

Chapter 1

Demography in a New Key: A Theory of Population Theory

1.1 Introduction

The status of theory in demography has been problematic ever since I can remember. Sixty-five years ago, Rupert Vance, in his Population Association of American presidential address, asked ‘Is theory for demographers?’ (1952). There is ample evidence that many demographers – then and now – would answer ‘Of course, but it’s not a high priority.’ But if demography is a true science – as opposed to a body of techniques or a branch of applied statistics – it must have theory, recognize that it has theory, codify its theory, and seriously teach theory to its students.¹

In his presidential Address to the Population Association of America, Nathan Keyfitz (1971) adopted what he termed a ‘liberal view of models.’ In this chapter, I sketch a liberal view of scientific theory, and discuss some of its implications for the way we think about demography and the way we present it to others.

This view of theory is known in philosophy of science circles as the ‘semantic’ view, or more recently and descriptively, the ‘model-based’ view of science. In describing this approach, I draw heavily on the work of Ronald Giere, an American philosopher of science (1988, 1999), but also on some methodological writings of Nathan Keyfitz (1971, 1975). Keyfitz introduced these ideas to demography years ago, although they never became mainstream.²

Based on a presentation to a Symposium on Theory in Demography, part of celebration of the new building of the Max Planck Institute for Demographic Research, 31 March–1 April, 2003, Rostock, Germany; originally published in *Demographic Research* 9(2003):263–284.

¹For a while, the cover of *Demography* (official journal of the Population Association of America) defined the field as ‘the statistical study of human populations,’ seeming to imply that demography is a branch of statistics, not a science in its own right.

²I also have benefited greatly from the following: Meehan (1968), Newton (1997), and Cartwright (1983, 1999). For a summary and assessment of the semantic school, see Teller (2001). I am grateful to John Wilmoth for reminding me that Keyfitz had written several papers on the role of models in demography.

In the model-based view, models, not empirical laws, are the central element of scientific knowledge. A model is any abstract representation of some portion of the real world. A model may contain basic principles generally regarded as ‘laws.’ In this case, the laws ‘function as true statements, but not as statements about the world. They are then truths only of an abstract model.’³ In this context, such statements are true in the way that explicit definitions are true’ (Giere 1999, p. 6). A model contains generalizations, but they are formal generalizations, not empirical ones. Empirical assessment of theory, therefore, relates not to whether a theoretical model is empirically true or false – strictly speaking all theories and models are false because they are incomplete and simplified representations of reality – but ‘how well the resulting model fits the intended aspects of the real world’ (Giere 1999, p. 6). This view stands opposed to many familiar teachings of logical empiricism, by which theory is based on empirical laws, and judged true or false solely by its agreement with data. The model-based view is equally concerned with empirical data, but these are used to judge whether a model fits some portion of the world closely enough for a given purpose, not whether the model is true or false in any absolute sense.

The model-based approach has two general implications for our view of demography

1. Much of formal demography (techniques, methods) can be viewed also as theory, that is, as a collection of substantive models about how populations and cohorts behave;
2. Many theories in behavioral demography which have been rejected because of empirical exceptions or on the grounds they are too simplistic can be viewed as perfectly good theory, especially if they were to be stated more rigorously.

Indeed, at the theoretical level, the classic distinction between formal/technical and substantive/behavioral demography loses much of its force. In both sub-areas of demography, theoretical models have essentially the same epistemological standing, even if they may differ on other dimensions such as scope and complexity, and even if different kinds of day-to-day work may be involved in their development and use.

The word *theory* is ambiguous in the non-pejorative sense of ‘having two or more meanings.’ It means different things to different people, both in everyday speech and in scientific discourse. It is futile to try to establish the ‘correct’ definition or the ‘true meaning’ of theory. But it is possible to suggest a new – though not entirely new – approach to theory that might prove more fruitful than older ideas to which we are accustomed. In the next section, I summarize the main elements of the model-based view, noting some ways in which it differs from, but also agrees with, logical empiricism. A key part of this exposition is a partial re-definition of such terms as *model* and *theory*. But terminology is not crucial,

³Cartwright refers to theoretical models as ‘nomological machines,’ that is, models generate laws, not the other way around. See (1999, p. 4).

and some may want to define these words differently, and to preserve a sharp distinction between *theory* and *model*. The central ideas I wish to convey are an emphasis on formal demography as substantive knowledge, and a plea that empirical exceptions to otherwise useful behavioral theories should not lead to their discard.

