Chapter 2
Data, Models, Theory and Reality: The Structure of Demographic Knowledge

2.1 Introduction

The development of demography as a science has been hampered by inadequate attention to theory and to methodology (broad issues of scientific procedure, as distinct from specific techniques or methods). Demography has been preoccupied with the analysis of empirical data, to the neglect of the systematic theory formulation. In turn, theory development has been hampered by the widespread acceptance, if only implicit, of the methodological ideas of logical empiricism. This is the philosophy of science of such mid-twentieth century writers as Nagel, Hempel, Popper, and Reichenbach. But there also is an older intellectual tradition deriving from nineteenth century scientists such as Karl Pearson and Ernst Mach. It is a tradition that finds expression in other social and behavioral science disciplines, of which Skinnerian behaviorism in psychology is a prime example.

The key ideas of logical empiricism are: (a) that empirical science should focus on observable phenomena, and avoid discussion of unobservable entities or processes; and (b) that scientific theory, if possible at all, must be based on empirical generalizations, preferably universal empirical generalizations or laws.[1]

With respect to abstract analytic theory and abstract models, demography has been schizoid. Few demographers would deny the validity of the stable population model or its fruitfulness in generating substantive conclusions, even though few

---

Based on a presentation to a workshop on 'Agent-Based Computational Demography,' at the Max Planck Institute for Demographic Research, Rostock, Germany, 21–23 February 2001; originally published in F.C. Billari and A. Prskawetz [eds.] 2003. *Agent-Based Computational Demography*. Heidelberg: Physica-Verlag, pp. 19–40.

[1] Another important influence on the scientific character of demography, not further discussed here, has been its heavy involvement in largely descriptive work, due to its close association with government statistical bureaus. The central mandate of these agencies is accurate data collection and description, not the overall advancement of demographic science as such. This close association clearly has had great advantages for demography, but also costs.

T.K. Burch, *Model-Based Demography*, Demographic Research Monographs,
DOI 10.1007/978-3-319-65433-1_2

real-world populations closely fit the stable model. By contrast, abstract behavioral theories such as transition theory or the microeconomic theory of demographic behavior are often dismissed because they admit of empirical exceptions or are 'unrealistic.' The tendency has been to draw a sharp line between formal demography, built on necessary relationships, and behavioral demography, built on empirical generalizations regarding contingent relationships.

Insight into these methodological issues can be garnered from a consideration of two leading North American demographers of the late twentieth century, Ansley J. Coale and Nathan Keyfitz. To the best of my knowledge, Coale never wrote systematically on the methodology of demography, so one must seek methodological remarks, often made in passing, in his other writings. For Keyfitz, we have two papers (Keyfitz 1971, 1975), in which he argues that much of our best demographic knowledge has been produced by use of abstract models, not by means of data analysis.

Interestingly, Coale relied heavily on abstract models in much of his earlier work on formal topics such as the demography of population aging, but also on behavioral topics such as the impact of high fertility and rapid population growth on economic development in low-income countries. In his later work on historical fertility transitions, Coale appears to favor a logical empiricist approach, pruning theoretical conclusions to fit the data, and ending with a restatement of transition theory reminiscent of Notestein's writings 40 years earlier. It is arguable that the *theoretical* returns to the heavy investment of time, money and personnel into the European fertility project were not as large as they might have been. Hobcraft (2000) has made the same argument with respect to the World Fertility Survey and its successors (notably the Demographic and Health Surveys).

Keyfitz's assessment of the relative fruitfulness of abstract modeling is neither new nor unique. Descartes favored thought over observation as the way to knowledge. The 'new scientists' such as Francis Bacon favored observation (data). John Locke attempted a synthesis that comes close to a balanced view of empirical science: experience and reflection on experience, or, observation and theory. More recently, Karl Pearson and Ronald Fisher are reputed to have parted ways over the issue of correlational studies of large samples (data) versus experiments on smaller samples to test ideas about mechanisms (theory).

John Platt, in a classic paper on scientific methodology (Platt 1964), recalls a 1958 conference on molecular biology, at which theoretical modelers were criticized by experimentalists. Leo Szilard is quoted as commenting about protein synthesis or enzyme formation that 'If you do stupid experiments, and finish one a year, it can take 50 years. But if you stop doing experiments for a little while and *think* how proteins can possibly be synthesized, there are only about 5 different ways, not 50! And it will take only a few experiments to distinguish these' (p. 348). An experimental researcher is reported to have replied 'You know there are scientists; and there are people in science who are just working with these oversimplified model systems – DNA chains and in vitro systems – who are not doing science at all' (p. 346). The subsequent history of molecular biology suggests who was on the right track.

Keyfitz's view is echoed and supported by a wide variety of writings on scientific methodology, all seeking an alternative approach to logical empiricism, seen as leading to dead ends and theoretical frustration. As examples, I cite and briefly discuss an early statement by political scientist Eugene Meehan (1968), some representative authors of the 'semantic' school in the philosophy of science (Cartwright 1999; Giere 1999), and a call for a return to abstract analytic theory and the search for social mechanisms by sociologists (notably, Hedström and Swedberg 1998).

The critique of logical empiricism common to all these writers suggests the need for greater respect for and attention to abstract models in demography. It also suggests dropping the sharp distinction between formal and behavioral demography, since all good scientific theories or models are in fact formal. The qualitative distinction between theory and models is minimized; at best, the difference is one of scope. And the pervasive but false characterizations of theory as verbal speculation, of modeling or simulation as quantitative speculation, and of empirical research as quantitative bedrock may seem less cogent.

Finally, given the complexity of the real world and the ability of the computer-aided scientist to handle larger amounts of complexity, in the future much of our fruitful theoretical work will consist of computer modeling (Burch 2002, and Chap. 3 below). The genre of *agent-based modeling* will likely occupy a central place in this work. It provides a feasible approach to study interrelations between the macro- and micro-levels in demography – exploring links between individual decisions and aggregate demographic patterns, a realm that up until now has resisted analysis (but see Courgeau 2004). It also can introduce rule-based behavior into complex demographic models that formerly relied on purely stochastic assumptions (e.g., Monte Carlo simulations). Microeconomic decision-making models (as well as decision models from sociology and social psychology) can be used in more than a 'heuristic' manner (see McNicoll 1992).

2.2 The Methodology of Ansley J. Coale

By all accounts, Ansley Coale was one of the most versatile, creative, and influential demographers of our era. His contributions range widely across the field – extensions of stable population models; stunning clarifications of the relative roles of fertility and mortality change on age composition; parametric modeling of demographic behavior (fertility, marriage, mortality); pioneering work on the impact of fertility and population growth on economic development; historical studies of fertility decline in Europe; the demography of China; powerful evaluations of the completeness of census enumeration – a complete list would be still longer. His work characteristically has a sure-footed and direct quality often lacking in social science: problems are stated with great clarity, and solutions provided; there is a sense of closure. His technical innovations are geared toward important substantive issues, and typically have proven useful in further empirical research by others. Although mathematically astute, he seldom if ever did mathematics for the sake of mathematics.

