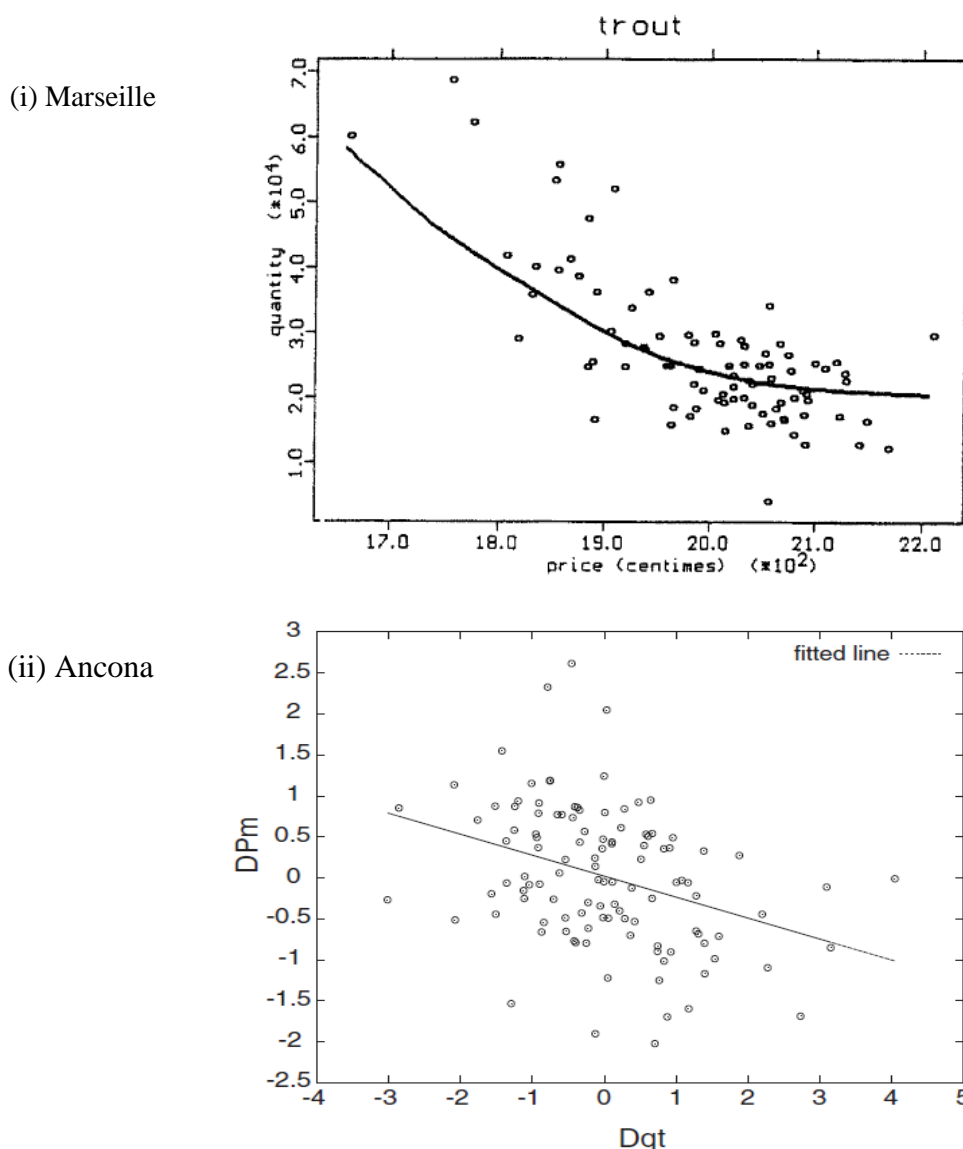
The second case study I shall examine is Alan Kirman's analysis of fish markets. Together with his collaborators, Nicolas Vriend, Mauro Gallegati, Gianfranco Giulioni and Antonio Palestrini, Kirman has conducted a series of studies of fish markets. These involve firstly examining the large data sets collected by the authors to discover social phenomena in need of explanation. The authors then look in detail at how the markets are organised to guide the construction of agent-based models. The insights of such analysis are summarised in Kirman (2011a, Ch. 3)²¹.

The markets Kirman has examined along with his collaborators vary in their organisational features. The Ancona market consists of 3 simultaneous Dutch auctions, where initial prices of 3 cases of fish are set by the auctioneer and displayed for the buyers to see. The prices fall, and buyers press a button to purchase a case. In contrast, the Marseille market involves pairwise transactions between specific buyers and sellers with no posting of prices. Modelling these in agent-based models involves taking account of the different ways in which these markets are organised. I shall focus on the agent-based model of the Marseille market constructed in Kirman & Vriend (2001), which draws on much data collected in an earlier study (Hardle & Kirman 1995). The Marseille model is not supposed to provide a general account applicable to all markets, but rather is attuned to the specifics of the way the market in question is organised.

²¹ This chapter also provides an interesting summary of Kirman's journey out of general equilibrium and game theory to an agent-based approach. I shall not dwell on his arguments in favour of agent-based modelling against a game theoretic approach to interactions, but the essence is that the game theoretic approach becomes unwieldy in its complication to little benefit. Equilibria of optimising agents may be found, but fail to produce many of the stylised facts that are displayed in the data. I shall confine myself more narrowly to examining the Marseille agent-based model and Kirman's interpretation of 'the demand curve'.

The modelling approach taken in Kirman & Vriend (2001) is to aim to produce as simple a model as possible that generates the emergent features observed on the Marseille market. The simplicity here pertains to the kind of learning that agents are capable of. Kirman summarises his methodology: “how simple can we make a buyer’s behaviour and still reproduce certain regularities we observe in the data” (Kirman 2011a, pp. 122). The emergent features of interest are several stylised facts observed in the data. Firstly, there are several facts about the prices that occur on the market. Plotting price-quantity data from the Marseille market gives rise to what is commonly known as ‘*the demand curve*’, a downward-sloping relationship that arises for the markets for individual fish and for the market for all fish. A similar downward-sloping relationship is also observed in Ancona (see figure 6).

Figure 6: Market price-quantity plots for Marseille and Ancona



Sources: Marseille figure is from Hardle & Kirman 1995, pp. 238; Ancona figure is from Gallegati, et al. (2011), pp. 4. Dp_m is difference in average market price, Dq_t is difference in daily average quantity.

