SAGE BENCHMARKS IN SOCIAL RESEARCH METHODS

QUANTITATIVE SOCIAL SCIENCE

VOLUME I

Overview and Major Issues

EDITED BY

Jacqueline Scott
and
Yu Xie

2005

SAGE Publications
London • Thousand Oaks • New Delhi

51. "Application of Hierarchical Linear Models to Assessing Change", *Anthony S. Bryk and Stephen W. Raudenbush*
Psychological Bulletin, 101(1) (1987): 147–158. Copyright © [1987] by the American Psychological Association. Reprinted with permission.

52. "Discreet-Time Methods for the Analysis of Event Histories", *Paul D. Allison*
Sociological Methodology, 13 (1982): 61–98. Published by the American Sociological Association. Reprinted with permission.

53. "Dynamic Analysis of Event Histories", *Nancy Brandon Tuma, Michael T. Hannan, and Lyle P. Groeneveld*
American Journal of Sociology, 84(4) (1979): 820–854. © 1979 by The University of Chicago. Reprinted with permission from The University of Chicago Press.

54. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models", *Luc Anselin*
Agricultural Economics, 27(3) (2002): 247–267. Published by Blackwell Publishing. Reprinted with permission.

55. "Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children", *Robert J. Sampson, Jeffrey D. Morenoff, and Felton Earls*
American Sociological Review, 64(5) (1999): 633–660. Published by the American Sociological Association. Reprinted with permission.

56. "An Introduction to Sample Selection Bias in Sociological Data", *Richard A. Berk*
American Sociological Review, 48(3) (1983): 386–398. Published by the American Sociological Association. Reprinted with permission.

57. "Models for Sample Selection Bias", *Christopher Winship and Robert D. Mare*
Annual Review of Sociology, 18 (1992): 327–350. Reprinted, with permission, from the *Annual Review of Sociology*, Vol. 18 © 1992 by Annual Reviews www.annualreviews.org.

58. "Anatomy of the Selection Problem", *Charles F. Manski*
The Journal of Human Resources, 24(3) (1989): 343–360. Copyright © 1989. Reprinted by permission of The University of Wisconsin Press.

59. "Heterogeneity Ruses: Some Surprising Effects of Selection on Population Dynamics", *James W. Vaupel and Anatoli I. Yashin*
The American Statistician, 39(3) (1985): 176–185. © 1985 American Statistical Association. Reprinted with permission from the American Statistical Association.

60. "A Latent Class Analysis of Tolerance for Nonconformity in the American Public", *Allan L. McCutcheon*
Public Opinion Quarterly, 49(4): 474–488. © 1985 by the American Association for Public Opinion Research. Reprinted by permission of Oxford University Press.

Editors' Introduction

Jacqueline Scott and Yu Xie

> Numbers gathered without some knowledge of the regularity to be expected almost never speak for themselves. Almost certainly they remain just numbers.
>
> Thomas Kuhn (quoted in Duncan 1984: 169)

Quantitative research that uses numerical or statistical information is commonplace in virtually all branches of social science. Disciplines such as sociology, economics, psychology, political science, social anthropology, social geography, public health, demography, public policy, and education all make use of quantitative data. In fact, we assert that there is no social science discipline that can afford not to embrace the quantitative approach, and there is no serious social science scholar who can deny the important contribution of quantitative research.

To appreciate the above statement one only needs to consider factual information about contemporary societies. Much of what we know as styled social facts has been provided or studied by quantitative social scientists. Examples include: population sizes and compositions, socioeconomic inequalities, economic development, political mobilization, educational attainment, trends in marriage and cohabitation, gender relations, economic well-being, and physical and mental health. It would be unthinkable to live and function in a modern and democratic society without accurate, and mostly numerical, information about ourselves and others around us.

However, the widespread use of quantitative information itself does not guarantee universal acceptance of the dominant role of quantitative and statistical methods in social science. For some time now there has been an ongoing debate about the respective strengths and weaknesses of quantitative and qualitative approaches in the social sciences. However, the debate is based on an exaggerated notion of the different epistemologies. Michael Mann (1981; in Volume 1) presents this point forcefully in his article on socio-logic. He claims that there are few real adherents to epistemologies of objective knowledge. The researcher has no choice but to act with a kind of 'as if positivism'. According to Mann, 'It matters not the slightest whether the researcher also advocates positivism or empiricism or neo-Kantian idealism or dialectical materialism or realism or pragmatism or whatever; if s/he undertakes research and seeks to have it evaluated by others, the methodology will be the same (p. 548).'

Unfortunately, there has been a backlash against 'positivism' in sociology, anthropology, and other social science disciplines. But if one looks at how 'positivism' is treated in those criticisms, a precise meaning of the term is completely missing. The anti-positivist sentiment can range all the way from distrust of numerical information and statistical methods to plain ignorance about contemporary quantitative research. John Goldthorpe (2003: 23), a British sociologist, puts it this way:

> Attacks on [sociology as social science] by proponents of expressive and critical sociology have of course been alike focused on 'positivism'. However, significant differences show up in what 'positivism' is taken to mean and why it is found objectionable, and one of the few common elements in such attacks is a rejection of quantitative methods in sociology, and, it seems, of any kind [of] systematic, reasoned and transparent procedures for data collection and analysis.

In our view, despite its serious limitations, the quantitative approach constitutes the core of contemporary social science and the foundation for generating and accumulating knowledge about contemporary societies. However, the methods that enable quantitative research in social science are problematic in one way or another. Given the subject matters that social scientists typically study, perfect methods simply do not exist. Indeed, all methodologies ever devised for the study of society and social relationships have been found to have limitations. Here arises the irony: how can we ever put so much stock into knowledge gathered through possibly faulty methods?