But like most demographers, he seems not to have been self-conscious and explicit about the methodology of demography (logic and epistemology) as opposed to technique.[2] Implicitly, his work might be taken to suggest ambivalence regarding the proper roles of data, models, and theory, or at least some changing emphases over time.

A leitmotiv of his career is formal mathematical modeling of demographic dynamics, popularizing and extending the work of Lotka and other early pioneers. This work relies on mathematically necessary relationships in highly abstract population models, with the stable model as central. Generalizations emerge from the models rather than from extensive empirical research. A prime example is his work on the demographic determinants of age structure (see, for example, Coale 1956). Using the stable model and the standard projection model, he concludes that fertility change generally has more impact on age structure than mortality change, and that the effects of mortality change are heavily dependent on the age-pattern of change, with the possibility that some patterns of change can make a population younger, not older. Coale uses data to calculate the models, of course, but his generalizations depend on the manipulation of relatively simple abstract models, not on the systematic analysis of empirical data. Plausible but constructed data could have served as well.

His monograph with Hoover (Coale and Hoover 1958) also involves abstract modeling rather than broad empirical research, but on issues generally viewed as behavioral rather than formal, and involving a considerably more complex model. The core of this work is a linking of a standard population projection model with a standard economic growth model. The model is fit to the Indian case at length, and to the Mexican case more briefly. There is some discussion – but no systematic empirical research – of the wider applicability of the analysis. It was what today is known as large-scale simulation. The researchers were no doubt hampered by the absence of computer tools we take for granted, although they did engage in what would now be called sensitivity analysis.

Coale was more deeply involved in the demographic parts of the work, Hoover in the economic. But they stress that there was co-operation of both in all parts. Sometimes they are clear about the abstract character of their analysis: 'Our calculations entail a "model" of the Indian economy, designed to take into account as realistically as an extremely simplified model can the main relevant features of that economy. . .' (p. 259). But the conclusions are set forth as general, applying to India, Mexico, and to most other low-income nations in the developing world.

[2]As has been pointed out many times, excessive self-consciousness of methodological issues can hamper scientific progress. The development of demography has probably benefited from the efforts of investigators who simply got on with the job with the tools at hand. It also is true that theorizing without an adequate empirical base often is sterile. My methodological critique is aimed more at present and future work than at the past.

They seem at times to overlook an element of circularity that is characteristic of all simulation, namely that the conclusions follow from the assumptions. For example, they conclude: '... through this whole gamut of projections, despite the wide variation in rates of progress that they imply, the differential associated with reduced fertility is remarkably persistent and stable' (p. 281).[3] But earlier they have told us that the model does not contain all important growth determinants, but only 'the growth determinants most clearly affected by the difference between our alternative rates of population growth' (p. 260). Since alternative rates of population growth in their scenarios depend mainly on alternative fertility assumptions, in hindsight the conclusion quoted just above is inevitable. And it is completely possible that the model is not relevant to concrete cases in which the omitted growth determinants are crucial.

Several passages in the book suggest that the abstract and speculative character of their work was partially obscured by the fact that the analysis was quantitative, that it dealt with concrete cases, and that it dealt with a very specific question – what difference would different fertility patterns make to economic development? In his preface to the work, Frank Notestein, who generally disliked abstract theory, seems to have been misled, describing it as a '...highly original demonstration of the way careful *factual analysis* can illuminate the vital interrelationships of economic and population change' (pp. v–vi, emphasis added). But clearly it is factual only if one accepts as facts theoretical and modeling assumptions as well as quantitative empirical data. And perhaps it is factual by contrast with pure theoretical speculation, not grounded in the details of a concrete case. But it certainly is not empirical work in the ordinary sense of that term. Of course, this is easier to see this in hindsight, given many decades of experience with computer modeling.

The important point is that general propositions emerging from the project were based on the model not just on empirical data. Indeed, critics of Coale-Hoover (e.g., Kuznets, Easterlin) criticized it precisely because comparative empirical research showed no strong or regular relationship between population growth rates and economic development. But this in no way diminishes the importance of their work. Like all good models, their simulation of population growth and development provided crucial insights into mechanism at work and a systematic framework for discussion and further research. In addition, their analysis almost certainly provides a relevant explanation as to why some nations with rapid population growth – then and now – have shown so little economic progress.

When Coale turned his attention to fertility transitions, the orientation became more empirical. In one of his earliest papers on the European fertility project (Coale 1965), he presents his indirectly standardized ratios and a few early results at the national level. The paper does not explicitly deal with classic transition theory, but implicitly calls it into question. Methodological comments made in passing suggest

[3]This language is almost identical to that would be used by the *Limits to Growth* researchers, some years later. But it did not deter harsh criticism of this study by many demographers and economists. See Bardi, 2011.

a radical logical empiricism. Speaking of the decline of marital fertility, he comments ‘There are few, if any, universally valid generalizations about the circumstances under which neo-Malthusian fertility reduction occurs’ (p. 5). After a list of frequently hypothesized causal factors, he notes that ‘Examples can be found illustrating the presumed influence of each of these factors, but counter-examples or exceptions are nearly as prevalent’ (p. 6). He concludes: ‘Fertility reduction seems to be a nearly universal feature of the development of modern, secular societies, but its introduction and spread cannot yet be explained by any simple, universally valid model or generalized description’ (p. 7). Looking to the future he expresses the hope that further empirical research ‘tracing the decline of fertility more systematically, and by geographic units smaller than nations, will certainly establish a fuller record of fertility reduction, and will perhaps make possible generalizations about the causes of the decline’ (p. 7).

Eight years later Coale (1973) deals with what would usually be called theoretical issues, in a paper on the demographic transition. But the emphasis is still on the search for universal empirical propositions. Interestingly, he never uses the word *theory,* either with reference to Notestein’s work or his own (the word does not appear anywhere in the paper). He speaks instead of the ‘idea’ of demographic transition. It is difficult to know just what was intended by his studious avoidance of the word *theory,* which would have seemed quite natural in the context.

In fact, the paper eventually produces very broad statements that most social scientists would view as theory. Coale posits ‘the existence of more than one precondition for a decline.’ ‘Three general prerequisites for a major fall in marital fertility can be listed’: (1) it must be within the calculus of conscious choice; (2) reduced fertility must be advantageous; (3) effective fertility control techniques must be available (p. 65). The language is borrowed from mathematics; the three preconditions or prerequisites are in fact ‘necessary conditions’ for fertility decline (p. 69). A weakness of ‘the idea [sic] of the transition is that it tells us that a high degree of modernization is *sufficient* to cause a fall of fertility, but does not tell us what degree (if any) of modernization is necessary to produce a fall’ (p. 69). He notes that one or more of the three preconditions can exist in the absence of modernization.