In the logical empiricist view of science, theory comes from data through a process of induction and generalization. Theoretical knowledge and empirical knowledge occupy different but parallel planes, layered upward into ever more general and abstract propositions. In the model-based view, theory and empirical studies occupy non-parallel planes. The planes must intersect, of course, since we are discussing empirical science. But the origin and character of the two kinds of knowledge are qualitatively different. In the model-based view of science, as the name suggests, models, not laws, are the central element of scientific knowledge. The prototype of scientific knowledge is not the empirical or theoretical law, but a model plus a list of real-world systems to which it applies. To quote Giere:

In this picture of science, the primary representational relationship is between individual models and particular real systems, e.g., between a Newtonian model of a two-body gravitational system and the Earth-Moon system. . . Here we have not a universal law, but the restricted generalization that various pairs of objects in the solar system may be represented by a Newtonian two-body gravitational model of a specified type. (Giere 1999, p. 93)

A model is any abstract representation of part of the real world, constructed to understand, explain, predict, or control. Giere distinguishes three types of models:

1. Physical models (for example, an automobile in a wind tunnel);
2. Visual models (for example, maps showing plate tectonics, or a diagram of the demographic transition);
3. Theoretical models (for example, Newton's Law of falling bodies, or the theory of evolution).

Physical models have little relevance to demography and other social sciences. Visual models have great potential, but are not as widely used as they might be, with the bulk of graphics in demography limited to the representation of data frequency distributions, time series, and age-structures rather than processes or systems.

Theoretical models can be expressed in ordinary language, formal logical systems, mathematics, computer code or diagrams.⁴

In the model-based view, no sharp distinction is made between model and theory. A collection of small models relating to the same realm can be called *theory* (for example, the theory of harmonic oscillators, or the theory of population aging).

⁴The idea that theory consists of purely verbal statements seems peculiar to social science. In the physical sciences, many of the most important theories are in mathematical form – Newton's law of gravity, Relativity, etc. For a recent indication of this way of thinking, see, for example, Baylis (1994). His book on *Theoretical Methods in the Physical Sciences* is an introduction to the use of a computer mathematics program, Maple V, to solve substantive problems in elementary physics.

These models typically contain a small number of variables, and are constructed to represent very limited portions of the real world. Or, *theory* can refer to a system of very general ideas (for example, the theory of relativity, or transition theory) attempting to represent larger, more complex real world systems. The difference is not qualitative, but relates to differences in scope, complexity, and other quantitative dimensions.⁵ There may be advantages to preserving fine distinctions among the words *theory*, *model*, and *theoretical model* in some contexts. In this chapter, they are used interchangeably.

Giere draws a useful analogy between scientific models and maps, viewed as simplified representations of our physical surroundings (1999, pp. 25–26; 81–82; 214–215). Like theoretical models, maps vary in purpose and scope. Some maps give a broad overview of nations or of whole states or provinces, and more detailed maps, often as insets, of smaller areas such as cities or metropolitan areas. Some maps are extremely simple. An example is the straight-line map found on metro trains or subways, which show only the stops and transfer points, which is all the rider needs to know. A topological map is useful for backpacking but no substitute for a road map. Maps differ in scope and detail, but all are abstract representations of reality.

A theory or theoretical model is a formal system: a set of propositions involving objects, variables, and relations among them. It must be clear and logically consistent. A model is constructed to represent or explain some empirical reality. But it need not be derived from empirical generalizations. And it does not have to be – indeed it cannot be – empirically true. In Giere’s words, models are true ‘in the way explicit definitions are true’ (1999, p. 6). They can never be absolutely and literally true because they are always partial and approximate representations of an infinitely complex real world. Scientific theories, he notes elsewhere (1988, p. xvi) can be viewed ‘not as empirical statements but as definitions of models variously related to the real world.’ And so, ‘Science does not deliver to us universal truths underlying all natural phenomena; but it does provide models of reality possessing various degrees of scope and accuracy’ (1999, p. 6). One can have, says Giere, ‘realism without truth.’