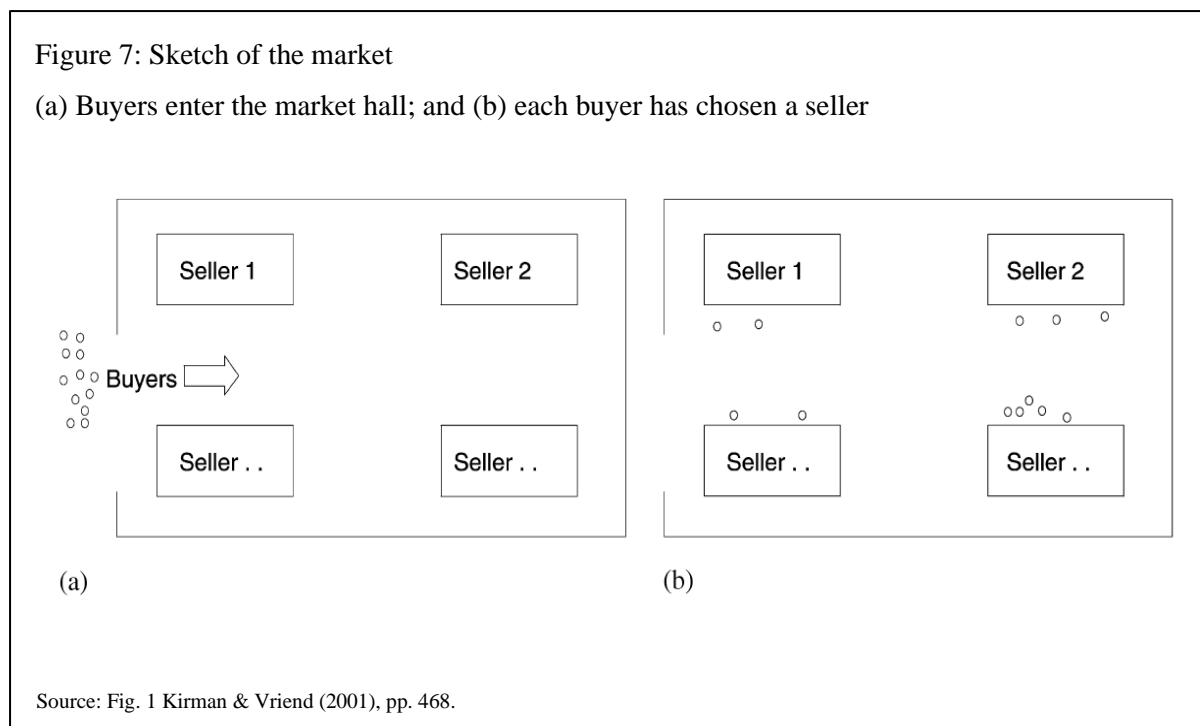
In addition to this downward-sloping relationship much *price dispersion* occurs; not everyone pays the same price each day, but a range of prices occur. Thirdly, it is also noticeable in the data that prices tend to be higher in the afternoon session than in the morning. This is an unusual feature since one would suppose that a rational consumer would always buy in the morning, leading prices to equal out between the morning and afternoon sessions. Finally, another sort of emergent feature that is analysed is the loyalty that occurs on the market; buyers tend to go to the same sellers day after day. Loyalty arises on both the Marseille market and the Ancona market. It is not that surprising that loyalty occurs between buyers and sellers on the Marseille market as the market is organised by pairwise trading. The loyalty on the Ancona market is more difficult to explain however, as the market is organised by auctions.

The model Kirman & Vriend (2001) produced is intended to replicate the salient features of the way in which the Marseille market is organised. Here is the stylised description of how the market is organised.

“In the morning before the market opens the sellers purchase their supply outside the market for a given price that was identical for all sellers and constant through time...The market opens and the buyers enter the market hall. Each buyer requires one unit of fish per day. All buyers simultaneously choose the queue of a seller. The sellers then handle these queues during the morning session. Once the sellers have supplied all the buyers who are willing to purchase from them the morning session ends. All those buyers who are still unsatisfied choose the queue of a seller in the afternoon. Of course, the only sellers present in the afternoon are those who did not sell all their stock in the morning. Sellers now sell to those buyers who are willing to purchase from them and the end of the afternoon session is then reached. All unsold stocks perish. Those buyers who did purchase fish, resell that fish outside the market, at a given price that is identical for all buyers, and constant through time. Each buyer can visit at most one seller in the morning and one seller in the afternoon.”

Kirman 2011a, pp. 101

Modelling this sequence of events first involves designing the environment that agents act within. A digital marketplace is programmed with 10 market stalls, one for each seller. In order to represent the idea that the possibility of trading later on in the day may influence what happens earlier, every day is divided into a morning and afternoon session. The day could be divided into more sessions, but two was sufficient to generate the desired emergent features. At the beginning of each session, the buyers enter the hall and choose a seller's queue to join (see figure 7).



Having designed the model's environment, the rules that define the way the agents interact are programmed. In the model, there are 10 initially identical sellers and 100 initially identical buyers. For simplicity, each buyer wants 1 unit of fish per day. First the sellers decide how much fish they are going to sell by purchasing it at a fixed price ($p^{\text{in}} = 9$). The market then opens and the buyers choose which queue to join in the morning session. The sellers then handle these queues by choosing what order to serve the buyers that have joined their queue. Sellers are able to recognise buyers, with their familiarity modelled as a weighted average of past visits to this buyer, assigning more weight to recent visits. The seller then offers the buyer a price. Sellers' offers are guided by rules that are conditioned on the loyalty of the buyer, the remaining stocks of fish and the remaining queue. The price is either accepted or rejected and the trade takes place accordingly. Once all customers have been served the morning session ends. The afternoon session then begins, those customers without fish select another queue to join, and another set of trading decisions take place.

Each of these decisions is made by agents who are endowed with some reasoning ability. The intention is to make this reasoning process as simple as possible. Programming the process is fairly technical, but the intuition is quite straight-forward: agents continue to do things that they have found successful in the past. This is achieved by a classifier system of various rules which the agent could follow. These rules are all *if... then...* statements, such as '*if the price proposed by the seller for one unit of fish in the morning is 11 euros then accept*'. Each agent has a large set of possible rules for use in his digital 'head', which are then selected for use according to a classifier system through which the

agents learn the best rules to use. The classifier system works by assigning a strength to each rule, which is calculated according to the profit that is generated by using the rule²². Each rule then makes a ‘bid’ in an auction to become the current rule that is in use, where each bid = current strength + ϵ , where ϵ is a random variable. The inclusion of this random variable introduces some experimentation in agents’ behaviour, as they explore different approaches to the decisions they have to make, which becomes less important once a strong rule is found.

Simulations of this specification give rise to all the emergent features of the Marseille market discussed above.

Now, as before with the Sugarscape models, let me now turn to examine the roles played by individuals, institutions and social wholes in Kirman’s analysis of fish markets..

3.3.1 Individuals in Kirman’s fish markets studies

As in Sugarscape the agents that populate Kirman’s models are supposed to represent human individuals. These are characterised first and foremost by whether they are buyers or sellers. In the agent-based model of the Marseille market there are 10 initially identical seller and 100 initially identical buyers. The individuals in Kirman’s models take up *social positions* that involve playing roles in social interactions, roles which involve making particular kinds of decisions. The individuals do not decide themselves whether they are buyers or sellers, or work out what it means to be a buyer or seller for themselves. Rather ‘buyer’ and ‘seller’ are defined by the model’s author as types of agents who undergo certain procedures. The decisions they have to make and the structure of the various strategies they apply are determined by the kinds of interactions they take place in.

Unlike in Sugarscape, there are no initial conditions for buyer and sellers that are specified by the model’s architect other than whether they are buyers and sellers; instead random variables generate the heterogeneity. Although not investigated, one can suppose that changes to the number of buyers and sellers would generate interesting differences in outcomes. Markets with fewer sellers and many buyers might produce more of a monopolistic outcome with lower price elasticity, while markets with few buyers but many sellers might well produce a more competitive outcome with higher price elasticity.