A short answer is that, despite its severe limitations, quantitative methodology remains far more attractive than the alternatives. To see why we make this potentially controversial statement, we need to examine the various reasons that make quantitative methodology

problematic. These reasons are usually labelled under certain types of 'biases', such as, sampling bias, missing data bias, interviewer bias, measurement bias, confounding bias, sample selection bias, truncation bias, omitted-variable bias, serial correlation bias, ecological bias, reactivity bias, unobserved heterogeneity bias, and endogeneity bias. While we can extend the list further, we have illustrated our point. What unites all these types of biases is a basic problem of variability: seemingly similar elements in a given population or sub-population can vary in important ways unknown or even unknowable to the researcher.

Social science resembles life science more than physical science in the role that variability plays. As Mayr (2001; in Volume 1) forcefully argued, variability as a scientific conceptualization did not receive much attention until Charles Darwin. Indeed, much of the success in physical science prior to Darwin can be attributed to the successful implementation of Plato's proclamation that (scientific) truths should be universal, unchanging, and thus abstract. According to Plato's definition, the scientist should not be concerned with observed variations, which are imperfect replicas of perfect ideas that exist independently and perpetually in the 'world of being'. It was Darwin who challenged this long-standing view and brought variability back to scientific inquiry. We agree with Lieberson and Lynn (2002; in Volume 1) that social science research should be primarily concerned with variability. Social phenomena and human behaviours are so diverse that no perfect methods can be devised to provide simple and unambiguous answers. Results from almost any study (even one based on an experiment) are subject to different interpretations.

No study in social science can be truly definitive because population variability can pose all kinds of threats under the various 'biases' discussed earlier. For example, subjects who are studied could be different from those who are not studied (that is, sampling bias). Respondents to a survey could vary, in different ways, their answers according to the characteristics of the interviewer (interviewer bias). Individuals who attend college could be systematically different from those who do not, in attributes such as aptitudes, work ethics, and ambition (unobserved heterogeneity bias). Given an option, workers may enrol in a training programme only if they know that such a programme will enhance their earnings capacity (endogeneity bias). All these and other 'problems' in social science research can be traced to the root of variability: we are not sure if the units of analysis under a study differ from each other and from those outside the study in observed and unobserved characteristics that may affect the validity of conclusions that can be drawn from the study.

In a Hegelian sense, what makes quantitative methodology unreliable and problematic is precisely what makes it imperative: the variability principle that Mayr rightfully extracted from Darwin. Variability is the essence of human society. Without a quantitative approach, it is simply not possible to characterize such variability. Other alternatives, such as, speculation, introspection, personal experience, observation, and intuition can and do advance our understanding. However, these other methods do not protect a social scientist from the threats posed by variability. Properly construed, quantitative methodology deals with population characteristics and their characterizations, yielding useful (albeit always tentative) empirical results.

Within the quantitative tradition, it has long been acknowledged that adequate information should be provided to allow the reader to assess the quality of the work and, in principle, to replicate the analysis. While this ideal is not always met, quantitative research, in using statistics, does subscribe to agreed standards of analytical precision. As Gary King (1986) notes, despite important cautionary tales of statistical lies and misuse (for example, Huff 1954), it is a lot harder (knowingly or not) to lie (and get away with it) with statistics than without them. Obvious misuses of quantitative methodology are not uncommon, but also easy to spot. In the four volumes that follow in the series, we propose something different to the reader. We present many excellent examples of research articles that should serve as a model of good statistical practice across a wide range of social sciences. While they do not (as they should not) solve all problems that a researcher may encounter in a research setting, the articles should provide enough ideas from which to learn and advance.

History of Quantitative Social Science

Auguste Comte (1853: 24) once said, 'Science cannot be completely understood without a knowledge of how it arose.' We hope that a historical overview in this section will help us understand why quantification became an important feature of the scientific method, and what consequences this had as social science emerged and became established. Marquis de Condorcet (1743–94), the French mathematician and philosopher, has been credited with coining the phrase 'social science'. More substantively, de Condorcet combined scientific methodology and commitment to evidence-based social reform and argued for social mathematics as an appropriate tool for studying society (Oakley 2000). By 'social mathematics' he meant 'political arithmetic' or statistics as social facts. In Britain, John Graunt (1620–74) and William Petty (1623–88) had pioneered political arithmetic, which harnessed the state's needs for information (hence the term political) for statistical analysis of numerical data. The term 'political arithmetic' has been retained in the UK, and today refers to a particular style of quantitative research that has focused on class inequalities in education (Heath 2000).

Adolphe Quételet (1796–1874), a Belgium mathematician, was instrumental in formalizing mathematically-based social science. His particular contribution was the application of probability theories for measurement, those associated with the normal distribution, to social phenomena. His application of the normal distribution to social phenomena led him to invent the concept of the 'average man' as the appropriate focus. To Quételet, averages in a population or subpopulation are attractive in that they seem to be stable and predictable, despite the apparent large variations and uncertainties in individuals' behaviours. That is, averages seem to satisfy Plato's high standards of invariance and absoluteness for universal truths. A contemporary of Quételet, Comte, who is credited with coining the term 'sociology', originally made his living teaching mathematics. Sociology was, according to Comte (1853), the last in line of the sciences to go through the three successive stages of knowledge: as theological or fictitious; the metaphysical or abstract; and the scientific or positivist.

The notion of the 'average man' in statistics could be said to be an instance of the typological or essentialist thinking that dominated the world-view prior to Charles Darwin's seminal work on the origin of species (1964). As Mayr (2001; in Volume 1) indicated, one of the most important and long-lasting influences of Darwin's evolution theory is 'population thinking' as distinguished from the 'typological thinking' prevalent in physical sciences. It is the variation among uniquely different individuals in which reality lies, while the statistical mean value of variation is an abstraction. This philosophical turn in the mid-19th century was crucial for Darwin's theory of natural selection and for today's social science generally.

Another strand of Darwinian thought that plays a crucial role in the development of modern social science is the acceptance of the role of chance in understanding social development, together with the importance of the time factor (historical contingencies). As Mayr rightly notes, we still lack a thorough analysis of the history of the gradual acceptance of chance in scientific explanation. In modern social science the role of chance remains somewhat under-theorized and contentious. John Maynard Keynes' *Treatise on Probability* (1921) took a fairly strong position about the theoretical, non-quantifiable, subjective elements involved in any judgements of probability, although for Keynes probability judgements are firmly rooted in careful analysis of empirical data.