Keyfitz, discussing models of the demographic effects of eliminating deaths from heart diseases, comments similarly that his conclusions ‘...are conditional statements, and as such they are true beyond debate, given their assumptions that death rates by age from all other causes and birth rates by age of mother will remain as they are’ (1971, p. 574). Conclusions drawn from a model follow inexorably from assumptions and model structure. Later, he contrasts the firmness of these conclusions with those established by ‘direct observation, which tend to provide enigmatic and inconsistent reports’ (1975, p. 267).

⁵Some authors distinguish theory and model, assigning the latter a role as intermediary between theory and empirical data. See for example, Gould and Tobochnik (1996) and Skvoretz (1998). Their distinction is on a general/specific axis and is not fundamental.

How then does one evaluate a model or theory? A model is a good model – Giere would not say a ‘true’ model – if it fits some portion of the real world (1) closely enough, (2) in certain respects, (3) for a specific purpose. All models are approximations. The question is whether the approximation is good enough for the purpose at hand. All models have a limited number of variables; none can mirror the numberless qualities of the real world. And finally, any model is to be evaluated with reference to the purpose for which it was designed or constructed.

The map analogy cited earlier helps clarify the last point. A highway map and a topographic map can both represent the same area. But the highway map is relatively useless for back-country hiking. It is not an incorrect or false representation, just the wrong one for the purpose. Similarly, a metro map correctly tells a rider where to get on and off the train, but is practically useless when one emerges above ground. A map of city streets is needed.

Over time in any science, some models receive widespread acceptance because they seem to embody central principles, or because they are widely applicable. In physics, classical mechanics provides an example. These models are taught in every introductory physics course. It is well understood that such models do not work as well at the sub-atomic level, or on the scale of the universe. But they are not therefore abandoned, since they serve many purposes in our everyday world.

The fit of a model to the real world is a matter for empirical examination. It is this empirical research that links model and data. But the conclusion that a model does not fit a particular case – perhaps not even closely – is only a conclusion that the model does not fit, not that the model is inherently false or invalid. It may well fit other cases. Decisions about whether or how well models fit the real world are based on scientific judgement, not on purely logical criteria. Giere again: ‘judging the fit of a model to the world is matter of decision, not logical inference’ (1999, p. 7).

The model-based view of theory has developed in opposition to logical empiricism, the dominant philosophy of empirical social science during the second half of the twentieth century. It differs from logical positivism in that the elements of a model do not have to be or be derived from or be logically consistent with broad – some would say universal – empirical generalizations or ‘laws.’ Such generalizations as exist may be incorporated into a model, but they are not essential. Many proponents of the model-based approach conclude that the logical empiricist program has been self-defeating precisely because empirical generalizations in social science are relatively rare.⁶ The model-based view agrees with logical empiricism in its emphasis on the importance of empirical observation. It is the

⁶Critics of social science often take the absence of universal empirical generalizations as evidence that the social sciences are not really science. See, for example, *The Economist* (8 May, 1999, p. 84): ‘...unlike physics, economics yields no natural laws or universal constants. That is what makes decisive falsification in economics so difficult. And that is why...economics is not and never can be a proper science.’ This statement reflects a common misunderstanding of physics. The constant in Newton’s law of gravity is only ‘relatively constant.’ Depending as it does on mass, it is much different on the moon, and even differs across earth, at different altitudes and locations.

real world, insofar as it can be observed, that one is trying to understand and explain, not some imaginary world, a pure construct. The imagination is at play in theorizing and model building. But it begins with some empirical observation to be accounted for, and it returns to empirical observation to see if the account is a good or useful one. Otherwise, there is endless speculation.