Although there is no heterogeneity built into buyers and sellers through parameters on which some distribution of initial conditions is specified, individual buyers and sellers come to distinguish themselves through their interactions. Some buyers become loyal to particular sellers, while others continue to shop around. These differences in behaviour arise not from any innate individual

²² The buyers here are not themselves end consumers. They are often restaurateurs, and so are modeled as realizing a profit through their actions (as are the sellers obviously), rather than having to make any assumptions about utility.

properties, but from differences that are learnt as different individuals have different experiences. These different experiences are generated out of the decision rule classifier system, which involves a random component to spark off different behavior. Initially, when agents have little experience and so have little information about each other to use in their decision making, the random aspect of the classifier system takes control so that agents explore each other randomly. These initial differences become hardwired into the current strength values of each of the possible rules that the agents consider. As a result, the agents become habitual, but it is important to stress that these habits are not built into the agents in the model's initial setup. Rather, they are determined endogenously.

The classifier system can be thought of as a meta-rule of reasoning that determines the agent's use of rules of interaction. This meta-rule of reasoning is hard-wired into individuals – it is an exogenous aspect to the model found within individuals. Kirman's model has another layer to it than the Sugarscape models, as the agents no longer apply fixed behavioural rules, but use a meta-rule to reason and learn the best behavioural rules to use. Kirman treats individuals in quite a different way to the Sugarscape models. Rather than mechanically applying a strict behavioural rule, the agents have some intelligence about them: they search out profitable ways of behaving and only continue to use a behavioural rule if they find it to be successful.

There are three aspects of individuals that are exogenous in the Marseille fish market model: (i) the number of agents, (ii) their character as buyers or sellers, and (iii) their reasoning.

All other aspects of individuals – which rules they rule, how much stock the sellers take to market, the prices they offer, the prices accepted and rejected by buyers, the loyalty that they display, etc. – all these are determined endogenously.

Unlike Axtell & Epstein, Kirman does not engage in any causal individualism: he does not examine how variations in the initial distributions of individual properties produce variation in social outcomes. In fact, Kirman embarks on quite the reverse: the models are intended to be robust to changes in individual properties. Multiple runs of the agent-based model for Marseille are run using randomly varied initial conditions for individual properties in order to demonstrate that variation of individual properties does *not* influence the emergent outcome. This does not preclude the possibility of causal individualism within Kirman's models; it is simply that Kirman has not looked at individualistic causes.

3.3.2 Institutions in Kirman's fish market studies

The focus of Kirman's fish market studies is on demonstrating the importance of the interactions between agents, where these interactions are determined by social institutions. In the same way that changing the interaction structure on Sugarscape effected emergent income distribution, the different interaction structures of the Marseille market and the Ancona market bring about differences in their

observed emergent features. The Marseille market is organised with pairwise trading with no posted prices whereas the Ancona market involves auctions. This difference in the organisation features of the market – the way in which agents interact – gives rise to different levels of loyalty displayed by agents. The Marseille market displays considerably higher loyalty than the Ancona market. This arises not because the market participants in Marseille are any different from Ancona; rather it is because these markets differ in the way they are organised.

In the Marseille model 9 out of 10 sellers and 9 out of 10 buyers actually get higher profits as a result of buyers showing loyalty, and sellers respecting this by giving them priority in service. This comes about because the transactions that take place are not a zero-sum game; the agents only realise a profit when a transaction actually takes place, but both buyer and sellers can decline to deal with each other. Sellers learn to give priority to loyal sellers, but charge them a higher price. Buyers find it more profitable to be loyal rather than shop around because they are more likely to be served. Loyalty is an emergent feature that comes about because of the way that agents interact with each other; it was not programmed into agents behaviour (as evidenced by the fact that one individual is not loyal) but rather it comes about through the interactions of agents.

Given this, one might expect that loyalty would not occur in Ancona because the kind of interaction that gave rise to it in Marseille – the pairwise trading – is simply not there and so the means by which relationships develop between agents is removed. Interestingly though, loyalty does still emerge in Ancona, but it takes a different form that reflects Ancona's distinctive organisation. At the Ancona market, the name of the vessel of the ship that the fish has come from is posted, giving an opportunity for buyers to become loyal to particular sellers. The extent of loyalty is much lower than at Marseille, as loyal buyers are not able to receive the *quid pro quo* of being served more reliably. The payoff they receive from being loyal to particular vessels simply arises because they work out that particular vessels are of good quality, and so bring them higher profits.

Kirman's two case studies of fish markets attest to the importance of interactions in complexity economics. The different outcomes observed on the markets arise not because the market participants at Ancona are any different to those in Marseille, but rather because different kinds of interactions led to different emergent outcomes. This comparison between different organisational features – the rules by which agents interact – is a case of causal institutionalism. The social outcomes Kirman investigates are analysed as being caused by specific social institutions involved in the market places he examines. He makes his institutionalism clear in concluding the chapter of *Complex Economics* dedicated to his fish market studies.

“The message of this chapter is rather simple. Markets are an important feature of all economies. Each market is characterised by an organisation and structure which will have an impact on the outcomes observed.... Comparison between the markets in

Marseille and Ancona highlights the link between organisational features and aggregate outcomes.”

Kirman 2011a, pp. 126-7

3.3.3. Emergent wholes in Kirman’s fish market studies

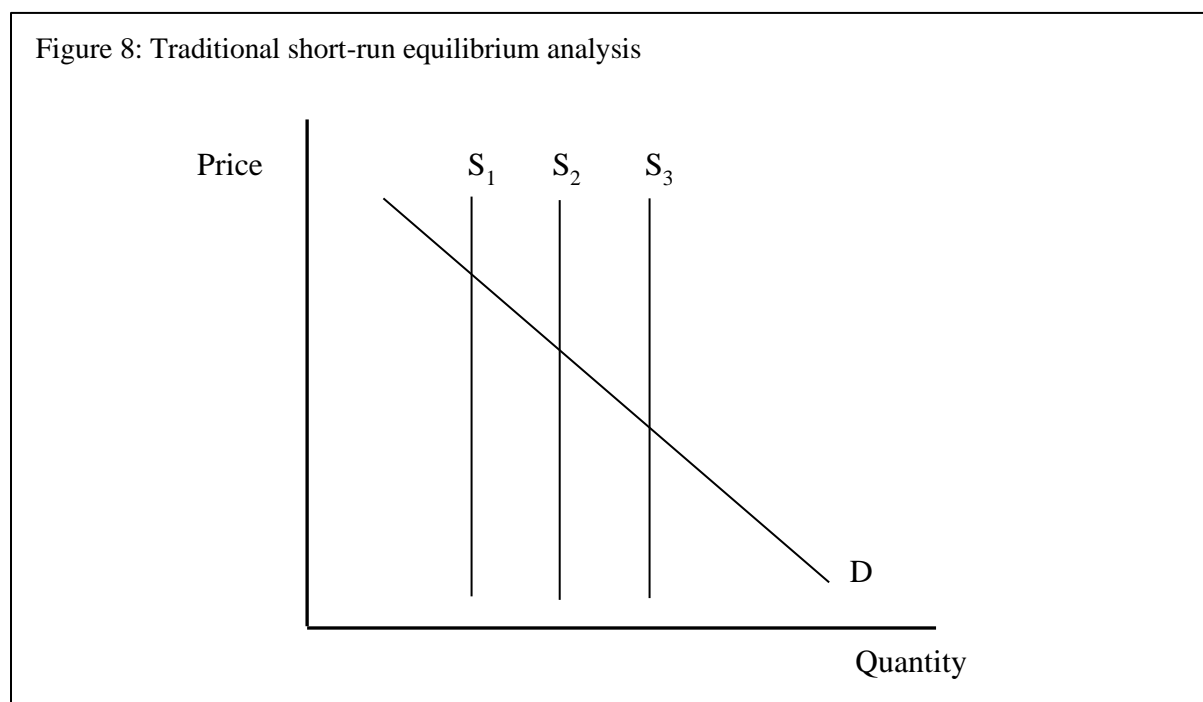
In the Sugarscape models the laws that were searched for were laws of the artificial societies being ‘grown’. So with Kirman’s fish market models; the laws that Kirman attributes are laws of the markets he examines. The laws are discerned by looking for macro stable patterns in the simulated data, and comparing them with the stylised facts of the actual markets. As with Sugarscape, the techniques applied to discern these laws are statistical; a downward-sloping price-quantity relationship is verified by running a regression and conducting hypotheses tests that there exists a negative relationship. Multiple runs of the simulation are used to inductively infer that an emergent feature consistently arises.

