4 Population regularities as basic explananda

For sociology understood as a population science, the basic explananda are probabilistic population regularities rather than singular events or events that are grouped together under some rubric but without any adequate demonstration of the underlying regularities that would warrant such a grouping.

Elster (2007: 9) has argued that 'The main task of the social sciences is to explain social phenomena' and that 'The basic type of explanandum is an *event*.' From the standpoint of sociology as a population science, Elster's argument needs to be qualified in one important respect. The events with which sociology is concerned are ones of a certain kind: that is, events that can be shown to occur within a given population or subpopulation *with some degree of regularity*.

Most of the explananda, or 'puzzles', for sociology that Elster (2007: 1–5) suggests by way of illustrating his argument do in fact relate to regularities in events: for example, 'Why are poor people less likely to emigrate?' and 'Why does an individual vote in elections when his or her vote is virtually certain to have no effect on the outcome?' However, some further instances that he gives refer to singular events: for example, 'Why did President Chirac call early elections in 1997, only to lose his majority in parliament?'[1] And, more

[1] Elster also gives examples which, while referring to singular events, are, it seems, to be understood as particular instances of regularities. For example, 'Why did none of thirty-eight bystanders call the police when Kitty Genovese was beaten to death?' (Kitty Genovese was a New York bar manager who was murdered while on her way home in the early hours of the morning one day in 1964.) What, one supposes, Elster is concerned with here is the so-called 'bystander effect', much discussed in the social-psychological literature, which could be formulated in general terms as follows: 'The probability of individuals providing help or taking other action in apparently emergency situations varies inversely with the numbers of those present.' Actually, the Genovese case is not a well-documented instance of this effect (Manning, Levine and Collins, 2007).

importantly for present purposes, there are other authors (e.g. Brady, Collier and Seawright, 2006; Mahoney and Goertz, 2006; Mahoney and Larkin Terrie, 2008) who recognise more explicitly than Elster the distinction between regularities in events and singular events and who still maintain that the social sciences should be as much concerned with explaining the latter as the former.

To see why difficulties arise when distinctive, singular events, rather than regularities in events, are taken as sociological explananda, it is necessary to consider further the role of *chance* in social life. In this regard, a distinction between two different understandings, or usages, of 'chance' in a scientific context, which has been proposed by the biologist Jacques Monod (1970), is of value: that is, the distinction between 'operational' and 'essential' chance.

Monod observes that chance is invoked in an operational sense when, in dealing with certain phenomena, a probabilistic approach is, in practice, the only methodologically feasible one – even if, in principle, a deterministic approach might be applicable. It is then chance in this operational sense that is basic to the idea of a population science as understood by Neyman. The infeasibility is accepted of seeking to account in a deterministic way for the states and behaviour of all individuals making up a population, whether because they are inherently indeterminate or simply because of the degree of complexity of the determination involved. Nonetheless, the possibility still exists of – to revert to Hacking's phrase – 'the taming of chance' by establishing regularities of a probabilistic kind at the aggregate, population level, and by then seeking explanations of these regularities as resulting from causal processes or mechanisms operating at the individual level which incorporate chance.

In contrast, essential chance is, for Monod, a far more radical idea and applies where some outcome results from the *intersection* of two or more quite independent series of events (see Hacking, 1990: 12). In the example Monod (1970: 127–31) gives, Dr Dupont goes out on an emergency call but, as he passes by a building where repairs are being carried out, roofer Dubois drops his hammer, which falls on Dr

Dupont's head and kills him. Even if the two series of events are themselves seen as in some way determined, their independence means that their joint outcome has still to be understood as a '*coïncidence absolue*'.[2]

Now, in human social life the operation of essential chance might seem pervasive. Individuals do often find themselves, like Dr Dupont, in the wrong place at the wrong time, or, more happily, in the right place at the right time. However, essential chance here operates in a context in which forces making for regularity are also present. Individuals pursue their diverse ends, often in situations of great uncertainty, but in an informed way guided by a common rationality and under various in part shared normative and non-normative constraints. What might then appear at first sight as essential chance at work can often be shown to be at least in some degree socially conditioned.