Darwin's evolution theory also inspired social scientists to engage with biological science in general, and investigate the interplay of genetics and the environment in particular. Regrettably, from the modern viewpoint, the earlier strands of social Darwinism and eugenics that exercised such influence in the late 19th and the early 20th century is unpalatable, with deplorable social consequences. However, a beneficial by-product of the eugenics movement was the development of statistical tools that laid many of the foundations for modern statistical analysis. The founding father of modern statistics Francis Galton (1822–1911), a cousin of Charles Darwin, was the person who coined the term 'eugenics' from the Greek root meaning 'noble in heredity'. An extraordinary man of multiple talents, Galton was prompted by his own personal experience to realize the importance of intellectual dynasties–talent running in families. Galton built on Quětelet's method of the normal distribution. Departing from Quětelet, however, Galton focused on variations and co-variations at the individual level. This led Galton to the idea of statistical regression for demonstrating the importance of genetic inheritance. Galton's idea was formalized mathematically by Karl Pearson (1857–1936), who founded biometrics as a discipline. Pearson's contributions to statistics include a correlation coefficient and a chi-squared test named after him. By the late 1920s Pearson's influence had waned, and R.A. Fisher (1890–1962), another committed eugenicist, began to dominate the field of statistics.

R.A. Fisher was a statistical genius who fully developed a unified system of statistical inference that came to be called the classical or frequentist approach, as opposed to the Bayesian approach (discussed later). Most undergraduate students today learn this system in their statistics classes, which was explicated in Fisher's seminar book *The Statistical Methods for Research Workers* (1925). (More details of Karl Pearson and R.A. Fisher's contributions can be found in the excellent source of statistical biographies accessible through the web site of the American Statistical Association.)

Fisher pioneered the random control trial (RCT) as a scientific method, based on his work on agricultural productivity and the importance of fertilizer. The RCT method has become the 'gold standard' for much evaluation work throughout science, especially in medicine.

The method has a great deal of appeal because randomization should, statistically, make comparable a group that receives a treatment and a group that does not receive treatment (also called control). Furthermore, Fisher's 1935 book *Design of Experiments* shows how matching techniques can be used as a means of control, thus making the experimental method more relevant to the social sciences, where random assignment is usually infeasible. The mathematician W.G. Cochran (1909–80), who was a young colleague of Fisher in agricultural research at the Rothemsted Experimentation Centre in England, authored of one of the earliest papers in this collection (Cochran 1957; in Volume 3). The relative merits of experimentation, quasi-experimentation, and observation continue to be an important design topic in quantitative social science research (see Meyer 1995; Rosenbaum 1999; Sobel 2000; all in Volume 1).

While statistical methods in today's social science were largely based on Fisherian statistics that developed primarily from agricultural science, methodological research has also been originated in social science to address real (and difficult) problems unique to social research (Clogg 1992). It is these substantive problems facing social scientists that have inspired (and will hopefully continue to inspire) many of the most useful developments in statistical methods, such as, log-linear models, latent variable models, discrete choice models, sample selection models, and multi-level models. In particular, rapid advances in econometrics in the past five decades have been driven by an earnest quest to quantify parameters that are structural, causal, and theoretically interpretable, rather than merely descriptive (Heckman 2000, in Volume 3).

The Rise of Quantitative Sociology in the US

In this section we examine the early history of sociology in the United States as an example of the rise of quantitative research in a social science discipline. Franklin Giddings (1855–1931), one of the founding fathers of American sociology, argued passionately in favour of quantitative work, although he did not produce any methods of lasting value. Giddings defined sociology as the study of social phenomena at the aggregate level and wrote, 'We need men not afraid to work, who will get busy with the adding machine and logarithms, and give us exact studies, such as we get in the ... laboratories' (Giddings 1901, as quoted in Oakley 2000: 163). He redefined the law of the three stages that were put forward by Auguste Comte (theological, metaphysical, and positivist) as 'speculative', 'observational', and 'metrical'. The pecking order was clear as far as Giddings was concerned and 'metrical' or quantitative social science stood at the apex of social research. Not all agreed, and in the 1930s there was a power struggle between rival schools of sociology in the United States, typified by Columbia (where Giddings steered social research as a quantitative science) and Chicago, which became associated with fine-tuned ethnographic studies and the theoretical reaction against 'scientism' that symbolic interactionism posed.

If Giddings' way of expressing himself seems rather too naïvely positivist for comfort, a far more nuanced and broad-minded understanding of the potential of quantitative analysis is represented by Paul Lazarsfeld (1901–76). Lazarsfeld was a German-trained psychologist who became the Quételet Professor of Social Sciences at Columbia. As Robert Merton (1979) notes in his paper 'Remembering Paul Lazarsfeld', students often came to work with Lazarsfeld because he was regarded as a shining light in the methodology of social science. Lazarsfeld's mathematical bent led some to conclude that he was narrowly quantitative. Yet nothing could be further from the mark. Throughout his career, Lazarsfeld took it as his moral duty to demonstrate the value and frequent necessity of combining quantitative inquiry with qualitative insights. Merton recalls how Lazarsfeld, faced with over-zealous socio-metric types who were overly awed by numbers, intoned the monitory words of St Augustine: 'So it is O lord my God, I measure it and know not what it is I measure.'

Lazarsfeld's contributions lay primarily in survey research and analysis of survey data in the form of partialing out spurious correlations between two variables with additional control variables. Lazarsfeld influenced the work of economist Herbert Simon, whose 1954 article on spurious correlation in turn helped Lazarsfeld refine his ideas on causal analysis. Most of Lazarsfeld's works are books and thus cannot be included in this series.