Coale acknowledges many good points about ‘the idea of the transition’ (Notestein’s transition theory] but faults it on its inability to make more than qualitative statements about the course of demographic and fertility transitions. He notes, for example, that with respect to developing countries, transition theory was ‘accurate in direction but inaccurate in detail, with respect to mortality’ (p. 68). Transition theory was qualitatively correct regarding the past of developed countries and qualitatively correct in its predictions for less developed countries. But, ‘In neither instance does it specify in terms that can be translated into quantitative measures, the circumstances under which the decline of fertility began’ (p. 68).

But Coale’s three preconditions clearly are subject to the same criticism, especially since they are not presented as quantitative variables. He speaks of ‘the degree of change that must occur before the preconditions are introduced,’ but does not always discuss the preconditions themselves as matters of degree, using words that suggest a 0–1 variable – whether the preconditions are 'present' or 'absent' (p. 66). There is little attention to the issue of how they might be quantified and operationalized.[4]

Coale's last major statement on fertility transitions is in his introductory chapter for the multi-authored summary volume on the project (Coale and Watkins 1986). The spirit of this essay is different from that of the 1973 paper, with a return to reliance on abstract models to gain insight into population dynamics. There is, for instance, considerable discussion of what might be called a cyclic logistic model to describe pre-modern or even pre-historical population dynamics. Population growth leads to rising mortality; populations react by reducing marriage and/or fertility or otherwise reducing population growth; mortality declines to former levels; and the cycle starts over. Interestingly, the model is purely qualitative, and, of course, there is little empirical evidence to support it, at least for the earlier, pre-modern periods. On transition theory, Coale seems to have given up the hopes expressed in earlier papers that the project would arrive at 'universal empirical generalizations.' The three preconditions are not mentioned. One long paragraph (p. 24) summarizes the causes of transitional mortality decline in broad terms that would not have passed muster by the standards of his 1973 paper. Ultimately, he writes of the fertility transition in language not so different from that of Notestein 40 years earlier, speaking of 'typical' patterns of transition and some exceptions (pp. 28–29). There is no attempt to quantify the 'idea of transition' beyond the presentation of empirical measurements of fertility and nuptiality, and their time trends and intercorrelations with a limited number of independent variables. It is as though the sheer mass of data has led to an abandonment of attempts to develop new and better theoretical ideas or models, including a revised and more rigorous statement of transition theory.

This is a long story, but it makes an important point: a massive 20-year project with substantial resources and collaboration by many first-rate demographers did not result in a major improvement in theory, nor in consensus on transition theory. This is not to deny the immense scientific value of the work as an empirical, descriptive study. And, in Coale's mind, description may well have been its central aim. Recall the quote cited earlier: 'tracing the decline of fertility more systematically... will *certainly* establish a fuller record of fertility reduction, and will *perhaps* make possible generalizations about the causes of the decline' (11, p. 7, emphasis added). But theory did not flow from the data; the methodological stance was such that theory would flow from universal empirical generalizations, and these were not forthcoming.

John Hobcraft has recently commented in a similar vein on the small theoretical returns to the large number of comparative fertility surveys under the aegis of the World Fertility Survey and its successors such as Demographic and Health Surveys. Entitled 'Moving beyond elaborate description: towards understanding choice about parenthood' (Hobcraft 2000), the paper argues that 'the results [of these

[4]But see Lesthaeghe and Vanderhoeft (1997).

surveys] did not live up to my own or to others' highest expectations; comparative analysis projects today are much less common; the Demographic and Health Surveys, the daughter of the WFS, have never had a serious comparative analysis capacity (beyond the mainly descriptive Comparative Studies)' (p. l). He speaks of 'meagre returns,' and of 'meagre progress to date in moving forward our real understanding of fertility behavior through cross-national comparative surveys' (p. 11). He adds that 'a profound shift of emphasis is required in order to make real progress' (p. l), that is, progress towards understanding. Hobcraft's diagnosis:

> . . .the main problem for comparative analysis, over and above the sheer scale of data manipulation, has always been the rather limited number of explanatory variables which are sufficiently standardized and accorded enough credibility to be collected in every country. In part, this problem arises from a lack of a commonly accepted theoretical framework for understanding fertility behavior, but it is also arguable that we shall never remedy the problem without better agreement and testing of comparable information [p. 2].

He seems to agree with Griffith Feeney, who earlier (Feeney 1994) noted that the surveys in question contained a lot of data but not necessarily the right data for testing or developing explanations and theory. Hobcraft comments, with respect to the comparative fertility surveys, that there has generally been 'a lack of conceptual and theoretical clarity about what elements should receive priority,' and that 'an explicit theoretical orientation has been lacking' (p. 4), noting that the surveys were done for policy or even political purposes, not primarily to advance science. Hobcraft's remedy would be to collect more and better data, with variables to be selected based on the best theoretical thinking about the determinants of fertility. The emphasis would be on variables that are comparable cross-culturally, and especially on variables relating to the fertility decision-making process (defined broadly to emphasize the decision to become a parent), and on community-level variables. Analysis of such data would aim at 'global models,' models involving 'not just the same range of regressors but also the same parameters'(p. 3):

> A deeper understanding would involve a closer specification, whereby the strength of the relationship was the same everywhere net of the correct range of other controls, or, the development of models which incorporate the factors which bring about variations in the strength of the relationship as a step towards the fuller model' [p. 4].

Hobcraft thinks 'progress toward such models is essential for good theory' (pp. 3–4), but not that theory will flow from the data. Rather, the appropriate data-collection and analysis will necessarily be informed and guided by theory. Of comparative description and 'detailed society-specific accounts,' he concludes that 'Neither holds out great hope for reaching general theoretical understanding' (p. 5). In the final analysis, however, Hobcraft's faith in future progress tilts towards better data and more sophisticated data analysis. And the goal of 'global models' suggests an assumption that widespread, if not universal, empirical generalizations are there, if only we can find them. The logical empiricist approach maintains some hold.

2.3 Nathan Keyfitz on the Fruitfulness of Abstract Modelling

A very different approach is found in the work of Nathan Keyfitz, one of a relatively few leading demographers who wrote specifically on scientific methodology (as opposed to technique) (Keyfitz 1971, 1975). In answer to the title question ‘How do we know the facts of demography?’ Keyfitz comments ‘Many readers will be surprised to learn that in a science thought of as empirical, often criticized for its lack of theory, the most important relations cannot be established by direct observation, which tends to provide enigmatic and inconsistent reports’ (Keyfitz 1975, p. 267). Citing E.O. Wilson, he speaks of ‘the resistance of data to generalization’ (p. 286).