The model-based view differs from a common view of economic theory, in which theory is derived from a limited set of axioms such as ‘impersonal markets,’ ‘maximizing behavior,’ and ‘well-ordered preferences.’ In the model-based view, the canonical axioms of economics may be incorporated into a model, but they need not be. Model construction is less constrained than in logical empiricism or mainstream economics. It is a creative leap from some empirical phenomenon that needs to be understood or explained, to the construction of a model that seems to do the job. Whether or how well it does so, as noted above, is a matter for empirical examination and scientific judgement.

The model-based view agrees with economics in an emphasis on the need for rigor in the statement of theories. The empirical assessment or use of models depends on their capacity to yield definite implications or predictions, and to support truly logical explanations.

In the model-based view, however, theory is not deductive in the sense of being inferred from a limited set of axioms. But explanation using a model is deductive, in the sense that the event or outcome to be explained must follow logically from the model, must be deducible from it. Nor is theory inductive in the sense of being derived from an examination of many cases to arrive at broad empirical generalizations. It is inductive in the broader sense that it starts with empirical observation and arrives at an abstract, and therefore general, model. But the process involves a creative leap of the imagination, not just generalization of the facts.⁷

A model or theory need not deal with general classes of phenomena. Otherwise, there could be no theory of the evolution of the human species or of the origins of the universe, both unique events. It is one of the strengths of the model-based view of science that it directs us to use abstract models to study unique events, unlike logical empiricism which requires empirical generalizations about classes of events. In the latter system, to the extent an event is truly unique, it cannot be subsumed under a class or a class-based generalization, and therefore resists explanation. The model-based approach to unique events enables us to pursue theoretical explanations, rather than falling back on the pure descriptions of ethnography or narrative history.

In this liberal view of theory, there are many different kinds. There are simple theories or theoretical models, and complex ones; cross-sectional and longitudinal or dynamic; and, as noted above, theories which apply to classes of phenomena, and those which apply to unique events. Clearly the latter kind of theory cannot be based on empirical generalizations based on the study of many cases; there is only one case. The generality lies in the model itself, not in data. In Meehan’s words, ‘...

⁷See Franck (2002), especially his introductory and concluding remarks, on ‘classical induction.’

timeless or general propositions are assumed to belong to the logical rather than the empirical world' (1968, p. 32).

To apply to a concrete phenomenon, of course, a theoretical model must be given greater specification. But even with such specification, it remains a theoretical model. The term *theoretical model*, for Giere, refers 'either to a general model or to one of its specific versions obtained by specifying unique values for all parameters and initial conditions' (1999, p. 177).

In physics, a distinction is sometimes made between *phenomenological* and *fundamental* theories (see Cartwright 1983). The former is essentially a description of what happens and how, without too much concern for why. A classic example is Newton's principle of gravity, which tells us that bodies released from a height will fall, and approximately how fast they will accelerate, but does not tell us what gravity is (Ekeland 1988). Fundamental theory delves more deeply into causes and mechanisms.

Meehan (1968) makes a similar distinction between models which can only predict a phenomenon, and those that can also explain why it will occur, by explicating processes or mechanisms. He views the latter as more difficult to construct, but also as more powerful, insofar as they make it possible to control events – at least in principle – not just adjust to them.

Other things equal, fundamental or explanatory models are of greater scientific value, because they involve deeper knowledge and understanding, and have more varied applications. But model assessment is related to purpose. And for some purposes, a phenomenological model may be just as effective and, often as not, easier to use.

1.2 Some Demographic Models Revisited

The model-based approach to science leads to a new perspective on demography; demographic knowledge, old and new, is seen from a different angle. Or, to use a musical metaphor, the same old demographic songs can be sung in a new key and reharmonized. This approach, I believe, greatly enhances demography's status as a science, notably its status as an autonomous discipline with its own large body of good theory. A few examples will illustrate the point.