At Princeton, John Tukey (1915–2000) pioneered approaches to exploratory data analysis (EDA), graph plotting methods that are today fixtures of many introductory statistical texts. Tukey's EDA methods have proven enormously useful for quantitative social science. Tukey, like Lazarsfeld, cautions against the too-ready assumption that data necessarily contain the answer, and he projected a healthy scepticism about what data does or does not represent. Yet, notwithstanding Tukey's scepticism, there is enough in his admonitions on letting data speak for itself for most social scientists who prioritize theory to feel concern. Data exploring is seen as uncomfortably close to data

ransacking, and both are regarded as worryingly atheoretical. Political scientist Nathaniel Beck (2000) tells an apposite story about the difference between statisticians and social scientists. He writes: 'To over-simplify, statisticians work hard to get data to speak; whereas [social] scientists are more interested in testing theory (p. 11–12).' It is a difference that matters. If push comes to shove, social scientists are likely to cling to their theory and search for flaws in their data or inadequacies in their analysis; whereas statisticians will be more inclined to throw out the model and believe the data.

Fortunately, some social scientists seem to successfully walk the tightrope of balancing theory and data, and do inspiring theoretically-led empirical work. One such person whose work has exerted an enormous influence on quantitative sociology and social demography was Otis Dudley Duncan (1921–2004). Duncan's best known work is a 1967 book that he co-authored with the late Peter Blau, the *American Occupational Structure.* Based on quantitative analysis of the first large national survey of social mobility in the United States, the book elegantly depicts the process of how parents transmit their social standing to their children, particularly through affecting the children's education. This work was subsequently elaborated by Duncan and other scholars to include the role of cognitive ability, race, and other factors in the transmission of social standing from one generation to the next.

Duncan also introduced 'path diagrams', 'path models', and 'path analysis' to the discipline of sociology, and he used these statistical tools in the Blau–Duncan book and his other studies of social stratification. A path diagram and a corresponding path model describe a set of equations summarizing complex scientific ideas in terms of statistical relationships (Duncan 1966; in Volume 3). Path analysis was first invented by Sewell Wright, a renowned biologist and evolutionary theorist. Duncan worked jointly with Arthur Goldberger, an econometrician, on the relationship between path analysis and other statistical methods in the social sciences. They showed that path analysis models were closely related to the simultaneous equation models of economics and the confirmatory factor analysis of psychology. These three different ways of analysing certain kinds of data can be viewed within a single general framework, called 'structural equation models'. Today, structural equation models are very widely used.

More than any specific statistical method such as path analysis, Duncan exerted enormous influence on quantitative sociology by establishing a new intellectual tradition in sociology that followed the model of his research. While some earlier sociologists tried to model sociology after physical science, Duncan was openly disdainful of the search for supposedly universal laws of society that would mimic those of physical science. The central tenet in Duncan's new paradigm for quantitative sociology is the primacy of empirical reality. Quantitative tools would not be used to discover universal laws that would describe or explain the behaviour of all individuals. Rather, quantitative analysis summarizes empirical patterns of between-group differences, while temporarily ignoring within-group individual differences. Over time, social scientists can improve their understanding of the world by incrementally adding greater complexities to their analyses.

This new approach was in large part built on a long-standing tradition in demography: it is of foremost importance to document and understand empirical patterns in real populations. This 'demographic turn' in quantitative sociology that was spearheaded by Duncan was highly successful. In a sense, Duncan simply borrowed Darwin's population thinking and emphasized variation, while discarding typological thinking, which is typical in theoretical work antithetical to empirical evidence. Yet, as Lieberson (1992) stresses, theory includes empirical research or knowledge. Theory is the whole body of knowledge. It can and should go *beyond* existing evidence, but it also must include existing information, if only to evaluate it. A theory that ignores existing evidence is, to Lieberson, an oxymoron. Lazarsfeld makes a quite similar point, stressing that theory needs to be rooted in empirical knowledge. He is reputed to have taken C. Wright Mills to task quoting the opening sentence of Mill's (1959) influential book, *The Sociological Imagination*: 'Nowadays men often feel that their private lives are a series of traps' (p. 3). Lazarsfeld's response was: 'How many men, which men, how long have they felt this way, which aspects of their private lives bother them, when do they feel free rather than trapped, and what kinds of traps do they experience?' (cited in Elcock 1976: 13). Here, again, the emphasis is on variation.

Sceptical of grandiose claims in social science, Duncan was the fiercest critic of quantitative sociology. To him, quantification alone is not equivalent to scientific reasoning and in fact can be misleading. In his own words (Duncan 1984: 226):

> We often find the syndrome that I have come to call *statisticism*: the notion that computing is synonymous with doing research, the naïve faith that statistics is a complete or sufficient basis for scientific methodology, the superstition that statistical formulas exist for evaluating such things as the relative merits of different substantive theories or the 'importance' of the causes of a 'dependent variable'; and the delusion that decomposing the covariations of some arbitrary and haphazardly assembled collection of variables can somehow justify not only a 'causal model' but also, praise a mark, a 'measurement model'.

The problem Duncan identifies remains a serious problem today: in sociology, as in other numerical social sciences, individual articles of exemplary quality are published side by side with transparent exercises in statistical numerology. What can be done to escape the trap of statisticism? Duncan suggested two possible paths: improvement of social measurement and an emphasis on the conceptualization of social processes and research designs that reveal such processes. Much of the methodological advances since the 1980s can be viewed as responding to Duncan's critiques, and some of the most important works are represented in this series.

Challenges to the dominance of quantitative research in sociology have come not just from inside–from the methodologists themselves–but also from outside. One strand of criticism has come from feminist scholars (for example, Oakley 1998, in Volume 1). The notion of gendered ways of knowing has been particularly influential in the latter part of the 20th century. Feminists have offered a powerful critique of the traditional canon of social science, as well as attacking what they saw as the misguided tenets of positivism. Oakley herself, in the 1970s, championed the view that quantitative methods were inappropriate for those interested in feminist research. Her early position is firmly renounced in subsequent works, where she claims that divisions between quantitative and qualitative research are unhelpful in the pursuit of useful knowledge. Indeed, she depicts the strengths of quantitative and experimental research for furthering the development of a critical and emancipatory social science. While there are still some who would champion a particular feminist methodology, most feminists, across a wide range of disciplines, acknowledge that both quantitative and qualitative research are needed to understand our gendered world.

