Introduction

This book follows on from, and at various points exploits, my earlier work *On Sociology* (2nd edn 2007). It has, however, a significantly different character. *On Sociology* was a collection of rather diverse essays that were brought together under the headings of 'Critique and Program' (volume 1) and 'Illustration and Retrospect' (volume 2). These essays were mainly written around the turn of the century – a time of intense debate over how sociology as an academic and intellectual enterprise should be viewed and of great uncertainty over the future course of its development. During more recent years I have become aware of a significantly changing situation. Some at least of the kinds of sociology that I earlier criticised – for example, 'grand' historical sociology and 'post-modernist' ethnography – would appear to have fallen into decline; and, of greater consequence, the version of sociology that I argued for programmatically and sought to illustrate has, at least in certain respects, flourished to a degree that I find surprising – although, of course, pleasantly so. I would naturally like to see evidence here of the influence of *On Sociology*; but, as a good Popperian, I have to accord crucial importance to the 'logic of the situation'. Research designed to address well-defined sociological problems and based on the quantitative analysis of extensive and high-quality data-sets – even if not backed by theoretical advance to the extent I might wish – has been increasingly recognised as having premium payoffs, in both its 'pure' and 'applied' aspects, and has in turn become increasingly attractive to working sociologists and to funding agencies alike.[1]

[1] The main exception to this general statement unfortunately arises with my own country, Great Britain, where, especially within university departments as distinct

Consequently, there now appears to be less need than previously for critical or programmatic interventions, and I would, in turn, wish to emphasise the following point regarding the present work. In seeking to make out the case for an understanding of sociology as a population science, my main concern is not to propose to sociologists how they should conceive of and practise their subject. It is rather to suggest a way in which a fuller and more explicit rationale than has hitherto been available might be provided *for what a large and steadily growing number of sociologists in fact already do* – although, perhaps, without a great deal of reflection on the matter. If asked what purpose the elaboration of such a rationale might serve, my response would be twofold. First, I believe that some greater awareness on the part of the sociologists in question of what it is that they are about should enable them to proceed more systematically and effectively in their everyday work. And second, I believe that, as well as helping to create such awareness, an understanding of sociology as a population science affords the best basis on which these sociologists can articulate and pursue a goal in which, I believe, they would largely share: that is, the goal of developing sociology as a science in a sense that allows for a meaningful degree of continuity with the natural sciences while still preserving sociology's proper distinctiveness.

It may well be that some of the sociologists I have here in mind will not be ready to accept my interpretation of the sociology in which they engage as constituting, or even as moving in the direction of, a

from – often interdisciplinary – research centres, a strong hostility to quantitative sociology persists. Interestingly, my earlier book, while widely reviewed elsewhere, whether in its original form or in later Italian, Polish and Spanish versions, was not reviewed in *Sociology*, the official journal of the British Sociological Association, nor in the *Sociological Review*. But even perhaps in Britain, 'the times they are a-changin'. The Q-Step programme, launched in 2013, with a budget of £19.5 million, aims to substantially extend and revitalise the quantitative training of social science undergraduates. It will, I hope, achieve its goals in sociology, despite efforts that are, apparently, in train to subvert it. A number of remarkably ill-informed claims have of late been made to the effect that quantitative methods of a kind labelled as 'conventional' are now largely outmoded and irrelevant and should be replaced by others (e.g. Byrne, 2012; Castellani, 2014). Some of the alternative methods proposed are critically considered in the course of the chapters that follow.

population science, nor again my view that this represents the most promising way ahead for a scientific sociology. Responses to my book made on these lines I would regard as very welcome ones – provided that they are accompanied by alternative interpretations of how sociology is in fact developing as a science and by some indications of how this development might best be furthered. Discussion of the issues that are likely to arise in this connection could be of particular value at the present time.

I do, of course, also recognise that there are many other sociologists who would disagree with me in a more fundamental way: that is, in doubting that sociology can claim scientific status and indeed in believing that it is not even desirable that it should try to do so. These sociologists I would regard as selling sociology short – as standing in the way of it realising its full potential – and I can have little common ground with them. Moreover, I can now see little point in engaging further in the already protracted debates that have taken place on this matter: the future will decide.