As with the Sugarscape example, what makes these properties ‘emergent’ is that they cannot be explained by a reduction to individual laws of behavior. To see why this is the case, it is instructive to look at Kirman’s critique of the mainstream derivation of the downward-sloping price-quantity relationship, as this is predicated on just such a reduction. In mainstream economics, it is held that the list of factors that are thought to influence prices can be neatly separated into those affecting demand for a good, and those affecting the supply. The demand side is affected by consumer preferences, which in turn might be influenced by things like the weather and other seasonal factors, say. And on the supply side we would find factors involved in the cost of production, such as the costs of labour, land and capital. Both demand and supply have been given specific meanings as economic theory has developed, and now form the heart of the education of an economist; one learns how to deduce demand and supply schedules from the theory of consumer choice and the theory of the firm. The two schedules are brought together to determine the market equilibrium price.

According to a mainstream view, the downward-sloping ‘demand curve’ should be clear to see when examining markets where the supply is fixed from day to day, like the fish markets Kirman studied²³. The ‘market demand curve’ is thought to be directly observable because each day gives us a different data point along it. With a fixed supply each day, the short-run ‘supply curve’ is simply a

²³ Indeed, this is why fish markets have long been studied by economists. Kirman’s fish market studies hark back to a 19th debate between John Stuart Mill (1869) and William Thornton (1870), which also examined the nature and existence of equilibria through fish markets.

vertical line. According, the equilibrium points of each day sketch out the demand curve (see figure 8).



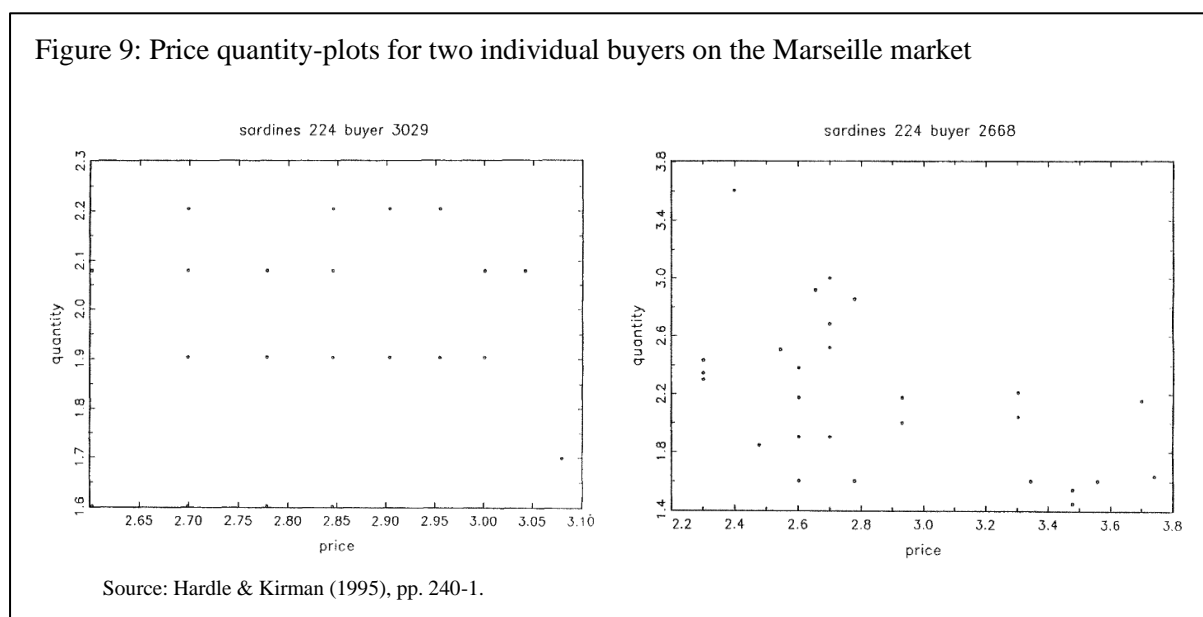
The neoclassical view of the demand curve arises by considering the response of a typical consumer. This typical consumer is understood as having certain preferences over the products available. The form of his preferences are constrained in various ways such that they are ‘well-behaved’, in order to yield sensible results. These constraints include things like the assumption that preferences are reflexive, i.e. if A is preferred to B at time t_1 then A will also be preferred to B at time t_2 , and that preferences are transitive, i.e. if A is preferred to B, B is preferred to C, then A is preferred to C. Having formalised preferences such that they are well-behaved, the consumer is also understood to be constrained in what he can buy because he has a particular budget. From this budget, the consumer then considers the various bundles of products that he can afford at the going prices. The consumer is assumed to be rational; this means that he maximises his utility, subject to a budget constraint, an inequality that states that the linear combination of the vector of goods purchased and the vector of their respective prices should be less than the consumer’s budget.

From such a formalisation, a demand schedule for each product can be derived. This demand schedule relates the quantity that the individual will demand for any given price. With preferences fashioned in the ‘well-behaved’ manner that they are, the demand schedule that results is downward-sloping; that is, as prices rise, the quantity that the individual demands falls. Thus, out of the well-behaved preferences, the budget constraint and the utility maximising assumption, the ‘law of

demand' is derived, that states that the price a consumer is willing to pay for goods falls as the quantity of those goods that the consumer buys increases.

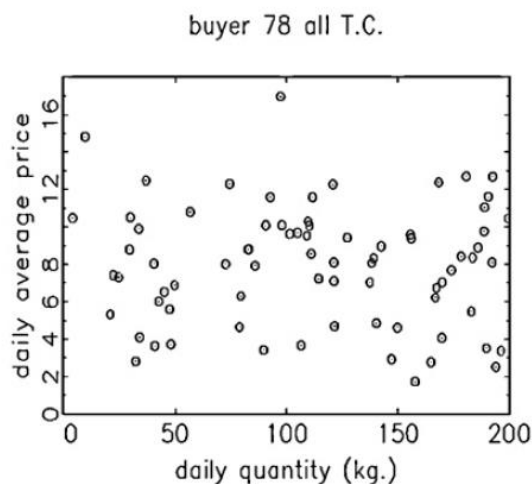
Having established demand schedules for individuals, these are typically assumed to add up to give rise to a market-level demand schedule that displays the same properties as at the individual level. The nice properties that are ensured by the assumptions on preferences are assumed to add up in a simple fashion, preserving those nice properties at the aggregate level. The mainstream view is that the downward-sloping price-quantity relationship observed for the market is a reflection of the aggregation of many individual downward-sloping price quantity relationships. Notice that this is a micro-reductionist view: aggregate demand curves observed empirically are presumed to come about because of individuals have downward-sloping demand curves. The aggregate relation is reduced to the law of demand that holds for individual behavior.