Some have claimed the 21st will be the 'social science century'. According to this view, the big issues of our century are about how human beings and societies interact, how they conduct their affairs, and how they capitalize on diversity in society (Rhind 2003). A related point is that social scientists are living in revolutionary times. The future holds out a promise of technological advance in both data access and analytical power that has the potential for allowing social scientists to build up large infrastructures for tackling 'big issues', in a way that parallels the large-scale infrastructures associated with scientific and medical advances.

While it is not obvious that the future of social science lies in large-scale enterprises, it is clear that the future will hold further rapid developments in quantitative methodologies. Indeed, most of the modern quantitative techniques have developed in their current form in the last three decades or so. The advent of the powerful desktop computer has changed not just how a quantitative social scientist obtains and analyses data, but also the nature of research questions that can be asked.

Major Themes

Given the space limitation, it is not possible to include all relevant works in this series. While we do not attempt to be comprehensive in including all classic works that should be read by any aspiring student who wishes to understand quantitative social research, we hope to cover, with representative works, all important themes in quantitative social research. They can be broken down as follows.

Theories, Debates, and Research Designs

We think that research design is the most important, fundamental part of quantitative research. As a first step, the researcher needs to know what the research question is and how to answer the question empirically. Before conducting any analysis of data, a researcher should first conceptualize the substantive problem. Research design to a researcher is like a roadmap and a compass to an explorer in the wild. With poor research design, no fancy statistical method can come to the rescue. There is a limit to what statistical methods can deliver, and the limit is set by the research design.

Time and Space

Humans are social animals that are located in time and space. They are not isolated individuals. If we wish to study variations in human populations, we first need to understand (or decompose) systematic variations by time and space. There are three primary reasons why variations in human populations vary systematically by time and space. The first is that, for a variety of complicated (including genetic as well as environmental) reasons, individuals who are closer in time and space tend to be more similar to each other than individuals located further apart in time and space. Second, individuals at the same time and location have opportunities to interact with each other and thus

affect each other's behaviours and outcomes. Third, individuals in a given time and space may choose to organize themselves in a different matter so that the social setting (such as the structures of governance and public financing) may change and as a result affect the individuals' lives within certain confines of time and space. Thus, it is extremely important that quantitative social science both accounts for, as well as takes advantage of, variations in social phenomena by time and space.

Levels of Analysis

Given variations that are present both across social contexts (as defined by time and space) as well as across individuals within a social context, should the context level, or the individual level, be the appropriate unit of analysis? The answer is both or either, depending on the research question asked. Both the individual level and the context level (or group level) can be units of analysis. When a social theory is purported to explain the variation at the individual level, individuals should be units of analysis. When a social theory is purported to explain the variation at the context level, contexts should be units of analysis. When individual level data are desired but lacking, aggregated group data may be used to address a research question under some strong and testable assumptions. Otherwise we run into potential problems (Robinson 1950; in Volume 1). Because we know that social processes for the variation at the context level may be different from those at the individual level, we sometimes saturate the former and focus on the latter in the form of the fixed effects model.

Surveys and Archives

Quantitative analysis cannot be carried out without raw material–numerical or quantitative data. No grand research design can yield results without data. There are two primary modes of gathering quantitative data: conducting a social survey (or census enumeration) or retrieving information from historical archives. For survey data, the researcher has a better control of sampling so that the data can be assembled to represent a target population. Spatial and longitudinal components can also be built into a survey. For archive data, representation is harder to maintain because there are likely selection biases as to what was initially recorded in the past and then preserved over time. Both surveys and archives are valuable sources of numerical information.

Measures, Scaling, and Missing Data and Context

Quantitative social scientists often discuss their research strategies and results in terms of 'variables'. How are the variables constructed? What should be done if there are missing cases in a variable? How do we measure and present contextual and network data? To answer these questions, we put together a list of articles under the heading of 'Measures, Scaling, and Missing Data and Context'. They only constitute a small number of possible topics that could be included under the title, but they represent these topics. For example, the Duncan and Duncan (1955) article is a classic statement on an aggregate measure of residential segregation by race. Bollen and Lennox's (1991) article provides an overview of the issues involved in constructing a composite measure from multiple items.

Interpretation, Covariates, and Controls

Interpretation of statistical results in social science, especially those based on observational data, always requires care. More often than not, results based on observational data change when we change model specifications (LaLonde 1986). One major factor is what other observed confounding factors should be controlled. For simplicity, let us say that the research objective is to study the causal impact of one binary factor, that is, to borrow the language of experimental designs, the comparison of 'treatment' versus 'control'. One possibility to control for other observed confounders is to encompass them into a catching-all propensity score of treatment (Rosenbaum and Rubin 1984; in Volume 3). The propensity score approach is attractive because it affords the researcher an easy way to examine differences in observed covariates and compare the treatment and control groups flexibly and non-parametrically.

The research potential of longitudinal data becomes increasingly rich over time. In the UK, for example, there is an extraordinary sequence of birth cohort studies that began in 1946 and have continued in 1958 (the National Child Development Study), 1970, and most recently the cohort born in the new millennium. Such a sequence provides a unique opportunity for studying social change. One such change concerns the far-reaching ones in family life, including the way family instability and parental divorce may adversely affect children's subsequent life course. The study of Bhrolcháin et al. (2000; in Volume 3) contrasts the changing impact of divorce between two birth cohorts. Yet, as they demonstrate, causal interpretation remains problematic even when the researchers have good contextualized data from multiple sources across time. Another potential benefit of the UK cohort studies is that they contain information that allows researchers to begin to explore the interplay of genetics and environment on human behaviour (see also Volume 1).