In writing this book, I have aimed at clarity and brevity. In the interests of clarity, the book is structured around nine propositions. One of these propositions stands at the head of each of the central chapters, and the chapter itself is then given over to elaborating and supporting the proposition. Readers who would like to have an initial overview of the argument of the book can simply read through the propositions. In advancing the case for sociology as a population science, I have found it necessary to cover a rather wide terrain and to refer to literature from a range of different fields apart from that of sociology itself. However, in the interests of brevity I have in general indicated only the essentials of the way in which I would see the work on which I draw as being relevant to the positions I take up, and I have then left it to readers to use the citations given in order to check, if they so wish, that my use of these sources is appropriate. It may be noted that, for a short book, there is a rather long set of references.

In one respect, I have given some weight to clarity as against brevity. Insofar as the arguments I advance are of a general and abstract

character, I have tried to bring out more clearly the main points that they seek to make by providing particular and concrete illustrations. In the case of specifically sociological arguments, these illustrations may perhaps be thought to be too often taken from my own fields of research interest, in particular social stratification and mobility. But, to the extent that my knowledge has allowed, I have entered into other fields as well.

The book is aimed primarily at professional sociologists and at more advanced students. I have therefore assumed a certain amount of background knowledge, including some basic technical knowledge in relation to methods of data collection and analysis. However, I have kept the text itself as non-technical as possible: there are no formulae or equations. At the same time, though, I have found expository advantage, in particular in the chapters dealing with issues of data collection and analysis, in taking a historical approach. Robert Merton once complained (1957: 4) that in discussion of sociological theory, too much attention was given to history at the expense of what he called 'systematics'. But as regards the discussion of research methods in sociology, one could almost make the opposite complaint. That is to say, too little attention is given to why methods presently in use are as they are. Why did these methods emerge? What preceded them? What were the problems for which they provided better solutions, and how? Addressing such questions seems to me to be often highly illuminating.

In completion of this Introduction, there are two further observations of a more personal kind that I might make. The first refers back to my earlier remark that over recent years I have become aware of a significant change in sociology in the research styles that are assuming prominence – a change that I find highly congenial and that led me to believe that there might be some point in a book of the kind I have now written. I would like here further to say that the main context within which this change was borne in on me was that provided by the *European* sociological research community: in particular, by conferences and workshops organised by the

European Consortium for Sociological Research and also under the auspices of two EU-funded 'Networks of Excellence' in sociology – CHANGEQUAL and its successor, EQUALSOC.[2]

With American readers chiefly in mind, I might add here that there has been some tendency within American sociology to see the most distinctive European contributions to the subject as being made at rather rarefied levels of theory or at the intersection of methodological and philosophical issues (with an attendant exaggeration of the importance of some, chiefly French and German, authors). While this view was always questionable, it is by now quite clearly out of date. Over the last twenty years or so there has been a rather remarkable expansion of sociological research, of a largely quantitative character, in almost all of the major European countries – research that is of a technical standard quite comparable to that of American work and that is often of at least potentially greater theoretical interest, in having a comparative cross-national or cross-regional basis.[3] I frequently draw on this body of research for my illustrative purposes. In addition, it could be noted that it is European sociologists who have been prominent in developing the 'mechanism-based' approach to causal explanation that, as discussed especially in Chapter 9, I would regard as most appropriate to sociology understood as a population science.

[2] The partner institutions in the CHANGEQUAL network were the Economic and Social Research Institute, Dublin; the Centre National de la Recherche Scientifique EHSS LASMAS, Paris; the Swedish Institute for Social Research, University of Stockholm; the Zentrum für Europäische Sozialforschung, University of Mannheim; and my own institution, Nuffield College, Oxford. In the EQUALSOC network, the CNRS institution became GENES/GRECSTA and eight further institutions were added: the Institute for Advanced Labour Studies, University of Amsterdam; the Centre for Social Policy, Antwerp; the Universita Degli Studi di Milano Bicocca; the Department of Political and Social Science, University of Pompeu Fabra, Barcelona; the Department of Sociology and Social Policy, University of Tartu; the Department of Social Sciences, University of Turin; the Department of Sociology and Social Research, University of Trento; and the Wissenschaftzentrum für Sozialforschung, Berlin.