Such a perspective is not borne out in Kirman's data however. When one plots price and quantity data for individual buyers, downward-sloping relationships are not apparent. Price-quantity plots for consumers on the Marseille market are shown below in figure 9.



Kirman found an identical problem in the data from the Ancona market.

Figure 10: Price-quantity plots for an individual buyer on the Ancona market



Source: Kirman (2011a), pp. 115.

Simple observation suggests there is no downward-sloping relationship between price and quantity for individuals, as would be required by a mainstream view²⁴. This is confirmed by the authors using regression analysis. Yet in both these markets downward-sloping price-quantity relationships occur at the market level. In spite of the fact that individuals do not generally display downward-sloping ‘demand curves’, a well behaved aggregate relation obtains. How does this well-behaved aggregate property arise? It is evidently not due to individuals behaving as per the standard model. “While recognising that individuals do not conform to the standard assumptions on rational behaviour subjects seem often to coordinate and to settle down to a particular collective state” (Kirman 2011a, pp. 17).

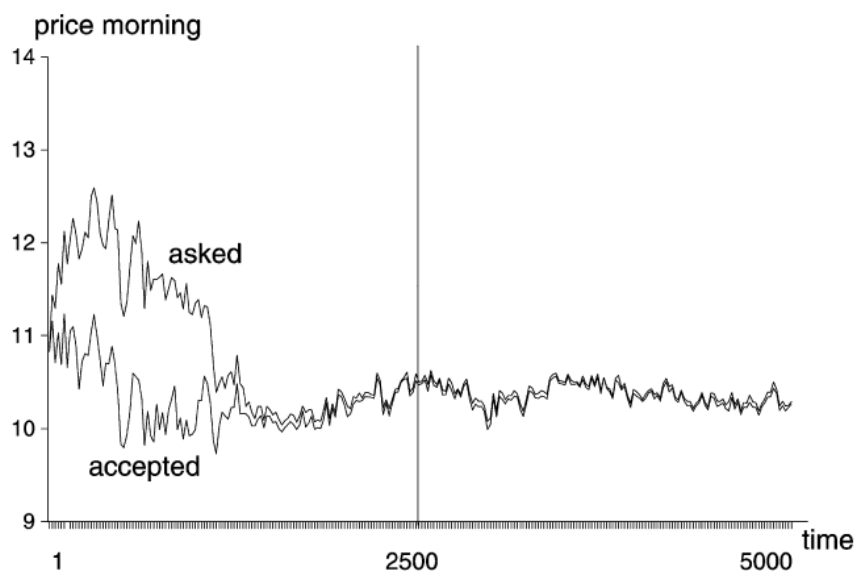
²⁴ This empirical observation is not the only problems with the seamless aggregation that is presumed in neoclassical economics. Theoretical aggregation problems have been understood by microeconomists for a long while. The Sonnenschein-Mantel-Debreu results showed that even with the usual strong assumptions about individual rationality such that individual demand curves are monotonic and decreasing, very little can be said about the market-level price-quantity relationship. Unless it is assumed that all individuals are essentially identical in every respect (i.e. with identical preferences and wealth), the market-level price-quantity relationship can take any form. As a result the derived behavioural laws have little to say about social phenomena. However, the ‘aggregation problem’ has been often ignored, especially by macroeconomists. Elision of the aggregation problem is seen in textbooks, which vary on how much caution they stress to the students in drawing aggregate implications from behavioural laws. Hal Varian’s popular *Intermediate Microeconomics* is notably dismissive. “Since each individual’s demand for each good depends on prices and his or her money income, the aggregate demand will generally depend on prices and the distribution of incomes. However, it is sometimes convenient to think of the aggregate demand as the demand of some ‘representative consumer’ who has an income that is just the sum of all individual incomes. The conditions under which this can be done are rather restrictive, and a complete discussion of this issue is beyond the scope of this book.” (Varian 2006, pp. 267). David Kreps’s *A Course in Microeconomics Theory* (1990) more helpfully informs students on the nature of the problem, although does little more to address it: “unless we are willing to make strong assumptions about the distribution of preferences or income throughout the economy (everyone has the same homothetic preferences, for example), there is little we can say... almost anything is possible for aggregate demand” (Kreps 1990, pp. 63).

Kirman's claim is that this relationship obtains through the *interactions* of buyers and sellers, rather than because of the properties of individuals. 'Demand' factors are not separated out from 'supply' factors and then combined analytically, rather the downward-sloping relationship arises out of the interaction of buyers and sellers. Buyers and sellers learn what offers to accept and reject through their interactions together, and this process leads them to converge on prices that coordinate their interests. Buyers and sellers learn what rules to use – what prices to offer and accept and so on – together through their interaction. It is this co-learning process that determines the outcomes of the simulations, and brings about the emergent properties of interest. One can gain insight into the co-learning process by looking at how the prices offered and accepted evolve throughout the day. Figure 11 shows the prices that are offered and the prices that accepted in the morning session. Each point is an average of 20 simulated days.

In opposition to the neoclassical prediction that a single price rules the entire market throughout the day, observation of the transactions data exposes how prices actually go through some adjustments throughout the day and gradually decline. Through the early part of the morning buyers and sellers go through a considerable learning period, before the average price eventually settles down to around 10 euros. This learning period determines the final outcome, and so serves to bring about the emergent downward-sloping price-quantity relationship. There are two separate learning processes going on: sellers learn about what prices are available, while buyers learn what prices they can charge. Sellers start out offering high prices, and some of these are accepted by buyers who have not learnt to refuse them. But the buyers start to find that they can obtain higher profits by refusing the high prices, and this leads sellers to gradually reduce their prices. As buyers start to observe that the prices being offered are falling, they begin to reduce the prices that they will accept. The important thing to realise here is that the learning of buyers and sellers co-evolves through their interaction together, demonstrating the importance of interactions in agent-based modelling.²⁵

²⁵ Interestingly, some other features emerge from the agent-based models that match what is observed on the real world markets, but contradicts the neoclassical view. One of these features is that the prices charged in afternoons are higher than those charged in the morning. Should this be the case, one would expect that a rational agent will always buy in the morning. But this does not happen. The reason is that just looking at the average hides the variety of prices paid by buyers. The price distribution that is observed generates the higher average price in the afternoon. If, for example, those buyers that encounter high prices in the morning reject in the morning, but accept (perhaps even a slightly lower average price) in the afternoon, the result would be that the average price paid would be higher in the afternoon than in the morning.

Figure 11: Time series average prices asked and accepted in the morning session



Source: Kirman & Vriend 2001, pp. 475.

This leads Kirman to reject the term ‘demand curve’, because it is not a property that arises from ‘demand’ factors alone²⁶. Rather it is way buyers and sellers interact with each other to learn what rules to use that brings about the emergent downward-sloping price-quantity relationship. Hence asking ‘what is it that brings about the downward-sloping price-quantity relationship?’ brings different answers from a mainstream and a complexity approach like Kirman’s. On the mainstream view it comes about because individuals are held to a ‘law of demand’ on their *individual* behaviour. Kirman’s view is that the downward-sloping price-quantity relationship is a complex one that arises through the interaction of both buyer and sellers and is a reflection of the particular interaction rules in place in the market in question. Accordingly it is a *social* law. It is a weakly emergent feature because it is not reducible from behavioural laws. “There is no direct [meaning, reductive] connection between micro and macro behaviour” (Kirman 2011a, pp. 85).