Bayesian Approach

The Bayesian approach refers back to the probabilistic theorem of Thomas Bayes the English Theologian (1702–61), whose theorem was published two years after his death. Bayesian statistics, as they developed over the past half century with the advent of significant computing power, does not bear much resemblance to its namesake. Indeed, it is not clear whether Bayes would have recognized or approved of the modern quantitative approaches that bear his name. At core of Bayesian statistics is the philosophical notion that probabilities are ultimately subjective. However, like Lieberson's (1992) social theories, current subjective probabilities are (or should be) based on accumulated evidence in the past. Thus, what empirical data from a particular study contribute to subjective probabilities is to update and refine prior subjective probabilities, converting them to a posterior probability distributions. The Bayesian approach is a complete statistical system that covers a large number of topics, such as statistical inference, imputation of missing data, and assessment of a model's goodness-of-fit. In particular, the Bayesian approach has led to many powerful computational methods for models that can be interpreted from the classic, frequentist perspective. The whole area of Bayesian statistics is exciting and fast-moving. We regret that we can only provide the reader with a taste for the approach, beginning with a review by Berger (2000; in Volume 3).

Structural Equation, Contingency Table, and Discrete Choice Models

As discussed earlier, structural equation models are systems of regression equations involving multiple dependent variables and multiple independent variables. A system of equations would allow a researcher to decompose a 'total effect' into a 'direct effect' component and an 'indirect effect' component (Alwin and Hauser 1975; in Volume 3). Structural equation models are used widely in sociology, economics, psychology, and other related fields. Advanced versions can also deal with categorical or truncated dependent variables (Muthén and Muthén

2000; in Volume 4). Contingency table models for multivariate discrete outcome variables were developed by biostatisticians and sociologists. An early important work was Bishop, Fienberg, and Holland's (1975) landmark book *Discrete Multivariate Analysis.* Goodman (1979; in Volume 3) later provided an important innovation of such models for ordinal variables.

While the contingency table model approach takes discrete data as if they are inherently discrete, an alterative approach views discrete data as discrete realizations, through a threshold model, of latent quantities that are theoretically continuous. First introduced by the psychologist Thurstone (1927), this approach has taken dominance in economics today, where the latent tendency means a utility function. This approach has given rise to the discrete choice model (also called 'conditional logit model') for studying choice behaviours (McFadden 1974). Taken together, the three classes of statistical models–structural equation, contingency table, and discrete choice–constitute the workhorses of statistical analyses in quantitative social science publications and should be understood by all sophisticated quantitative researchers today. They do not require special features of data and can be used widely in applied settings.

Multi-Level, Longitudinal, and Spatial Analyses

The three classes of statistical models discussed under the preceding heading all assume that the units of analyses are independent of each other. Multi-level models, longitudinal models, and spatial models all relax this independence assumption, and indeed make use of the systematic dependencies across units of analysis. In a way, this second set of statistical models extends the first discussed earlier, utilizing extra, known information about relationships among units of analysis from data collection. For example, multi-level models allow a shared variance across units of analysis at a lower level (say, students) within a unit of analysis at a higher level (say, a school). Thus, multi-level models allow the researcher to study the determinants of variation at multiple levels (say, variation in academic performance at the individual level and the school level). Similarly, spatial models allow units of analysis that are closer in distance to be more correlated than units of analysis that are further apart.

Longitudinal models are facilitated by the availability of longitudinal data, repeated observations on the same units of analysis over time. The challenges of appropriate dynamic analysis go hand in hand with the challenges of research design and data collection. The dynamics

of poverty, labour market behaviour, or family composition were not well understood until adequate longitudinal data became available. Analysis of longitudinal data, such as, the US Panel Study of Income Dynamics (1968 onwards), has produced results that indicate these processes were far more fluid than had been previously expected. It was clear that inferences from cross-sectional data had erroneously and systematically underestimated change. In the early 1960s poverty was conceptualized as a static state, with an immutable distinction between the poor and not poor. It is now recognized to be a dynamic process, and people move between finite spells of poverty that have distinct beginnings and ends. The focus on 'spells' has led to an important conceptual distinction between transitory and long-term poor. Dynamic modelling of longitudinal data makes it possible to study transitions into and out of spells, and the relevance of time-varying predictors.

Selectivity and Latent Class

Let us now return to the earlier discussion of estimating the causal effect associated with a treatment. Average treatment effect can be reliably estimated from an experiment–random control trial. However, social scientists usually have access to observational data only. With observational data the causal inference of a treatment may be incorrect due to two types of selectivity: selection on observable factors and selection on unobservable factors. The problem of selection on observables can be solved with the help of covariate controls or propensity score. However, the problem of selection on unobservables is the most difficult to handle (Manski 1989; in Volume 4). Nonetheless, we can conceptualize them and logically infer their consequences if they were present.

Modelling selection on unobservables is possible with a class of models closely associated with the economist James Heckman (for example, 1978, 1979). These models were introduced to the mainstream social science audience in the mid-1980s by Maddala (1983). However, researchers later found results from such models to be highly sensitive to an untestable, parametric assumption about multivariate normal distribution for errors across multiple equations. A different approach, the latent class approach, drawing on Paul Lazarsfeld's insight of conditional independence within a relatively homogeneous group, is to partition a sample of units of analysis into subgroups, each with its own model parameters. The latent class approach has proved to be a useful descriptive device that attempts to classify data into different groups. However, it is always subjective as to how many classes should be allowed to capture group-level variation while tolerating a certain level of residual within group variation.

Comparative Research

The preceding themes correspond to the 10 groups of articles (called 'parts') in the series. One important theme we do not give a separate heading to is comparative research. The usefulness of quantitative cross-national comparisons is indisputable (although the virtues of comparative research are certainly not confined to nations). Comparative research can serve as a powerful check against parochialism and unwarranted generalizations, as well as serving to identify rare and unique socio-economic processes. Our reason for not giving a separate heading to comparative research is because methodological issues for comparative research permeate all of the parts in the series, plus additional complications involving comparability across societal contexts. For example, non-response, measurement, model specification, and observed heterogeneity all raise questions at two levels. First, to what extent are the methods applicable to a particular setting? Since no absolute, perfect methods exist, some sorts of 'errors' (such as non-response, measurement, model specification, and observed heterogeneity errors) are inevitable. However, for comparative research it is important to consider whether the errors are similar across different settings. Otherwise, the cross-national differences found in the results may reflect the differences in errors rather than real differences for a substantive question.