For example, in early work, Jencks (1972) emphasised the role of 'sheer luck' in relative success or failure in economic life; but, as he himself came later to accept (Jencks, 1979), individuals' experience of such luck, whether good or bad, can be significantly influenced by their social milieux. And as Granovetter's work has served to show, the occurrence of one kind of luck cited by Jencks – 'chance acquaintances who steer you to one line of work rather than another' – is very likely to be conditioned by features of individuals' social networks, so that it is possible to 'pursue a systematic analysis of this variety of "luck" by placing it in a social structural context' (Granovetter, 1995: xi).

Thus, in analyses made at the population level – that is, covering relatively large numbers of individuals – probabilistic regularities in

[2] As a biologist, Monod's main concern was to establish the essentially chance element in evolution. The processes through which mutations in DNA sequences occur, he sought to show, have no connection with – are quite independent of – the effects that follow from the modified protein, the interactions it ensures, the reactions it catalyses and so on. At King's College, Cambridge during the 1960s, I benefited greatly from several conversations with Jacques Monod on 'chance and necessity' in biology and in social life.

social life of many kinds do still emerge despite the pervasiveness of essential chance – though often they are regularities of a complex and not readily visible, let alone transparent, kind. And it is then these regularities that, under the auspices of operational chance, can be treated as the explananda for which sociological explanations may properly be sought.

In contrast, with distinctive, singular events, essential chance tends to take on a far more dominant role and to operate in a way that is far less easily 'tamable'. Although it may be possible, at least after the fact, to suggest certain prevailing regularities and perhaps underlying causal mechanisms that could have been conducive to such events, in their actual occurrence the – inherently improbable – intersection of preceding events tends to be involved, and often quite crucially so. It is this fact that appears to be overlooked by those who urge that sociologists should seek to explain singular events or singular complexes of events, such as – to take examples from Mahoney and Goertz (2006: 230) – the outbreak of the two world wars or the collapse of the Soviet Union. In arguing thus, these authors claim that natural scientists are very ready to apply their theories to account for 'particular outcomes', and they give as an example the explanation provided by the physicist Richard Feynman for the NASA shuttle *Challenger* disaster of January 1986. However, this example actually serves rather well to bring out the weakness of the position that Mahoney and Goertz take up.

The immediate cause of the disaster at the launch of *Challenger* was the failure of the rubber seals on a solid rocket booster joint, through their loss of resilience at low temperatures. And the physics of this failure were indeed dramatically demonstrated by Feynman. At a session of the Presidential Commission of Inquiry into the disaster, he soaked a sample of the rubber used for the seals in a glass of iced water on his desk – and then snapped it. But the problem of the resilience of the seals was in fact well-known, and what was crucial was that during a teleconference the night before the launch, confusion occurred over the relationship that existed between air temperature and the probability of the seals losing their effectiveness. The statistical data hurriedly considered came only from previous launches in which some damage to the seals had actually occurred, and analyses of these data did not strongly indicate that, with the temperature that was forecast for the morning, the launch should be postponed – whereas what subsequently emerged was that *if data from all launches had been examined*, the results would have far more clearly shown that a major risk existed (Dalal, Fowlkes and Hoadley, 1989). After much argument, the launch was allowed to go ahead and the seals failed.

The key point to note here is that while sociological theory could perhaps, as claimed by Vaughan (1996), be of help in explaining prevailing contextual features that made it more likely that a shuttle disaster of some kind *might* happen – for example, NASA's organisational culture and its 'normalisation of deviance' in regard to safety issues – it could scarcely lead to an explanation for the fact that this particular disaster *did* happen. As Popper (1957: 116–17, 143–7) has maintained, with due acknowledgement to Max Weber (1906/1949), when, in order to explain some 'actual, singular, or specific event', several different causal processes have to be invoked, with different theoretical groundings, the explanation that results *will not itself be a theoretical one*. It will be an explanation of a quite different kind: that is, a *historical* explanation, which involves an account of the unfolding of all relevant prior events, including their quite contingent intersections and their consequences – essential chance at work – up to the point at which the event of interest was actually brought about. In short, what is involved is a narrative of a highly specific, place- and time-linked kind, and one that is necessarily given *ex post*.[3]

[3] In her book on the *Challenger* launch decision, Vaughan (1996: xiii) claims that she provides 'a sociological explanation' of this decision, but also, on the same page, that she provides 'a historical ethnography' of the sequence of events that led up to it. The latter claim is more compelling than the former. Vaughan's detailed account of this sequence of events (see ch. 8 especially) does in fact serve well to show how, at a number of different points, essential chance came crucially into play.