To illustrate his point, he first looks at the issues of the interrelations among growth and proportion of elderly, and of the relative impact of fertility and mortality on age structure, both of which are best answered using population models. In another section, entitled ‘No model, no understanding,’ he notes that statistical observations of differential incidence of breast cancer remain largely unexplained, and comments ‘Here is just one more question that is unlikely to be solved by any volume of statistics by themselves’ (p. 276).

He then considers the issues of the effect of marriage delay on completed fertility, of promotion in organizations, and the relationships between development and population growth – all questions involving behavioral models, on which there is less consensus than on the stable model.

The important point is that Keyfitz attributes our accepted answers to these issues to work with theory or models. In addition, he does not make a sharp distinction between formal models (e.g., the stable model) and behavioral models (e.g., transition theory). The logical procedures involved in the statement and use of the two sorts of models are seen to be much the same. In a final section entitled ‘The psychology of research,’ he comments:

> The model is much more than a mnemonic device, however; it is a machine with causal linkages. Insofar as it reflects the real world, it suggests how levers can be moved to alter direction in accord with policy requirements. The question is always how closely this constructed machine resembles the one operated by nature. As the investigator concentrates on its degree of realism, he more and more persuades himself that his model is a theory of how the world operates’ [p. 285].[5]

Keyfitz’s thought and language anticipates that of leading proponents of the semantic or model-based school of the philosophy of science, such as Giere and Cartwright.

[5]Note the equation of model and theory in the closing sentence.

2.4 A Model-Based View of Science

The logical empiricist view of science has dominated social science, including demography, in the latter half of the twentieth century. By this view, theory – a summary of what is known in a field – is made up of statements derived by further generalization and abstraction from universal empirical generalizations or laws. Explanation, in this perspective, consists of subsuming some fact under a broader general proposition, which in turn is subsumed under a still broader generalization, and so forth – the so-called 'covering law'approach to explanation. Laws are subject to empirical test, to be 'proven,' or, in keeping with the widespread Popperian view, to survive efforts at falsification.

Not all social scientists have adhered to the dominant view, as Keyfitz's essays attest. As a discipline, economics has departed from literal logical empiricism at least to a degree sufficient to allow and encourage abstract analytic theory and models, even if they seemed to some 'unrealistic.' Milton Friedman's essay on 'The Methodology of Positive Economics' (Friedman 1953) is representative. Theories and models are viewed as analytic tools that may or may not be useful for analyzing specific empirical phenomena, with their usefulness judged by their ability to yield understanding, and to predict phenomena not previously observed. Friedman acknowledges the formal or logical character of models, but still speaks of 'falsifying' or 'validating' them in terms of their ability to predict with regard to whole classes of empirical phenomena. The notion of the search for universal laws, so characteristic of logical empiricism, lurks just below the surface of the essay.

Many sociological theorists have attacked logical empiricism, but the attack often has been aimed at positivism or empiricism as such, especially in its quantitative forms, not just at the approach advocated by Nagel and Hempel. The resulting schism between theory and empirical research in sociology remains strong.

A frontal attack on logical empiricism that rejects neither empiricism nor formalization is to be found in an extraordinary but neglected work by the political scientist Eugene Meehan, *Explanation in Social Science: A System Paradigm* (Meehan 1968). Meehan argues that the search for universal empirical generalizations is largely doomed to failure in social science, since such generalizations are, and are likely to remain, few and far between. He proposes instead explanation by 'systems' (contemporary terminology would call them 'models' or 'theory'), formal structures which entail or logically imply the phenomena to be explained. The systems are true only in the sense of being logically coherent. The relevant empirical question is not whether they are empirically true or false, but whether, when 'loaded' with specific empirical information, they sufficiently resemble some portion of the real world (in his words, whether the system is sufficiently 'isomorphic' with reality) to enable the analyst to accomplish his or her purpose.[6] The

[6]In mathematics, the word isomorphism refers to a one-to-one correspondence between two systems. In chemistry and biology, the word refers only to some similarity between systems. Meehan clearly has in mind the latter meaning, which admits of degrees of isomorphism.

purpose can be prediction, which enables human beings to adjust to the world, or, even better, explanation, which provides insight into process and mechanisms such that one could at least in principle control the real-world system in question.

Meehan is wary of trying to explain large classes of events for the same reason that he is pessimistic about discovering universal empirical generalizations – classes of social events are often defined independently of attempts at scientific explanation, and typically are not particularly homogeneous.[7] The focus is on specific concrete events, but the key tool is abstract analytical reasoning.

A more recent, but similar, approach to explanation in social science is *Social Mechanisms: An Analytical Approach to Social Theory* (Hedström and Swedberg 1998), a collection of essays calling for a return to so-called 'middle-range theory,' and generally rejecting the logical empiricist emphasis on empirical generalizations.

In their introductory essay, the editors call for 'an analytic approach that systematically seeks to explicate the social mechanisms that generate and explain observed associations between events' (p. 1). They contrast a mechanism approach to science with pure description, with theory as labelling or classification, and with the search for 'laws'. They quote Francis Crick, co-discoverer of the structure of DNA, to the effect that contemporary biologists prefer to think in terms of mechanisms not laws, commenting that 'The reason for this is that the notion of "laws" is generally reserved for physics, which is the only science that can produce explanations based upon powerful and often counterintuitive laws with no significant exceptions' (p. 3).[8] Mertonian middle-range theory, in their view now out of favor, is seen as an appropriate middle ground between pure description and the search for social laws.

The search for mechanisms, or underlying processes, is contrasted with statistical analyses of interrelationships among variables: 'The search for mechanisms means that we are not satisfied with merely establishing systematic covariation between variables or events: a satisfactory explanation requires that we are also able to specify the social "cogs and wheels" that have brought the relationship into existence' (p. 7). This comment is taken to apply, not just to simple regression models, but also to path models and other structural equations models. Another way to put it is that reasoning in terms of mechanisms tries to figure out what is happening in the black box between a measured input I (including multiple inputs, as in a regression model) and a measured output O. A mechanism is a systematic set of statements that provide a plausible account of how I and O are linked to one another (compare Meehan's 'system').

---

[7]Chemical compounds may behave the same everywhere because they are the same everywhere. But demographic categories like *marriage* or *fertility transition* pertain to very heterogeneous classes of events. See Burch and Belanger (1999).

[8]For a different view of the status of laws in physics, however, see Giere (1999) and Cartwright (1983, 1999).