One particular difficulty in cross-national research is the problem of posing counter-factual questions. These 'what would have happened if' questions often face extreme issues of selection bias and endogeneity, which can only be resolved (if they can be resolved at all) by strong theoretically derived understandings of the mechanisms of causal links (Esping-Andersen and Przeworski 2001). Thinking through counter-factuals raises issues about what constitutes appropriate evidence. This is a challenge for all social research.

In a thought-provoking review of quantitative cross-national comparisons, Esping-Andersen and Przeworski (ibid.) suggest that comparative research has been spurred on, in part, by the increasing availability of good data for a growing number of countries and, in part, by the diffusion of statistical skills among researchers. Both observations require qualification. There is still a surprising shortage of high-quality cross-national survey data that contain carefully crafted functionally equivalent socio-economic indicators, together with random samples

and high-quality field procedures. This is to be rectified as newer studies come on line. These include the European Social Survey, the International Social Survey Programme, and the European Household Panel Study, which are all serving to heighten methodological standards and enhance research potential. Methodological tools are also needed to specifically facilitate comparative research. Comparative social mobility research, for example, has been facilitated by a class of models that focus on cross-context comparisons (Xie 1992; in Volume 3).

Selection of Articles

We believe that decisions about appropriate evidence and analysis are rooted in theory and research goals. Our conviction that methodology should serve a researcher's need to address substantive problems is the first of several principles that guided the selection of articles for this series.

Altogether, there are five main principles that underpin the selection of quantitative approaches that are contained in this collection. First, we take it that methodology for social science is a means to an end. The appropriate use of methodology is bound with substantive objectives. In these volumes we do not attempt to cover the vast history and range of topics relevant to social statistics. Rather, we focus on the *advances* and *uses* of quantitative research to address substantive issues in the social science. Exemplars of research using quantitative methods across a wide range of topics in various disciplinary and interdisciplinary studies are included in each section.

Second, we wish to illustrate how long-established and newly developed quantitative techniques can be used to support innovative forms of enquiry. We do not aim to provide a recipe book for the use of quantitative methods. There are many excellent statistical guides and quantitative methodological textbooks that will guide interested readers through the decision processes about what technique to choose given a particular problem, and what statistical approach to employ given particular analytical goals.

Third, we illustrate more recent developments for quantitative research design, data collection, and analysis. In particular, we examine how quantitative techniques have developed to meet the challenges posed by the demands for the analysis of complex data, such as, longitudinal data, multi-level data, spatial data, and network data.

High-quality social science is dependent on rigorous methodology to underpin the usefulness, validity, and reliability of its results. A fourth principle, on which this collection is based, is that social science is best served by appropriate methodology. As we indicated at the outset, we do not subscribe to the unhelpful quantitative and qualitative divide that pits one approach against the other. Rather, we view each approach as appropriate for different purposes. In addition, we note that within quantitative methods there is a range of approaches that makes different assumptions and offers different analytical approaches. We present exemplars that illustrate 'best fit'–the use of appropriate methods, given specific research objectives and data constraints.

We have been at pains to point out that quantitative methods, like any methodological approach, can be misused. Used properly, however, some quantitative approaches support many different applications, while others are relatively specific in their use. The articles in this collection range from simple to highly sophisticated techniques, but the common focus is on the appropriateness of application for a given research question. Much of quantitative social research rests on a sophisticated mathematical base. However, the fifth principle that guides our selection of articles is that, in general, they should be accessible to the numerate and interested reader. Some articles do include the mathematical formulas that underpin the techniques, but many do not. In either case, we emphasize logical reasoning rather than statistical techniques. Thus, we do not subscribe to the view that it is necessary to be a social statistician in order to produce high-quality quantitative research. The advent of the microcomputer brought with it a whole range of statistical software that renders quantitative techniques that rest on complex algorithms within the reach of the moderately numerate.

Our selection, based on these five criteria, remains subjective. We include articles that we regard as well written, methodologically important, and substantively illuminating. Most are by leading authorities in the field. Many of the articles have helped shape our own understanding of quantitative research and have proved useful for teaching purposes. We also consulted colleagues and trawled the many course outlines from well-established centres of quantitative excellence. We consulted the Social Science Citation Index (SSCI), but, if anything, this increased our scepticism of citation metrics as a useful indicator for assessing the quality or value of articles. Nonetheless, citations are a useful indicator of how wide-ranging an article's influence has been. Some of our articles are 'citation classics' and are flagged as such in the SSCI. Others are of similar quality, but have not yet received the attention they deserve. (This, however, assumes a relationship between citation and influence that may or may not be warranted.) Our hope is that the articles complement each other well, and together serve to

illuminate the wide range of themes and topics that inform the fast-evolving knowledge base of quantitative social science.

The focus of this collection inevitably reflects, in part, the editors' disciplinary backgrounds and research expertise. Both are sociologists: Scott is particularly interested in family sociology, and the design and analysis of longitudinal data; and Xie has expertise in demography and social mobility research. Our collection does refer to work from disciplines other than sociology, including psychology, demography, history, education, geography, economics, and social policy.

We are aware that the collection omits some important topic areas. Some topics are omitted because they are not sufficiently central to main-stream social science concerns. For example, we make only passing reference to data-mining, which, though important in data-driven commercial and marketing activities, has had less impact on mainstream social sciences to date. Other topics are excluded because they are still in early stages of development. This group includes micro-simulation techniques, which are already used in specialized areas, but are now undergoing rapid development for mainstream social science applications.

Our selection is also limited to works published as journal articles. There are many social scientists who have made immensely important contributions to quantitative social science, but are omitted from this collection because their key works have been in monographs or edited books. There is also the fact that this collection only contains articles written in English, and unquestionably this language restriction has led to some influential non-English articles being excluded.

Despite these limitations, the range of topics that we are able to include is immense. We range from the theoretical to the practical, from the simple to the complex, and from feminist critiques to debates about nature versus nurture. For readers who should find the coverage limited, we hope that this collection of articles stimulates their interests and thus serves as a springboard for their plunge into the vast sea of quantitative social science methodology.