It may be noted that Mahoney has more recently associated himself with the view that historical explanations are indeed distinctive in being concerned with the causes of particular past occurrences, and that 'the question of whether and how the resulting explanation might then be generalised is a secondary concern' (Mahoney, Kimball and Koivu, 2009: 116). But, very strangely, this is in a paper devoted to the logic of historical explanation *in the social sciences* – in which, one might suppose, the search for theoretically grounded explanations than can extend beyond particular cases must be a *primary* concern.

Fortunately, despite the misguided encouragement that some would offer, social scientists do not all that often attempt to explain singular events, and thus the main importance of the foregoing argument actually arises in another case: that is, where sociologists seek to explain events or complexes of events that are grouped together under some rubric *as if* they were characterised by significant regularities, but where no compelling demonstration of this has been provided. For purposes of illustration of the difficulties that arise, I will take what is claimed to be the sociology of revolutions, although I could have equally well taken, say, the supposed sociology of economic crises or of various historical trajectories, such as, say, 'routes' to authoritarianism or democracy or 'paths' to modernisation.

In a review article, Goldstone (2003: 50, see also 1995) has maintained that 'steady progress' has been achieved in the sociology of revolutions. This progress has resulted from the application of essentially inductive procedures to detailed case studies of 'finite sets' of revolutions, rather than from the analysis of samples of revolutions taken from some 'pre-defined universe'. As studies of particular revolutions have accumulated, Goldstone claims, knowledge of the different causal processes that may be involved has grown, allowing wider-ranging and more secure generalisations about the occurrence of revolutions to be made, and, indeed, creating the possibility of their prediction. Goldstone (1995: 45) himself has sought to integrate this body of work into what he terms a 'conjunctural process model' of revolutions. This proposes that a society 'is careening [sic] toward revolution' when three conditions apply: (i) the state loses effectiveness in its ability to command resources and obedience; (ii) elites are alienated from the state and in heightened conflict over the distribution of power and status; and (iii) a large or strategic proportion of the population can be readily mobilised for protest actions.

As Tilly (1995: 139–40) has observed, these conditions are in themselves so close to *defining* an actual revolutionary situation as to make Goldstone's model of limited potential in explaining how revolutions come about. But what may further be objected is that what is empirically claimed does not even amount to a number of established regularities in the social processes leading up to revolutions that could constitute appropriate sociological explananda. For, as Goldstone appears to accept, his 'conjunctural' conditions for revolution *have no inherent tendency to come together*; whether or not they do so in particular cases has to be regarded as quite contingent. And he does indeed explicitly acknowledge that his model 'says nothing' about the causes of societies moving towards a revolutionary situation, and suggests that no fixed set of such causes exists (Goldstone, 1995: 45).

It is therefore difficult to see that any compelling argument is made out for the viability of a sociology of revolutions. If induction from successive case studies does not lead to the establishment of empirical regularities in the pre-conditions for revolutions, in their outbreak, in their development, in the factors associated with their success or failure and so on, but instead reveals ever-widening variation from case to case, then the possibility of some *theoretical* explanation of revolutions in terms of the systematic processes that are at work is clearly undermined.[4] And what, one may argue, is in turn indicated is that major weight must be given to factors that are specific

[4] I should stress here that although, as indicated later in the chapter, I have serious doubts about the methodology that Goldstone favours – that is, relying on induction from 'finite sets' of cases – I do not believe that this methodology is itself the source of the failure to establish empirical regularities in relation to revolutions. I would entirely agree with Goldstone (2003: 43) that large-N, sample-based studies of revolutions have 'not been terribly fruitful' and have in fact done no better in this regard.