The approach is explicitly contrasted with the covering-law model of explanation advocated by Hempel, Nagel and their followers. In this latter approach, if the covering law is only a statistical association, which is the norm in social science according to Hempel, then 'the specific explanation will offer no more insights than the law itself, and will usually only suggest *that* a relationship is likely to exist, but it will give no clue as to why this is likely to be the case' (p. 8). Finally, there is no attempt to prove that a model is true in the sense of empirically valid:

> The choice between the infinitely many analytical models that can be used for describing and analyzing a given social situation can never be guided by their truth value, because all models by their very nature distort the reality they are intended to describe. The choice must instead be guided by how useful the various analytic models are likely to be for the purposes at hand [Hedström and Swedberg p.15].

Keyfitz, Meehan, and Hedström, and Swedberg gain considerable support from recent work in the philosophy of science by members of the semantic' school. These philosophers challenge the classic logical empiricist view, arguing that it is neither an accurate description of what scientists actually do, nor a good guide to what they should do for their work to be fruitful. In this newer view, scientific laws are seldom, if ever, true representations of reality, but at best idealizations of certain features of an indefinitely complex real world. Nor are they so much discovered in nature as constructed by the human mind. Cartwright (Cartwright 1983, 1999) speaks of *nomological machines:* models created by the scientist generate laws rather than vice-versa (recall Keyfitz's use of the machine analogy).

Giere (1999) notes that most scientific laws are not universal, and that they are in fact not even true: 'Understood as general claims about the world, most purported laws of nature are in fact false. So we need a portrait of science that captures our everyday understanding of success without invoking laws of nature understood as true, universal generalizations' (p. 24). The reason is that any law of nature contains 'only a few physical quantities, whereas nature contains many quantities which often interact one with another, and there are few if any isolated systems. So there cannot be many systems in the real world that exactly satisfy any purported law of nature' (p. 24).[9]

For Giere, the primary representational device in science is not the law but the *model,* of which there are three main types: physical models; visual models; and theoretical models (Giere prefers the term 'model-based view' of science to the older philosophical term 'the semantic view'of science). Models are inherently abstract constructions that attempt to represent only certain features of the real world. They are true only in the sense that definitions are true. The question of whether they are empirically true is irrelevant, since they cannot be. The relevant question is whether they correspond to some part of the real world in (a) some respects (b) to a sufficient degree of accuracy for (c) certain well-defined purposes

---

[9]Giere, with considerable training in physics, draws many of his examples from that field. If his conclusions apply to physics, they would seem to apply with even more force to other scientific disciplines, for example biology or the social and behavioral sciences.

(compare point b to Keyfitz's phrase 'degree of realism' or Meehan's notion of sufficient isopmorphism). Giere gives the example of the model for the earth-moon system, which is adequate to describe and account for the moon's orbit and perhaps for putting a rocket on the moon, but is inadequate to describe the Venus-earth system. The prototype of scientific knowledge is not the empirical law, but a model plus a list of real-world systems to which it applies.

A model explains some real-world phenomenon if (a) the model is appropriate to the real-world system in the three respects noted above, and (b) if the model logically implies the phenomenon, in other words, if the phenomenon follows logically from the model as specified to fit part of the real world. It would never occur to most physical scientists to add the second condition. But in social science, including demography, we are so accustomed to loose inference that its explicit statement is necessary.[10]

Note that in this account of science, all models *are formally* true (assuming, of course, no logical errors or internal contradictions), that is, true in the way that explicit definitions are true. The empirical question then becomes one not of empirical truth or validity, but whether a valid model applies to a specific empirical observation.

Of course, some models are more widely applicable than others, and, other things equal, science will prefer the model with the widest applicability. In demography, for example, the fundamental demographic equation is true by definition, and applicable to every well-defined real population (neglecting error in data). The exponential growth formula is true by definition, and, for calculating the average annual growth rate over a period is also applicable to every real-world population. For describing a population's actual growth trajectory, however, the exponential growth formula applies to some populations, but not others.

A behavioral model such as the theory of demographic transition can be stated in such a way that it is formally true. Its status has been a matter of debate for over 50 years. But it is worth noting, in terms of Giere's criteria of applicability, that it correctly represents many actual cases of mortality/fertility decline, at least in qualitative terms.[11]

In my reading of these various accounts of science, they come close to what has long been the standard approach in the literature on mathematical modelling, and more recently of computer modelling. A model is an abstract construct that may or may not be useful for some well-defined purpose. In science, that purpose often will be explanation or prediction as opposed to practice. And in some schools of

[10] As noted above, the notion of explanation as logical inference from a model is central to Meehan's *Explanation in Social Science* (1968). The need for rigorous logic is emphasized by Platt (1964).

[11] An interesting point about transition theory is that there has been a tendency to dismiss it as not fitting all cases or as not providing details of timing, pace, etc. There seems to have been relatively little effort to accept it as a valid model and work towards a more precise specification by defining functional forms for fertility or mortality decline as functions of development, and parameters representing size of time lags, slopes, etc.

computer modelling, the emphasis is on less abstract models, trying to capture more of the complexity of the real world. But the central ideas are the same.

The model-based approach to science described above prefers not to make a sharp distinction between a model and a theory. Some authors distinguish the two on a general/specific axis; but then differences are in degree only not in kind. Giere speaks of 'theoretical models,' and sometimes describes a 'theory' as a collection of such models.

Note that this position has nothing to do with the view that science is totally a social construction. A good model is good precisely because it captures some important aspects of the real world. In Giere's words, there is 'realism without truth.'

2.5 Elements of Science

In thinking about some of the above issues involving the description and evaluation of scientific knowledge, it is useful to think in terms of four distinct but interrelated sets of elements:

1. Reality: the real-world as it exists independently of human knowledge;
2. Theory: coherent sets of ideas about how some portion of reality works;
3. Data: observations and measurements on some real-world system;
4. Models: abstract, but rigorous and specific, representations of reality, based on theory or data or both.[12]

In some sciences, but not in sociology or demography, there is another important element, namely, the controlled laboratory experiment. This is reality in the sense that it exists independently of the human mind, but it often is artificial reality since it does not occur in the natural world without human intervention.

Some further comments on these elements:

The first element implies a belief in the existence of objective reality, an assumption that will be taken for granted and not further discussed here.

The second element implies that it is possible to have theories that are not mere fantasy, but express important insights into the real world, insights that can provide a basis for explanation, prediction, and, sometimes, control. If one believes that such theory is not possible with respect to human behavior, whether because of complexity, free will, ideological bias, or some other reason, then there is no discussion.

[12]As noted earlier, the distinction between a theoretical model and a theory is not fundamental; differences are on a general/specific axis. There are, however, big differences between theoretical models and purely mathematical or statistical models.