[3] The British exception has again to be noted. At the conferences and workshops referred to in the text, the virtual absence of young British researchers has been sadly apparent.

My second, more personal observation is the following. I have written this book somewhere towards the end of a rather long life in academia (a circumstance in itself favouring brevity), and my views have obviously been in various respects influenced by my own experiences over the years in question. Indeed, in some instances I have made this quite explicit, and especially where I have to acknowledge the influence of a teacher or a colleague. I would like to think that in this way also the book benefits from a historical perspective – and one that is, I would suggest, much needed in order to offset sociology's rather manifest lack of collective memory, leading to an unfortunate neglect of the deeper origins of current problems and often in turn to the rediscovery of wheels. But I realise that I might equally well be regarded as harking back unduly to issues that have been long forgotten, and with good reason.[4] At all events, examples of what might be regarded as my anecdotage have been largely confined to the notes.

[4] A reviewer of a paper that a similarly aged colleague and I recently submitted to a leading sociological journal objected to the fact that articles were cited that were published before he or she was born. Clearly, nothing of importance could have happened before that date.

1 Sociology as a population science: the central idea

Sociology should be understood as a population science in the sense of Neyman (1975).

On the occasion of the 500th anniversary of the birth of Nicholas Copernicus (1473–1543), the US National Academy of Sciences sponsored a collection of essays on 'quasi-Copernican' scientific revolutions. The volume was edited by the Polish-born statistician Jerzy Neyman, who supplied brief introductions to its several sections. In one such introduction, to a series of essays on 'The Study of Chance Mechanisms – A Quasi-Copernican Revolution in Science and Mathematics', Neyman (1975: 417) made the following observation:

> Beginning with the nineteenth century, and increasing in the twentieth, science brought about 'pluralistic' subjects of study, categories of entities satisfying certain definitions but varying in their individual properties. Technically, such categories are called 'populations'.

Neyman emphasised that populations in this technical sense could, substantively, be of quite different kinds. They could be human or other animal populations, but also populations of, say, molecules or galaxies. The common feature of such populations was that, while their individual elements were subject to considerable variability and might appear, at least in some respects, indeterminate in their states and behaviour, they could nonetheless *exhibit aggregate-level regularities of a probabilistic kind*.[1]

[1] I was first directed to Neyman's remarks by a reference to them in Duncan (1984: 96). As will become readily apparent, Dudley Duncan is an author to whom I am

The aims of a science dealing with such pluralistic subjects of study – or, that is, of what could be called a 'population science' – were then twofold. The initial aim was to investigate, and to establish, the probabilistic regularities that characterise a particular population, or its appropriately defined subpopulations. In this regard, Neyman saw the use of statistical methods of both data collection and analysis as being essential. And indeed, fifty years previously, R. A. Fisher (1925: 2) had already *defined* statistics as 'the study of populations, or aggregates of individuals', and had represented statistics as foundational for all sciences that were primarily concerned with the properties of aggregates rather than of their individual members. It may moreover be noted, in view of what is to follow, that Fisher then added the remark that 'Statistical methods are essential to social studies, and it is principally by the aid of such methods that these studies may be raised to the rank of science.'[2]

However, Neyman also made it clear that once population regularities had been empirically established, the further aim of a population science had to be that of determining the processes or 'mechanisms' which *in their operation at the individual level* actually produced these regularities. And since the regularities – the explananda of a population science – were probabilistic, the mechanisms that would need to be envisaged would be ones that, rather than being entirely grounded in deterministic laws, *incorporated chance*. A new form of scientific explanation was implied.

Neyman's claim that from the nineteenth into the twentieth century the increasing study of 'pluralistic' entities on a statistical

indebted in many other respects. He must be regarded as one of the great pioneers in conceptualising and practising sociology as a population science. Another who contributed significantly, although in a less explicit way, was my former teacher at the London School of Economics, David Glass – now shamefully disremembered in British sociology – under the influence of his own teacher, the extraordinary polymath Lancelot Hogben (see Hogben, 1938).