to individual cases, including their interaction through the operation of essential chance. In other words, what would appear appropriate is not a sociology of revolutions but, at most, a comparative history, which, following Popper, would have to be recognised as an intellectual undertaking of a quite different kind. That is to say, one in which explanations advanced for different revolutions are compared, with due account being taken of the particular concatenations of events involved, and with a concern as much for the diversity of possible revolutionary processes as for common features.[5]

It is in this regard of interest to find that among professional historians studying revolutions, the tendency has in fact increasingly been to downplay regularities that might be supposed to exist across cases and to emphasise their individual distinctiveness. For example, the author of the most comprehensive study to date of 'the English Revolution' of the seventeenth century remarks in his concluding chapter that 'I am sceptical about the quest for a morphology of revolution that will accommodate the upheavals that began in France in 1789, in Russia in 1917, in China under Chairman Mao, and other later convulsions elsewhere' (Woolrych, 2002: 792).[6]

The position that I have taken up in the foregoing might be regarded as unduly negative: that is, as seeking to impose unnecessary limits on sociological ambition. But those who are inclined to such a view might wish to reflect on the fact that there is no reason to suppose that the range of sociological explanation is infinite – no

[5] Goldstone and others seeking to create a sociology of revolutions do sometimes alternatively describe their objective as that of providing a 'comparative historical analysis' of revolutions – but without any evident recognition of the differences that arise.

[6] See also the observations of the pioneering revisionist historian of the French Revolution, Alfred Cobban (1965: esp. chs 1–3), whom I was fortunate enough to have as a teacher. His work in particular serves to underline the point that scepticism of the kind that Woolrych expresses in no way precludes historians of revolutions from making use of sociological concepts, or indeed of sociological theory, that have been developed in other areas. But the distinction between historical and sociological explanation remains. Anyone doubtful of it, and thus inclined to follow Mahoney and Goertz in believing in the possibility of a sociological explanation of the outbreak of the First World War, would do well to read Clark's (2013) superb historical account, or at least pp. 361–4 and the Conclusion.

reason to suppose that there can be a sociology of anything and everything – and that it is therefore important to have some idea of where, and on what grounds, the *boundaries* of sociological explanation are to be drawn. It is moreover relevant to note that excessive ambition has its costs. Thus, after the events of 1989–90, sociologists did in fact have to face much criticism for their failure to anticipate the collapse of the Soviet Union – criticism that could have been avoided, or at least effectively rejected, if within the discipline a clearer awareness and more explicit recognition had been present of where a historical rather than a sociological explanation is called for. If this had been the case, then, as Hechter (1995: 1523) has observed, sociologists would have had no need 'to hang their heads in shame'. As for Goldstone's (1995) claim that, in the light of his model, he would in fact have been able predict the revolutions associated with the Soviet collapse, Runciman's (1998: 16) comment is brutal but to the point: 'So maybe you could, Jack. And if you could, you should. But you didn't.'[7]

As a coda to this chapter, some remarks may be apposite on the use of *logical* rather than statistical methods in sociological analysis, since the difficulties that arise in the application of logical methods serve to bring out in another way the dangers of pursuing sociological explanations where it is historical explanations that are required. It was in fact with Skocpol's (1979) early study of revolutions – specifically, with her use of John Stuart Mill's (1843/1973–74) 'method of agreement' and 'method of difference' – that logical methods came into prominence in sociology; and while Goldstone would appear to apply such methods in only an informal manner, others who would share in his commitment to working inductively from case studies have sought to develop them beyond Mill. Most notably, Charles Ragin (1987) has proposed the use of set theory and Boolean algebra in

[7] Economists have of course come under criticism for their failure to predict the financial crash of 2008, and it is notable that some at least of their number have taken the view that such prediction lies beyond the scope of economics as a social science, not least on account of the part likely to be played in events of the kind in question by historical specificities.

what has become known as 'qualitative comparative analysis' (QCA), and it is on this method that I focus.