As noted in my discussion of Sugarscape, an emergentist perspective can be associated with methodological holism. Here the irreducible downward-sloping price-quantity relation is useful pattern that can be deployed to support holist causal claims. Empirical knowledge of the elasticities that occur in the different markets can be used to show how variation in the amount of fish supplied to these markets will result in variation in the average prices that result on them. That is, a difference in a social aggregate (the total volume of fish supplied on a given day) is causally linked to differences in

²⁶ See Kirman 2011a, pp. 72-8 for a detailed discussion.

another (the average price). This is not a reflection of an individual law of behavior, but is a law that pertains to the market itself. The aggregate relationship cannot be simplified or reduced in any way; the downward sloping price-quantity relation gives us knowledge about the markets in question that cannot be broken down any further. Simulations can provide an explanation of how they come about, but this involves a complete description of the aggregates at hand, and in doing so the causal claim of the regularity retains its holistic character.

3.3.4 Conclusions on Kirman's fish market studies

Kirman's analyses of fish markets involve both causal institutionalism and holism. Kirman shows how differences in the organisational institutions involved in the markets he has studied bring about differences in their emergent features: causal institutionalism. He also believes that the downward sloping price-quantity relationship commonly observed in most markets is an emergent feature, which enables holistic causal claims to be made; knowledge of this irreducible regularity permits one to show how changes in the quantity supplied to the market place causes changes in the average price that will obtain: causal holism.

He does not demonstrate any individualistic causal claims. On the contrary, by randomising the initial conditions for individual properties, Kirman is trying to provide further support for his institutional causal claims, by showing how the institutional causes are robust to changes in individuals. This does not preclude the possibility of some individualistic causal claims however; it is simply that they are not part of Kirman's analysis, but there may be other unexplored individualistic factors that make a difference to the emergent outcomes.

As with Sugarscape then, there is methodological plurality in Kirman's analyses of fish markets.

3.4 Conclusions

The case studies of complexity economics put forward in this chapter demonstrate that complexity economics is pluralistic when making causal claims. Axtell & Epstein made an individualist causal claim when showing how variation in the initial conditions of metabolic rate and vision produced variation in the carrying capacity achieved on Sugarscape. They also showed institutional causal claims: they showed how changing the institutions in Sugarscape, including reproduction, trade and credit, produce variations in the inequality observed. Kirman makes institutional causal claims when contrasting the Ancona and Marseille markets, noting how their different organisational features produce variation in the degree of loyalty between buyers and sellers emerges. He is also able to make holist causal claims, utilising the price-quantity relation that emerges

from the interaction of agents. Knowledge of this irreducible pattern enables one to make causal claims about how variation in the quantity of goods supplied to the markets in question brings about variation in the average price that obtains: causal holism.

Chapter 4

Conclusions

4.1 The pluralism of the complexity era

This thesis has achieved the following:

- I have overcome the difficulties involved in distinguishing individualism, institutionalism and holism by defining them in causal terms.
- I have shown how neoclassical economics is committed to causal individualism exclusively, by working through an example and by showing how the micro-reductionist ontology of neoclassical economics leads to causal individualism.
- I have shown the plurality of complexity era. At various points, complexity economists have advanced individualist, institutional and holist causal claims.

This provides support for the following conclusion:

- The transition to the complexity era from the neoclassical era involves the transformation of mainstream economics from being a science preoccupied exclusively with individualistic causes to being a science that examines a plurality of different kinds of causes. That is, the complexity era rejects the normative thesis of neoclassical methodological individualism.

The reason for this is the belief that economy is a complex system, offering stable weakly emergent patterns at the aggregate level. Such a belief implies that:

- Economic phenomena will not submit to the reductionist methodology that was central to the neoclassical era's individualist methodology.
- The belief in irreducible aggregate laws leads complexity economists to make holist causal claims of the sort produced by earlier holist economists.
- Since it is the iterated interaction of components that is of primary interest in complexity science, the complexity era will primarily look for institutional causes to explain emergent patterns.

I believe that should the complexity era complete its transition, future historians will note along with me that this was the moment when economics rejected the methodological individualism of the neoclassical era, turning away from exclusively locating the causes of social phenomena within individuals and becoming a more pluralist science, happy to accept either individuals, institutions or social wholes as being the causes of social change.

References

- Agassi, J., (1960), "Methodological individualism", *British Journal of Sociology*, 11(3), pp. 244–70.
- Agassi, J., (1975), "Institutional individualism", *British Journal of Sociology*, 26(2), pp. 144–55.
- Alexander, S., (1920), *Space, Time, and Deity*, 2 Vols. London: Macmillan.
- Anderson, P. W., (1972), "More is different", *Science*, 177(4047), pp. 393-6.
- Anderson, P. W., K. J. Arrow, and D. Pines, eds., (1998), *The Economy as an Evolving Complex System*, Proc. Vol. V, Santa Fe Institute Studies in the Sciences of Complexity. Redwood City, CA: Addison-Wesley.
- Arrow, K., (1959), "The Role of Price Adjustment", pp. 41-51 in M. Abramovitz (ed.) *The Allocation of Economic Resources*. Stanford: Stanford University Press.
- Arrow, K., (1994), "Methodological individualism and Social Knowledge", *American Economic Review (Papers and Proceedings)*, 84(2), pp. 1-9.
- Auerbach, F., (1913), "Das Gesetz der Bevölkerungskonzentration", *Petermanns Geographische Mitteilungen*, 59, pp. 74-76.
- Axelrod, R., (1984), *The Evolution of Cooperation*. New York: Basic Books.
- Ayres, C., (1962), *The Theory of Economic Progress*. Austin: University of Texas Press
- Bain, A., (1870), *Logic*, Book II & III. London.
- Batty, M., (2005), *Cities and Complexity: Understanding Cities Through Cellular Automata, Agent-Based Models, and Fractals*. Cambridge, MA: MIT Press.
- Bedau, M., (2003), "Downward Causation and Autonomy in Weak Emergence" in Bedau, M & Humphreys, P., (2008), *Emergence: Contemporary Readings in Philosophy and Science*. Cambridge, Mass.: MIT Press.
- Beinhocker, E. D., (2007), *The Origin of Wealth: The Radical Remaking of Economics and What It Means for Business and Society*. Boston, MA: Harvard Business School Press.
- Braudel, F., (1987), *A History of Civilisations*. New York: Penguin [1993]
- Broad, C., (1925), *The Mind and Its Place in Nature*. London: Routledge and Kegan Paul.
- Cass, D., (1965), "Optimum Growth in an Aggregative Model of Capital Accumulation", *Review of Economic Studies*, 32, pp. 233-240.

Colander, D., (Ed.) (2000a), *The Complexity Vision and the Teaching of Economics*. Cheltenham: Edward Elgar.

Colander, D., Föllmer, H., Goldberg, M., Haas, A., Juselius, K., Kirman, A., Lux, T., & Sloth, B., (2009), “The Financial Crisis and the Systemic Failure of the Economics Profession”, *Critical Review*, pp. 249-67.