The word *data* refers to a limited set of measurements on some portion of the real world. Note that virtually all non-trivial data sets involve error and various kinds of abstraction.

The word *model* is troubling in its ambiguity, but the ambiguity reflects actual usage of the word in different contexts.[13] A richer set of terms would help. In the meantime, several distinctions are necessary. The emphasis here is on the distinction between models that represent, usually in more specific form, a set of theoretical ideas about some real-world system, and models that represent a set of data or a structure contained therein – *theoretical models and empirical/statistical models.* But these two categories are not mutually exclusive; there are many hybrid or intermediate types. The distinction between theory and model is appropriate in this context, but as noted earlier, not in all contexts.

There are many modeling exercises in which the data have the upper hand, as it were, with little input from theory. An extreme case is that of *approximating functions,* where a functional form is found that best represents a data set (two or three-dimensional), without any regard to the substantive meaning of the function or its parameters. This typically will be done for purposes of smoothing, interpolation, or extrapolation. The TableCurve software provides a good explanation of this approach to modeling (Jandel Scientific 1996). It contrasts the use of approximating functions with what it terms *parametric functions,* in which the parameters can be given meaning. Parametric functions in turn can be theory based, with the functional form and its parameters mirroring some relevant theory. ‘A parametric model with an underlying physical foundation relates dependent variable Y to the independent variable X via a postulated physical relationship between them’ (pp. 3–19). Or, in the absence of an underlying theoretical model, the parameters may simply characterize the data set. ‘Here features of the data are known to have a direct relationship with the underlying ... factors, but there is no quantitative underlying theoretical model’ (pp. 3–19) – the system producing the data is treated as a black box. Parameters may describe such features of the data as maxima and minima, transition points, and slopes.

Multivariate statistical modeling often resembles atheoretical parametric modeling as described above. The parameters have substantive meaning (often as slopes), but the mechanism or underlying process leading from inputs to outputs is not represented by the model. The selection of variables may be based on theoretical considerations, but just as often is based on common sense, availability of data, or previously observed empirical correlations. The statistical model is a representation of the structure, typically linear structure, of a data set. It is an abstract representation of an abstract data set using abstract concepts for measurement. But a statistical model is not purely empirical, fully determined by data. There must be

[13]The English language does not help in that there is only one verb form in current use, that is, *to model*. French distinguishes *modeler,* as in the acts of a sculptor, and *modeliser,* as in the fitting of a statistical model. The *Oxford English Dictionary* lists *modelize* as an obsolete English word, but the meanings of it and *model* do not seem to parallel the French distinction. Insofar as theory construction is a creative act of imagination, the French *modeler* seems the more apt word for it.

a priori assumptions about what variables to include and about functional form. Such a model may be best fitting, but only from a very limited set of all possible models. A single-equation multivariate statistical model is not, or at least should not claim to be, a representation of the real-world system on which the data were measured.[14] Structural equations models such as path analysis, introducing such concepts as direct and indirect paths of influence or of mutual causation, are a move toward a more faithful representation of the actual real-world system.

A good example of a theory-based model is provided by Hernes' (1972) study of first marriage. Hernes starts with behavioral assumptions: a cohort has an underlying 'marriageability' or eligibility for marriage; for the individual, this marriageability tends to decline with age; pressures to marry increase as the proportion already married rises, but only up to the point where the scarcity of potential partners begins to take effect. These theoretical ideas are then expressed by means of a mathematical function, an asymmetric logistic, with the asymmetry due to an exponential decline in marriageability with age. Finally, the parameters are estimated from census or survey data on proportions ever-married by age. The parameters are estimated from data, and they take their meaning from a theory. In a sense, the equation models both the theory and the data. This analysis can be contrasted with a blind, mechanical fit of some mathematical function to a given data set on first marriage by age (an approximating function), or with a statistical model of event-history data on first marriage in relation to a set of co-variates, as in a hazards model, based only loosely on theory.

These distinctions are far from definitive. But the central distinction between modeling theoretical ideas and modeling a specific data set is paramount. ModelMaker, systems modeling software, makes a distinction between *empirical models* and *simulation or mechanistic models* (Walker 1997). Empirical models describe variation in 'some observed data for a phenomenon which shows how it varies in relation to other factors' (p. 7). 'Simulation models try to describe a number of sub-processes which can be combined to represent the behavior of a larger more complex system' (p. 9). The description of 'empirical models' – models of data – by Edwards and Hamson (1989) is worth quoting in full:

> An empirical model is one which is derived from and based entirely on data. In such a model, relationships between variables are derived by looking at the available data on the variables and selecting a mathematical form which is a compromise between accuracy of fit and simplicity of mathematics.... The important distinction is that empirical models are *not* derived from assumptions concerning the relationships between variables, and they are *not* based on physical laws or principles. Quite often, empirical models are used as 'submodels,' or parts of a more complicated model. When we have no principles to guide

[14]See the important distinction by Abbott (1988) of the *representational* versus the *entailment* interpretations or uses of the general linear model. The entailment use of the general linear model reasons: given a theory about the social world, a certain linear structure should be observed in relevant data. The representational use involves the fallacy, in Abbott's view, of mistaking the linear model for a representation of the real-world system – the fallacy of reification or misplaced concreteness.

us and no obvious assumptions suggest themselves, we may [with justification] turn to data to find how some of our variables are related [p. 102].

Both previous citations raise the issue of complexity, which cuts across the theory/data distinction. Hernes uses a relatively simple theory and model to account for a relatively simple data set, describing a relatively simple observation – proportions first married by age. Hernes does not attempt to model the complete marriage system, which would have to deal with issues such as endogamy/exogamy, spouse selection (by couple or by parents), premarital conception, betrothal, and post-marital residence – all factors bearing on age at first marriage.

And so, we must keep in mind at least three kinds of modeling: modeling a limited set of theoretical ideas; modeling a limited set of data; and modeling a relatively complex portion of the real world, coming closer to providing a replica – generally known as *large-scale simulation*. The last will of necessity require many theoretical ideas and much data, if it is to be at all successful, but success, as always, must be defined in terms of purposes.[15]

2.6 Assessing Scientific Knowledge

The aim of empirical science is to understand some portion of the real world. How well we do in this regard can be seen in terms of the closeness of fit among theory, data, various kinds of models, and the real world. Empirical social scientists are used to thinking in terms of goodness of fit between a statistical model and data. We are less used to thinking in terms of the other relationships, shown schematically in Fig. 2.1.[16]

The broken lines and question marks between each pair of elements are inserted to emphasize that the closeness of one element to another is an open question. The distance of each of the above links may vary, and no one of them is necessarily or consistently shorter than another. It is important to include reality in a diagram such as this, since, insofar as we can know it, reality is the ultimate reference point of all empirical scientific work. But it often does not appear in schematic diagrams of the scientific process.