QCA aims at showing the conditions under which a certain outcome does or does not occur, with the various conditions considered being themselves also treated as binary: in effect, as present or absent. On the basis of case studies, a 'truth table' is constructed, showing which sets of conditions are associated with the outcome occurring or not occurring, and Boolean algebra is then used in order to collapse the truth table into a minimal formula. In the strong version of QCA, this Boolean equation is taken as a 'causal recipe' that gives all of the combinations of conditions that are necessary and/or sufficient for the outcome in question to be realised. In a weaker version, which of late appears more often to be advanced – if with a good deal of equivocation (see e.g. Rihoux and Marx, 2013: 168–9) – the Boolean equation is represented as simply a means of summarising results from a number of case studies in a form that invites, or, at most, could suggest, causal explanations. In this weaker version, QCA might then appear as a possible way of establishing empirical regularities in regard to some social phenomenon that would in turn constitute sociological explananda of a legitimate kind.

However, whether QCA is understood as a source of causal recipes or as an essentially descriptive method, it is open to serious objections on grounds that have been most cogently expressed by Lieberson (2004) and by Lucas and Szatrowski (2014). What these authors are concerned to stress is that QCA as a method of logical rather than statistical analysis must assume a quite deterministic rather than a probabilistic social world.[8] That is to say, it can make

[8] In a later work, Ragin (2000) has moved beyond his original formulation of QCA in proposing that in place of 'crisp' sets, implying strictly binary categorisations, 'fuzzy' sets may be used. This new approach entails a significant shift away from logical and towards statistical analysis, in that it involves measurement, albeit rather crude and often arbitrary, of the degree to which cases belong to particular sets; and it can also be used to implement a probabilistic rather than a deterministic understanding of causation, although at cost of such oxymorons as 'almost necessary' and 'nearly sufficient'. I do not therefore consider fuzzy-set QCA in the context of the present

no allowance for the operation of essential chance in this world, nor, moreover, for chance simply in the form of error in our – supposed – knowledge about it (see further Goldthorpe, 2007: vol. 1, ch. 3; Hug, 2013). In turn, therefore, QCA can take no account of the extent to which a truth table derived from case studies may contain results that are in effect *random*: that is, in consequence of the social world not in fact being deterministic or in consequence just of data error. And for this reason, Boolean summary equations can then easily lead, as Lieberson (2004) argues, to 'massive over-interpretation', or, as Lucas and Szatrowski (2014) would claim, to causal explanations that are simply mistaken.[9] Another way of putting the point here being made would be to say that QCA may often be incapable of distinguishing signal from noise. And Lieberson has indeed demonstrated the possibility that QCA could well produce Boolean equations from truth tables that were *entirely noise*: that is, that were generated by entirely random processes.

In response to this, it has been argued (e.g. Ragin and Rihoux, 2004) that such random truth tables would in fact be readily shown up as such, in that they would contain many contradictions – that is, instances of identical sets of conditions being associated with the

discussion of difficulties associated with purely logical analysis. However, as shown by Krogslund, Choi and Poertner (2015), the results of fuzzy-set QCA are highly sensitive to quite small changes in the parameters applied in the 'calibration' of set membership and in then carrying out the Boolean minimisation. And the question does thus arise of what fuzzy-set QCA can do that cannot be done more simply and reliably through existing statistical methods such as, say, loglinear modelling or latent class analysis (on which, see further Chapter 7). As Achen (2005: 29) has commented, Ragin's repeated assertion that quantitative methods of data analysis in the social sciences are restricted to regression analyses estimating context-free net effects is 'a mystifying claim indeed'.

[9] Interestingly, Lucas and Szatrowski (2014) seek to illustrate their case by showing that a QCA analysis of the – immediate – cause of the *Challenger* disaster, previously discussed in this chapter, gives an explanation that is in contradiction with that generally accepted by engineers and also by the Presidential Commission of Inquiry: that is, that the disaster was precipitated simply by the failure at low temperatures of the rubber seals on the rocket booster. The QCA analysis would imply that an interaction with other factors was also necessary to the failure – despite there being no independent evidence of this.

outcome of interest both occurring and not occurring – and that it is a standard task in QCA to resolve such contradictions, usually by introducing new conditions, before proceeding further. But this response turns out to be inadequate. On the basis of simulations, Marx (2010) has shown that contradictions do not necessarily arise with randomly generated truth tables. What is crucial is the number of cases covered and the number of conditions involved. More specifically, Marx's (2010: 155) results lead him to suggest that '[QCA] applications with more than 7 conditions (including the outcome) and applications where the proportion of conditions on cases is higher than .33 are not able to distinguish real from random data'. And his overview of studies using QCA (Marx, 2010: Table 4) then reveals that a substantial proportion were in fact of this very questionable kind.