[2] Neyman and Fisher were of course the leading antagonists in what has been described as 'the widest cleft in statistics' over issues of hypothesis-testing. But, as Louçã (2008: 4) has observed, in their vision of statistics as the language for a new form of science, they were in fact 'quite close'.

basis marked a scientific revolution has been amply justified by later work in the history of science. What has in fact become known as the 'probabilistic revolution' (Krüger, Daston and Heidelberger, 1987; Krüger, Gigerenzer and Morgan, 1987) is now widely recognised as one of the most – if not the most – momentous intellectual developments of the period in question. 'In 1800', to quote Hacking (1987: 52), 'we are in the deterministic world so aptly characterised by Laplace. By 1936 we are firmly in a world that is ultimately indeterminate . . . Chance, which, for Hume, was "nothing real" was, for von Neumann, perhaps the only reality.'[3] However, as Hacking goes on to stress (see also Hacking, 1990), it is important to see that complementary to 'the erosion of determinism' was 'the taming of chance': that is, the process of making chance and its consequences intelligible and manageable on the twin bases of assemblages of numerical data and the application of probability theory.

In the early stages of the probabilistic revolution, the social sciences did in fact play a leading part. In particular, Quetelet's application (1835/1842, 1846, 1869) of the Gaussian 'error curve' – or the normal distribution – to the display of regularities in the 'moral statistics' of marriage, illegitimacy, suicide and crime represented a pioneering attempt to show how a higher-level probabilistic order could emerge from out of individual actions that were generally supposed to be non-deterministic in character, or, that is, to express individual will and choice (Porter, 1986: chs 2, 6 esp.). And the notable development then was that the influence of Quetelet's work extended *from* the social *into* the natural sciences – somewhat ironically, given his great ambition to create a 'social physics'. As Krüger (1987: 80) has observed, at this point 'the familiar hierarchy of the disciplines' was inverted.

[3] Hacking is here referring to von Neumann's mathematical formulation of quantum theory. This aimed to preclude the possibility of 'hidden variables' that, if identified, would allow for phenomena that otherwise appeared probabilistic to be understood as deterministic – so that particles possessed a definite position and velocity at all times. For an accessible account, see Kumar (2008: ch. 14).

Most notably, Quetelet's use of the error curve provided a model for James Clerk Maxwell in his development of the kinetic theory of gases (Gillispie, 1963; Porter, 1982). Within a gas, the lower-level processes of colliding molecules were, in principle, subject to deterministic Newtonian laws; but the vast numbers of molecules involved meant that, in practice, a probabilistic treatment – 'statistical physics' – was required. In work carried out in the later 1860s, Maxwell took a version of the error curve to represent the *distribution* of molecular velocities within an ideal gas, so that, while nothing could be said about individual molecules, it became possible to calculate the *proportion* of molecules with velocities within a given range at any given temperature. Maxwell was generous in his acknowledgement of his borrowing in this regard from Quetelet and his followers. In speaking to the British Association for the Advancement of Science, he referred to physicists adopting a method of analysis new to them but which 'has long been in use in the section of Statistics' (cited in Gigerenzer et al., 1989: 62; see also Mahon, 2003: ch. 6).[4]

Subsequently, Fisher (1922), in seeking to integrate Mendelism into Darwin's evolutionary theory, adopted a model closely analogous to that which Maxwell had taken over from Quetelet, with biological populations corresponding to the populations of molecules. Under this model, natural selection could be seen as operating amid a multiplicity of random causes – any of which might have a predominant influence at the level of a particular individual – while, however, the probabilistic processes of natural selection remained the key determinants of the evolution of the population as a whole (Morrison, 2002).

In association with such developments, evolutionary biology, as Ernst Mayr (2001; see also 1982: ch. 2) has described, became the field in which the most explicit development of 'population thinking' occurred. In a scientific world dominated by physics and chemistry, what Mayr characterises as 'typological thinking' had prevailed,

[4] Ludwig Boltzmann, another pioneer of statistical physics, was also influenced by the work of Quetelet and his followers and expositors (Porter, 1986: 125–8).