Colander, D. Holt, R. P. F. & Barkley, J. B., (2004), “The Changing Face of Mainstream Economics”, *Review of Political Economy*, 16 (4), pp. 485-99.

Colander, D. Holt, R. P. F. & Barkley, J. B., (2011), “The Complexity Era in Economics”, *Review of Political Economy*, 23 (3), pp. 357–369.

Debreu, G., (1974), “Excess demand functions”, *Journal of Mathematical Economics*, 1 (1), pp. 15-21.

Durkheim, E., (1897), *Suicide*. New York: The Free Press, 1951.

Durkheim, E., (1893), *The Division of Labor in Society*. London: Routledge Classics, 2002. First published 1952.

Elster, J., (1982), “Marxism, Functionalism and Game Theory: The Case for Methodological Individualism”, *Theory and Society*, 11, pp. 453-82.

Epstein, J. M. & Axtell, R., (1996), *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge MA: MIT/Brookings Institution.

Field, A., (1981), “The Problem with Neoclassical Institutional Economics: A Critique with Special Reference to the North/Thomas Model of Pre-1500 Europe”, *Explorations in Economic History*, 18, pp. 174-198.

Field, A., (1984), “Microeconomics, Norms and Rationality”, *Economic Development and Cultural Change*, 32, pp. 683-711.

Galbraith, J., (1958), *The Affluent Society*. Boston: Houghlin Milton.

Galbraith, J., (1967), *The New Industrial State*. London: Hamilton.

Galbraith, J., (1973), *Economics and the Public Purpose*. Boston: Houghlin Milton.

Gallegati, M. Giullioni, G., Kirman, A., & Palestrini, A., (2011), “What’s that got to do with the price of fish? Buyers behaviour on the Ancona fish market”, *Journal of Economic Behavior and Organization*, 90 (1), pp. 20-33.

Gallegati, M. G & Richiardi, M., (2010), “Agent-based models in Economics and Complexity” in *Complex Systems in Finance and Econometrics* (ed. Meyers, R. A.). New York, NY: Springer.

Gaspard, M., (2003), "Ramseys' Theory of National Saving: A Mathematician in Cambridge", *Journal of the History of Economic Thought*, 25 (4), pp. 413-435.

Goodwin, R., (1951), "The Nonlinear Accelerator and the Persistence of Business Cycles," *Econometrica*, 19, pp. 1-17.

Hahn, F. H., (1984), *Equilibrium and Macroeconomics*, Oxford: Basil Blackwell.

Hardle, W. and Kirman, A., (1995), "Non-classical demand: a model-free examination of price quantity relations in the Marseille fish market", *Journal of Econometrics*, 67, pp. 227-257.

Harsanyi, J., (1968), "Individualistic and Functionalistic Explanations in the Light of Game Theory: The Example of Social Status", in *Problems in the Philosophy of Science* (eds. Lakatos, I. & Musgrave, A.), pp. 305-348. Amsterdam: North Holland.

Hart, H. & Honoré, A., (1985), *Causation and the Law*. Oxford: Clarendon Press.

Hempel, C. & Oppenheim, P., (1965), "On the Idea of Emergence" in Bedau, M & Humphreys, P. (2008), *Emergence: Contemporary Readings in Philosophy and Science*. Cambridge, Mass.: MIT Press, pp. 61-67.

Hicks, J. R., (1980), "IS-LM: An explanation", *Journal of Post-Keynesian Economics*, 3(2), pp. 139-154.

Hildenbrand, W., (1994), *Market Demand: Theory and Empirical Evidence*. Princeton, N.J.: Princeton University Press.

Hitchcock, C., (2003), "Of Humean Bondage", *The British Journal of the Philosophy of Science*, 54, pp. 1-25.

Hobbes T., (1968), *Leviathan*. Harmondsworth: Penguin, [1651].

Hodgson, G. M., (2006), "What are institutions?", *Journal of Economic Issues*, 40(1), pp. 1-25.

Hodgson, G. M., (2007), "Meanings of methodological individualism", *Journal of Economic Methodology*, 14(2), pp. 211-226.

Horgan, J., (1997), *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. New York: Broadway Books.

Jevons, W., (1879), *Theory of Political Economy*. London: MacMillan [1871]

Keen, S., (2011), *Debunking Economics: The Naked Emperor Dethroned?*, 2nd Ed.. New York: Zed Books.

Keynes, J.M., (1936), *The General Theory of Employment, Interest and Money*. Macmillan, London.

Kincaid, H., (1986), "Reduction, explanation, and individualism", *Philosophy of Science*, 53 (4), pp. 492-513.

Kirman, A., (2007), "The basic unit of economic analysis: Individuals or markets? A comment on "Markets come to bits by Phil. Mirowski", *Journal of Economic Behavior and Organization*, 63, pp. 284-294.

Kirman, A., (2011a), *Complex Economics: Individual and Collective Rationality*. London: Routledge.

Kirman, A., (2011b), "The Economic Entomologist: An interview with Alan Kirman", *Erasmus Journal for Philosophy and Economics*, 4(2), pp. 42-66.

Kirman, A. & Vriend, N., (2001), "Evolving market structure: An ACE model of price dispersion and loyalty", *Journal of Economic Dynamics and Control*, 25, pp. 459-502. Kirman & Vriend (2001),

Kreps, D. M., (1990), *A course in microeconomic theory*. Princeton, NJ: Princeton University Press.

Krugman, P., (1996a), *The Self-Organizing Economy*. Cambridge, MA: Blackwell Publishing.

Krugman, P., (1996b), *What economists can learn from evolutionary theorists*, talk delivered to the European Association for Evolutionary Political Economy, Nov 1996, available at <http://web.mit.edu/krugman/www/evolute.html>.

Krugman, P., (2000), "How complicated does the model have to be?", *Oxford Review of Economic Policy*, 16(4), pp. 33-42.

Kuhn, T. S., (1970), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Koopmans, T., (1947), "Measurement without Theory", *The Review of Economics and Statistics*, 31, pp.86-91.

Koopmans, T., (1965), "On the Concept of Optimal Economic Growth." In *The Econometric Approach to Development Planning*. Amsterdam: North Holland, pp. 225–300.

Kuznets, S., (1930), *Secular Movements in Production and Prices: Their Nature and their Bearing upon Cyclical Fluctuations*. Boston, Mass. and New York: Houghton Mifflin.

Kuznets, S., (1955), "Economic Growth and Income Inequality", *American Economic Review*, 65, pp. 1-28.

Lewis, D., (1973), "Causation", *Journal of Philosophy*, 70, pp. 556-567.

Lewes, G., (1875), *Problems of Life and Mind*, Vol. 2. London: Kegan Paul, Trench, Turbner, & Co.

Lehtinen, A. & Kuorikoski, J., (2007), "Computing the Perfect Model: Why Do Economists Shun Simulation?", *Philosophy of Science*, 74 (3), pp. 304-329.

Lucas, R. E., (1976), "Econometric Policy Evaluation: A Critique", pp. 19-45 in K. Brunner and A.H. Meltzer (eds) *The Phillips Curve and Labor Markets*. Amsterdam: North-Holland.

Mackie, J., (1974), *The Cement of the Universe*. Oxford: Oxford University Press.

Mantel R. R., (1974), "On the characterization of aggregate excess demand", *Journal of Economic Theory*, 7 (3), pp. 348-353.