At the root of the problem that arises here is the fact that, as a logical method assuming a deterministic world, QCA must aim to account *fully* for all cases considered. In statistical language, it must aim to account for 100% of the variance in the outcome of interest. This means, as Seawright (2005: 16–18) has pointed out, that in QCA *all* conditions that are causally relevant to the outcome of interest in the population of cases studied have to be included: not only those that operate with some regularity across cases, but also quite 'idiosyncratic' – and, one might suspect, purely chance – conditions that may be of relevance only in this or that particular instance. Thus, it is always likely that the number of conditions that will need to be considered and their ratio to the number of cases will be pushed up towards the danger levels that Marx identifies, and that would seem in practice to be frequently exceeded.[10]

[10] Even where QCA is applied to data-sets with quite large Ns (see e.g. Cooper, 2005), the problem of over-interpretation may still arise in that Boolean equations can be arrived at that imply complex interaction effects which, if incorporated into, say, a loglinear model, would not prove statistically significant or, in other words, could easily reflect merely chance aspects of the data (see Krogslund, Choi and Poertner, 2015: 50–1).

Marx (2010: 147) himself aptly characterises the problem as one of 'uniqueness'. As the number of conditions distinguished approximates the number of cases analysed, the point is being reached at which each case has to be seen as representing a unique configuration of conditions: that is, there are *no* regularities. At this point, the possibility of contradictions is eliminated but the Boolean equation produced becomes meaningless, in that it could apply as well to random data as to those actually derived from the cases under study. An alternative way of putting the matter, to revert to the earlier discussion of this chapter, would be as follows. Where a situation of the kind in question arises, what is being indicated is that – as in the case of the supposed sociology of revolutions – an attempt is being made to provide a sociological explanation for events that are not in fact characterised by sufficient regularity to allow for this, nor indeed for a theoretical explanation of any kind. Or, in other words, to the extent that the problem of uniqueness threatens in the case of events or complexes of events, it is historical rather than sociological explanation that is called for.

Insofar as proponents of QCA propose a solution to this problem, it would appear to be that of limiting explanatory analyses to populations of cases that are deemed to be 'comparable' in the sense of being 'causally homogeneous' or, in other words, in allowing a contradiction-free truth table to be obtained with a relatively small number of conditions. But to resort thus to 'constructed' populations (Ragin, 2013: 173; and see Goertz and Mahoney, 2009) rather than ones that are defined independently of the explanatory model to be put forward – and, presumably, in the light of the analyst's prior substantive interests – must mean that the scope conditions of the model are set in a quite arbitrary way. The description of population regularities and their explanation are confounded; and only those cases are to be considered where a particular explanatory model can be shown to fit. While all theories in the social sciences are likely to require scope conditions of some kind, the appropriate procedure (see further Chapter 9) must be to develop theories in order to account for quite independently established explananda and then to discover, through further research, how adequate to the task these theories are and what their limitations, including their scope conditions, might be. Otherwise, one has from the start explanatory models that, like an ill-cut suit, just fit where they touch.

Examination of the difficulties faced by logical methods of analysis, based on the assumption of a deterministic social world, does then help to highlight distinctive features of sociology understood as a population science based on the assumption of a probabilistic social world, and thus reliant on statistical methods. In this latter case, two limiting conditions are recognised at the outset. First, it is accepted, on the grounds set out in this chapter, that appropriate explananda for sociology will be only events of a kind that can be shown empirically to be expressed in aggregate-level, probabilistic regularities, emergent from the states and behaviour of individual members of populations. And second, as will be discussed at greater length in the chapters that follow, it is accepted that variance in the outcomes of interest will be accounted for not totally, but only to the extent that this variance results from factors that can be regarded as operating in a systematic rather than an idiosyncratic or quite random manner (see King, Keohane and Verba, 1994: ch. 2 esp.).