In an older version of positivism, associated with Karl Pearson, theory and theoretical models drop out of the diagram. The most one can expect is to find stable correlations among data (Turner 1987). In the newer logical empiricism of Nagel and Hempel, theory is possible, but it must be based on empirical generalizations arrived at by statistical analysis of data. In either case, it is assumed that data and empirical models are closest to reality. In the latter case, data and empirical

[15]This approach favored in Casti's *Would-be Worlds* (1997), although he is clear that there are many kinds of modeling, suitable to different purposes.

[16]Note: this diagram is a modified version of that in the original publication.

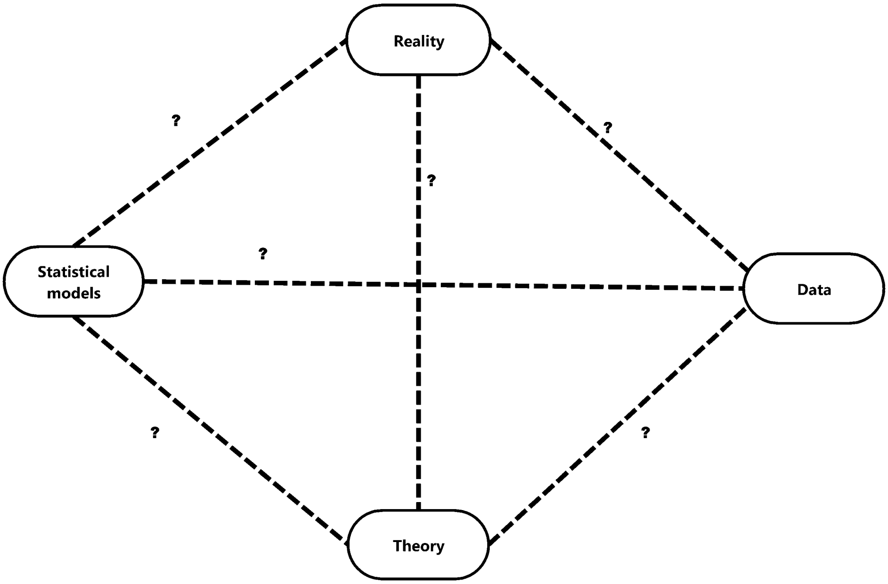

Fig. 2.1 Relations among data (observation), statistical models, theory, and reality T*heory* includes general theory and smaller theoretical models. Broken lines and question marks suggest that, in any concrete instance, the distance between any two elements is an open question; it cannot be specified *a priori*

models are the explicit foundation for theory, and a gatekeeper with respect to theoretical ideas.

But this assumption is gratuitous; it needs to be demonstrated in each case and in general. A good theory or a theoretical model can come closer to reality than data and statistical models. If nothing else, theoretical models can include variables which we know almost certainly are important, but which have not been measured or cannot easily be measured. This point is seldom recognized in the demographic literature. An exception is an eloquent statement by Bracher *et al.* (1993), in their state-of-the-art analysis of Australian divorce, using unusually rich event-history data. They comment in the concluding section:

> However detailed and comprehensive the ‘explanatory’ factors we have had at our disposal, they are, after all, only dim reflections of the possibly un-measurable factors that keep marriages together or drive them apart; and it would be unfortunate if the apparent sophistication of the measures were to obscure this fact [p. 423].

Theoretical models also can more easily incorporate non-linear functional forms and feedbacks, and deal with process and mechanisms. They can paint a plausible picture of how a real-world system works, in a way that no data-bound statistical model ever can. Whether that picture fits that real-world system closely enough to serve some purpose remains an empirical and practical question. And so, empirical data are necessary to describe selected features of real-world systems and observable phenomena to be explained. And, they may be necessary to help establish how closely a theoretical model fits the real world – they can never completely determine this. But data and empirical models do not close the understanding gap. This requires theory.

A critical history of twentieth century demography would show, I believe, that demography has tended to hold theory at arm's length, only grudgingly giving it a small place in our discipline. The same has been true with traditional mathematical modeling [apart from the core mathematical models such as the stable population] and with later developments in computer modeling or simulation. The demographic mind has viewed data as solid and real, theory and simulation as airy and fanciful. A growing body of opinion in the philosophy of science suggests that this view is faulty, as does the scant theoretical progress in the discipline, symbolized by the fact that transition theory is not in appreciably better shape now that it was 40 or 50 years ago. More, and more informed, attention to theory is needed if demography is to be a science as opposed to a body of techniques, or a branch of applied statistics.

2.7 Coda: On the Dangers of Dichotomies

Demography, along with other social science disciplines, has been plagued by the tendency to think, act, and institutionalize in terms of strict dichotomies: theory vs. empirical research; quantitative vs. qualitative; model vs. theory; scientific law vs. hypothesis; formal vs. behavioral demography, to mention a few of the mere obvious. Different kinds of work and different approaches to science tend to be seen as discrete, polar opposites. Often, they are seen as opposed, in the sense of being hostile to one another.

But some of these dichotomies are false; others get in the way of the harmonious working together of the many parts of the scientific endeavor. The distinction, for example, between quantitative and qualitative overlooks the historical fact that that most quantitative tool of science – differential equations – often deals only with qualitative solutions. In the more contemporary realm of computer simulation also, it is often the qualitative results that are of most importance. The precise quantitative results, after all, are essentially arbitrary.[17] In the burgeoning field of 'qualitative methods,' on the other hand, it is ironic – but a healthy development – to see the relevant software beginning to include utilities for the formation of frequency distributions.

[17] In a recent book review, two immunologists have commented that in biology the real contribution of mathematics is not that it introduces quantification, but that it provides '... a precise qualitative framework of reasoning. As biological knowledge becomes ever more complex and detailed, so natural language becomes more inadequate for certain types of biological questions. Mathematics provides an efficient, precise, and rigorous alternative; as the authors note, "mathematics is no more, but no less, than a way of thinking clearly"...' (Bangham and Aquith 2001).

Giere's and Cartwright's model-based views of science, by dethroning the concept of scientific law – in physics of all disciplines, the one considered most lawlike – have underlined the abstract, unreal character of all scientific theory and theoretical models. In a sense, all scientific theory is hypothetical. But the hypothesis is not whether it will eventually be proved true or valid, but whether it fits a particular part of the real world well enough to accomplish a particular human purpose. The application of this thinking to demography appears to me to abolish the sharp distinction between formal and behavioral demography, as traditionally defined. The propositions of a behavioral theory or model must be formally true, true in the sense that definitions are true. Otherwise, the model cannot serve as what Cartwright terms a *nomological machine*; it cannot generate implications or predictions that follow in strict logic from the model assumptions and structure. The idea of necessary relationships in our mathematical models in demography and contingent relationships in our theoretical models needs re-examination.