Mandeville, B., (1714/29/32), *The Fable of the Bees, or Private Vices, Public Benefits*. Indianapolis: Liberty Fund, [1988].

Marx, K., (1847), *The Poverty of Philosophy*, English translation in (1977) Karl Marx and Friedrich Engels: *Collected Works*. New York: Progress Publishers.

Maudlin, T., (2004), "Causation, Counterfactuals and the Third Factor", in Collins, J., Hall, N & Paul, L. (eds.), *Causation and Counterfactuals*. Cambridge, Mass.: MIT Press, pp. 419-443.

McLaughlin, B., (1992), "The Rise and Fall of British Emergentism", in Bedau, M & Humphreys, P. (2008), *Emergence: Contemporary Readings in Philosophy and Science*. Cambridge, Mass.: MIT Press.

Mellor, D., (1982), "The Reduction of Society", *Philosophy* Vol. 57, No. 219, pp. 51-75.

Menger, C., (1976), *Principles of Economics*, Trans Dingwall, J. & Hoselitz, L.. Urbana: University of Illinois Press [1871].

Menzies, P., (2007), "Causation in Context", in Corry, R. and Price, H. (eds.), *Russell's Republic*. Oxford: Oxford University Press, pp.191-223.

Menzies, P. & Price, H., (1993), 'Causation as a Secondary Quality', *British Journal for the Philosophy of Science*, 44, pp. 187-204.

Mill, J. S., (1843), *System of Logic*. London: Longmans, Green, Reader, and Dyer. (Eighth edition, 1872).

Mill, J. S., (1869), *Thornton on Labour and its Claims*, in *Collected Works*, 1967, Essays on economics and society. Toronto: Thornton University Press, pp. 631-668.

Mirowski, P., (2007), "Markets come to bits: Evolution, computation and markomata in economic science", *Journal of Economic Behavior and Organization*, 63, pp. 209-242.

Mitchell, W., (1927), *Business Cycles: The Problem and Its Setting*. New York: National Bureau of Economic Research.

Mitchell, W., (1950), "Quantitative Analysis in Economic Theory", reprinted in *The Backward Art of Spending Money and Other Essays*. New York: Augustus M. Kelley, pp. 20-41 [1925].

Morgan, C. L., (1923), *Emergent Evolution*. London: Williams & Norgate.

Nagel, E., (1961), *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace and World.

North, D., (1981), *Structure and Change in Economic History*. New York: Norton.

North, D., (1990), *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

North, D. & Thomas, R., (1973), *The Rise of the Western World: A New Economic History*. Cambridge: Cambridge University Press.

Oppenheim, P. and Putnam, H., (1958), "Unity of Science as a Working Hypothesis", pp. 3-36 in H. Feigl, M. Scriven and G. Maxwell (eds) *Minnesota Studies in the Philosophy of Science*, vol. II, Minneapolis: Minnesota University Press.

Pareto, V., (1896), "La Courbe de la Repartition de la Richesse." Busino, Giovanni, ed., *Oevres Completes de Vilfredo Pareto*, 3, pp. 1-15, *Ecrits sur la Corbe de la Repartition de la Richesse*. Geneva: Librairie Droz, [1965].

Perona, E., (2007), "The Confused State of Complexity Economics: An Ontological Explanation" in Colander, D. & Solzano, M (Eds.) (2007), *Complexity Hints for Economic Policy*. Milan: Springer.

Persky, J., (1992), "Retrospective: Pareto's Law", *Journal of Economic Perspectives*, 2, pp. 767-73.

Picton, E., (2000), "Competition, rationality, and complexity in economics and biology" in Colander, D. (Ed.) (2000b), *Complexity and the History of Economic Thought*. London: Routledge.

Popper, K. R., (1944-5), "The Poverty of Historicism, I-III", *Economica*, 11, 86-103, 119-37; 12, 69-89.

Ramsey, F., (1928), "A Mathematical Theory of Savings" *Economic Journal*, 38, pp. 543-59.

Romer, D., (2006), *Advanced Macroeconomics*. New York: McGraw-Hill.

Rutherford, M., (1994), *Institutions in Economics: The Old and New Institutionalism*. Cambridge: Cambridge University Press.

Schaffer, J., (2005), "Contrastive Causation", *The Philosophical Review*, 114 (3), pp. 327-358.

Schotter, A., (1981), *The Economic Theory of Social Institutions*. Cambridge: Cambridge University Press.

Schelling, T. C., (1978), *Micromotives and Macrobehavior*. New York: W. W. Norton and Co.

Schumpeter, J. A., (1909), "On the Concept of Social Value", *Quarterly Journal of Economics* 23, pp. 213-32.

Smale, S., (1976), "Exchange Processes with Price Adjustment", *Journal of Mathematical Economics* 3, pp. 211–26.

Smith, A., (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*. Chicago: Chicago University Press [1976].

Sonnenschein, H., (1972), "Market Excess Demand Functions", *Econometrica* ,40, pp. 549-563.

Sperry, R., (1964), *Problems Outstanding in the Evolution of Brain Function*, James Arthur Lecture. New York: American Museum of Natural History.

Spinner, H., (1973), 'Science Without Reduction: A Criticism of Reductionism with Special Reference to Hummel's and Opp's "Sociology Without Sociology"', *Inquiry*, 16, pp. 16-94.

Steglich-Peterson, A., (2012), "Against the Contrastive Account of Singular Causation", *British Journal of the Philosophy of Science*, 63, pp. 115-143.

Udehn, L., (2001), *Methodological Individualism: Background, History and Meaning*. London: Routledge.

Udehn, L., (2002), "The Changing Face of Methodological Individualism", *Annual Review of Sociology*, 28, pp. 479-507.

Thornton, W.H., (1870), *On Labour, its Wrongful Claims and Rightful Dues, its Actual Present and Possible Future*, 2nd ed. McMillan, London.

Varian, H., (2006), *Intermediate Microeconomics*, 7th edition. New York, NY: W.W. Norton & Co..

Veblen, T., (1898), "Why is Economics not an Evolutionary Science?" *The Quarterly Journal of Economics*, 12 (4), pp. 373-397.

Veblen, T., (1899), *The Theory of the Leisure Class*. New York: Dover [1994].

Veblen, T., (1900), "Preconceptions of Economic Science, part III", *Quarterly Journal of Economics*, 14, pp. 240-269.

Veblen, T., (1909), "The Limitations of Marginal Utility", reprinted in *The Place of Science in Modern Civilisation*. New York: B. W. Huebsch [1961].

Vining, R., (1949), "Methodological Issues in Quantitative Economics", *Review of Economics and Statistics*, 31, pp. 77-86.

Walras, L., (1954), *Elements of Pure Economics*, Trans. Jaffé, W. Homewood. IL: Richard D. Irwin [1874].

Watkins, J. W. N., (1952), "Ideal Types and Historical Explanation", *British Journal of Philosophy of Science*, 3, pp. 22-43.

Weber, M., (1922), *Economy and Society*, 2 Vols., Berkeley: University of California Press, [1978].

Wible, J., (2000), "What is Complexity?", in Colander, D. (Ed.) (2000b), *Complexity and the History of Economic Thought*, London: Routledge.

Woodward, J., (2003), *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Zipf, G.K., (1949), *Human Behavior and the Principle of Least Effort*, Cambridge, MA: Addison-Wesley Press.