References

Alwin, D.F. and R.M. Hauser. 1975. The decomposition in effects of path analysis. *American Sociological Review, 40*(1): 37–47.

Beck, N. 2000. Political methodology: A welcoming discipline. *Journal of the American Statistical Association*, 95: 651–54.

Berger, J.O. 2000. Bayesian analysis: A look at today and thoughts of tomorrow. *Journal of the American Statistical Association, 95*(452): 1269–76.

Bhrolcháin, M.N., R. Chappell, I. Diamond, and C. Jameson. 2000. Parental divorce and outcomes for children: Evidence and interpretation. *European Sociological Review, 16*(1): 67–91.

Bishop, Y.M., S.E. Fienberg, and P.W. Holland. 1975. *Discrete multivariate analysis: Theory and practice.* Cambridge, MA: MIT Press.

Blau, P. and O. Duncan. 1967. *The American Occupational Structure.* New York: Wiley.

Bollen, K. and R. Lennox. 1991. Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*(2): 305–14.

Clogg, C.C. 1992. The impact of sociological methodology on statistical methodology. *Statistical Science,* 7(2): 183–96.

Cochran, W.G. 1957. Analysis of covariance: Its nature and uses. *Biometrics, 13*(3): 261–81.

Comte, A. 1853. *The positive philosophy of Auguste Comte* (trans. H. Martineau; 2 vols), London: John Chapman.

Darwin, C. 1964. *On the origin of species* (with an introduction by Ernst Mayr). Cambridge, MA: Harvard University Press.

Duncan, O.D. 1966. Path analysis: Sociological examples. *The American Journal of Sociology,* 72(1): 1–16.

——. 1984. *Notes on social measurement, historical and critical.* New York: Russell Sage Foundation.

Duncan, O.D. and B. Duncan. 1955. A methodological analysis of segregation analysis. *American Sociological Review, 20*(2): 210–17.

Elcock, H. 1976. *Political behaviour.* London: Methuen.

Esping-Anderson, G. and A. Przeworski. 2001. Quantitative cross-national research, pp. 12649–55. In N. Smelser and P. Baltes (eds), *International encyclopaedia of the social sciences.* Oxford: Elsevier.

Fisher, R.A. 1925. *Statistical methods for research workers.* Edinburgh: Oliver & Boyd.

——. 1935. The design of experiments. Edinburgh: Oliver & Boyd.

Giddings, F.H. 1901. *Inductive sociology: A syllabus of methods, analyses and classifications, and provisionally formulated laws.* New York: Macmillan.

Goldthorpe, J.H. 2003. Sociology as social science and cameral sociology: Some further thoughts. (Sociology Working Papers No. 2003-07). Oxford: University of Oxford.

Goodman, L.A. 1979. Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association, 74*(367): 537–52.

Heath, A. 2000. The political arithmetic tradition in the sociology of education. *Oxford Review of Education, 26*(3&4): 313–31.

Heckman, J. 1978. Dummy endogenous variables in a simultaneous equation system. *Econometrica, 46*(4): 931–59.

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica, 47*(1): 153–61.

Huff, D. 1954. *How to lie with statistics.* New York: Norton.

Keynes, J.M. 1921. *Treatise on probability.* London: MacMillan.

King, G. 1986. How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science, 30*(3): 666–87.

LaLonde, R.J. 1986. Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review, 76*(4): 604–20.

Lieberson, S. 1992. Einstein, Renoir, and Greeley: Some thoughts about evidence in sociology–1991 presidential address. *American Sociological Review, 57*(1): 1–15.

Lieberson, S. and F.B. Lynn. 2002. Barking up the wrong branch: Scientific alternatives to the current model of sociological science. *Annual Review of Sociology, 28*: 1–19.

Maddala, G.S. 1983. *Limited-dependent and qualitative variables in econometrics.* Cambridge: Cambridge University Press.

Mann, M. 1981. Socio-Logic. *Sociology, 15*(4): 544–50.

Manski, C.F. 1989. Anatomy of the selection problem. *Journal of Human Resources, 24*(3): 343–60.

Mayr, E. 2001. The philosophical foundations of Darwinism. *Proceedings of the American Philosophical Society, 145*(4): 488–95.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior, pp. 105–42. In P. Zarembka (ed.), *Frontiers in econometrics.* New York: Academic Press.

Merton, R.K. 1979. Remembering Paul Lazarsfeld, pp. 19–22. In R. Merton, J. Coleman, and P. Rossi (eds), *Qualitative and quantitative social research: Papers in honour of Paul Lazarsfeld.* New York: MacMillan.

Meyer, B.D. 1995. Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics, 13*(2): 151–61.

Mills, C.W. 1959. *The sociological imagination.* New York: Oxford University Press.

Muthén, B. and L.K. Muthén. 2000. Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24*(6): 882–91.

Oakley, A. 1998. Gender, methodology and people's ways of knowing: Some problems with feminism and the paradigm debate in social science. *Sociology, 32*(4): 707–31.

——. 2000. *Experiments in knowing.* Cambridge: Polity Press.

Rhind, D. 2003. Great expectations: The social sciences in Britain. Available at http://www.city.ac.uk/vco/davidrhind/expectations.html.

Robinson, W.S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review, 15*: 351–57.

Rosenbaum, P.R. 1999. Choice as an alternative to control in observational studies. *Statistical Science, 14*(3): 259–78.

Rosenbaum, P.R. and D. Rubin. 1984. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*(387): 516–24.

Simon, H.A. 1954. Spurious correlation: A causal interpretation. *Journal of the American Statistical Association, 49*(267): 467–79.

Sobel, M.E. 2000. Causal inference in the social sciences. *Journal of the American Statistical Association, 95*(450): 647–51.

Thurstone, L. 1927. A law of comparative judgement. *Psychological Review, 34*: 273–86.

Xie, Yu. 1992. The log-multiplicative layer effect model for comparing mobility tables. *American Sociological Review, 57*(3): 380–95.

PART ONE
THEORIES, DEBATES, AND RESEARCH DESIGNS