A final distinction that is breaking down is that between theory and modeling. Older texts on mathematical modeling and on social theory (as always, economics is an exception) have tended to occupy different worlds. Mathematical modeling books seldom have much to say about theory, and vice-versa.[18] The model-based view of science, on the other hand, sees the model as the primary representational device in science, encapsulating its theoretical ideas. The theory/model distinction is blurred. And, given the complexity of many social and demographic systems, the theoretical tool of choice becomes computer modeling or simulation. A particularly promising genre of simulation is agent-based modelling, which promises to link individual demographic behaviors to aggregate patterns, and to explicate the social – as distinct from stochastic – mechanisms underlying demographic dynamics.

This does not mean that empirical, statistical studies will become unimportant. Indeed, their importance may be all the greater, given a larger and stronger set of theoretical tools with which to interpret their findings. Nor does it mean that simpler models are obsolete. For some purposes, a simple model may be the model of choice. The ideal theoretical toolkit will contain a wide array of models – some old, some new; some mathematical, some verbal; some simple, some complex; some based directly on data, some more speculative; some explanatory, some more lawlike and predictive.[19] This inclusive approach to theory and models can only enhance the status of demography as science.

Old words in a language often lead to striking insights. In English, there is an old word, *theoric* (also spelled *theorick*), described as obsolete and archaic by the *Oxford English Dictionary*. Its first meaning is 'theory.' Its third meaning is: 'a

[18] An interesting exception is Doucet and Sloep (1992), whose book on mathematical modeling in the life sciences, includes a substantial discussion of the semantic school of philosophy, of which Giere and Cartwright are contemporary representatives.

[19] See Cartwright (1983) on two kinds of theory in physics. Newton's 'law' of universal gravitation describes the attraction of two bodies [ignoring other bodies], but does not give the mechanism underlying the attraction, nor explain the nature of gravitation. But few physicists would suggest it is not good theory.

mechanical device theoretically representing or explaining a natural phenomenon.' Sixteenth century examples relate to astronomy. To many social scientists, the separation between theory and a mechanical device is total. In Giere's model-based view of science, there are three kinds of models: theoretical; visual; and physical, that is, mechanical devices. Things have come full circle, back approximately to where they belong.

References

Abbott, A. (1988). Transcending general linear reality. *Sociological Theory, 6*, 169–186.

Bangham, C. R. M., & Aquith, B. (2001). Viral immunology from math. In M. A. Nowak & R. M. May (Eds.), *Review of virus dynamics: Mathematical principles of immunology and virology* (Vol. 291). Oxford University Press, 2000, in *Science* [9 February 2001], p. 992.

Billari, F. C. (2003). In A. Prskawetz (Ed.), *Agent-based computational demography; using simulation to improve our understanding of demographic behaviour*. Heidelberg: Physica-Verlag.

Bracher, M., Santow, G., Morgan, S. P., & Trussell, J. (1993). Marriage dissolution in Australia: Models and explanations. *Population Studies, 47*, 403–425.

Burch, T. K. (2002). Computer modeling of theory: Explanation for the 21st century. In R. Franck (Ed.), *The explanatory power of models: Bridging the gap between empirical and theoretical research in the social sciences*. Dordrecht: Kluwer Academic Press. See also Ch.4 in this volume.

Burch, T. K., & Belanger, D. (1999). L'etude des unions en demographie: des categories aux processus. *Cahiers Quebecois de Démographie, 28*, 23–52.

Cartwright, N. D. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.

Cartwright, N. D. (1999). *The dappled world: A study of the boundaries of science*. New York: Cambridge University Press.

Casti, J. L. (1997). *World-be worlds: How simulation is changing the frontiers of science*. New York: Wiley.

Coale, A. J. (1956). The effects of changes in mortality and fertility on age composition. *The Milbank Memorial Fund Quarterly, 34*, 79–114.

Coale, A. J. (1965). Factors associated with the development of low fertility: An historic summary. *United Nations world population conference, WPCAVP/194*, Belgrade.

Coale, A. J. (1973). The demographic transition. International Union for the Scientific Study of Population. *International Population Conference* (pp. 53–72).

Coale, A. J., & Hoover, E. M. (1958). *Population growth and economic development in low-income countries*. Princeton: Princeton University Press.

Coale, A. J., & Watkins, S. C. (1986). *The decline of fertility in Europe*. Princeton: Princeton University Press.

Courgeau, D. (2004). *Du groupe à L'Individu: Synthèse Multiniveau*. Paris: L'Institut National D'Ètude Démographiques.

Doucet, P., & Sloep, P. B. (1992). *Mathematical modeling in the life sciences*. New York: Ellis Horwood.

Edwards, D., & Hamson, M. (1989). *Guide to mathematical modelling*. Boca Raton: CRC Press.

Feeney, G. (1994). Fertility decline in East Asia. *Science, 266*, 1518–1523.

Friedman, M. (1953). The methodology of positive economics. In *Essays in positive economics* (pp. 3–43). Chicago: University of Chicago Press.

Giere, R. N. (1999). *Science without laws*. Chicago: University of Chicago Press.

Hedström, P., & Swedberg, R. (1998). *Social mechanisms: An analytic approach to social theory*. Cambridge: Cambridge University Press.
Hernes, G. (1972). The process of entry into first marriage. *American Sociological Review, 37*, 173–182.
Hobcraft, J. (2000). *Moving beyond elaborate description: Towards understanding choices about parenthood*. Paper presented at UN Economic Commission for Europe, Fertility and Family Surveys Flagship Conference, Brussels, Belgium.
Jandel Scientific. (1996). *TableCurve 2D: Automated curve fitting and equation discovery*. San Rafael: Jandel Scientific.
Keyfitz, N. (1971). Models. *Demography, 8*, 571–580.
Keyfitz, N. (1975). How do we know the facts of demography? *Population and Development Review, 1*, 267–288.
Lesthaeghe, R., & Vanderhoeft, C. (1997). *Ready, willing and able: A conceptualization of transitions to new behavioral forms*. Vrije Universiteit, IPD Working Paper, Interface Demography, Brussels.
McNicoll, G. (1992). The agenda of population studies: A commentary and complaint. *Population and Development Review, 18*, 399–420.
Meehan, E. J. (1968). *Explanation in social science: A system paradigm*. Homewood: The Dorsey Press.
Platt, J. R. (1964). Strong inference in scientific research. *Science, 146*, 347–353.
Turner, S. P. (1987). Underdetermination and the promise of statistical sociology. *Sociological Theory, 5*, 172–184.
Walker, A. (1997). *Modelmaker: User manual. Version 3*. Oxford: Cherwell Scientific Publishing.