The Exponential Growth Model No one would question the validity or 'truth' of the expression $P[t] = P[0]e^{rt}$. It is a standard mathematical function. In demography, the empirical question cannot be whether it is true, but only whether it applies to concrete human populations. And this depends on purpose. It is a good theoretical model to describe the basic character of the growth of the human and many other biological species, namely, that growth is proportional to population size. It is a good model to calculate an annual average rate of growth over some historical period, although like many averages the resulting figure may be misleading. The exponential model, however, is not a good model to describe the actual growth

trajectory of many, perhaps most, real-world populations; consider the many examples of supra-exponential growth during the last three centuries. But we do not therefore say that the exponential model has been falsified, only that it doesn't fit the cases at hand.

Is the exponential model demographic theory? Perhaps some would prefer to call it a theoretical model or just a model. But when taken together with several others – the logistic, a supra-exponential model, the stable model, the cohort-component projection model – one can legitimately speak of the resulting collection as a 'theory of population growth.'

The Life Table The life table usually is presented as a complex measurement device, primarily a measure of current mortality. But it is basically a model of cohort survival. The algorithm for calculating a life table from assumed death rates or probabilities is true, depending as it does on the straightforward application of basic arithmetical operations. And, a life-table based on observed rates is a true summary of those rates. Again, the relevant question is not whether a specific life table is true or false, but whether it fits a real-world population closely enough for the purpose at hand. To summarize current age-specific death rates and re-work the information they contain into a more useful form (for example, for calculating e_0 or survival ratios), the life table works quite well. Whether the input rates somehow misrepresent some true, underlying mortality level is another issue, as is the question of whether a current life table can be used to forecast future mortality. The very best life table for contemporary humans would do a poor job of characterising the survival patterns of early humanoids or of other species, say insects.⁸ But we can only say that it does not apply in these cases, not that it is invalid or false. Incidentally, I would describe the life table and other objects from formal demography as being *behavioral*, in the sense they characterize the survival behavior of a cohort – aggregate behavior to be sure, but behavior nonetheless.

The Coale-McNeil Marriage Model The Coale-McNeil model of first marriage (Coale and McNeil 1972; Coale 1977) began life as an exercise in mathematical curve fitting. Only later was it interpreted in terms of waiting times for entry into various stages of the marriage process. I once criticized the model compared to that of Hernes (see Burch 1993, and Chap. 6 below) because it was lacking in behavioral theory; I characterized the waiting-times interpretation as 'semi-behavioral.' This earlier assessment reflected my logical empiricist training and heritage, and acceptance of the conventional distinction between formal and behavioral demography. I would now say that to find a parametric model that closely fits a large collection of age patterns of first marriage is a considerable theoretical achievement – in the category of phenomenological theory. It is behavioral in the sense mentioned above – it captures important features of cohort behavior. It is not, of course, the only good model of marriage. There are several others, some of which may be better for

⁸The consideration of survival curves for other species or of unrealistic curves for humans (e.g., calculating a life table with a typical age pattern of q 's reversed) helps put human survival and its implications for age structure and social structure in better perspective. See Carey (2002).

certain purposes. And there is both need and opportunity to develop rigorous models of marriage that are more complex and more richly behavioral. Some recent agent-based models of marriage (Todd and Billari 2003; Billari and Prskawetz 2003; Billari *et al.* 2003) represent one promising direction for these efforts.

Coale and Trussell's later (1996) discussion of the character of parametric models is instructive:

The models are descriptive and were never intended to be anything else. No deep theory, or even shallow theory, underlies the search for empirical regularities. (p. 483)

The quality of a model, in their view, 'depends on how usefully it can be exploited for empirical research' (p. 469). Three uses are highlighted: testing data; building blocks for estimates; forecasting. The value of models is closely tied to working with 'inaccurate and incomplete data' (p. 484). Later they lament 'the virtual absence of the development and steep decline in the use of demographic models during the past decade,' related in part to increasing availability of good survey data (p. 484). But if the value of demographic models is tied primarily to the absence of good data, then there is some logic in a decline in their use as data improve. If they are viewed instead as substantive models of demographic behavior, then they have permanent value and application. This is recognized implicitly when they note that models 'can be used to make broad inferences about behavior...'; but the emphasis is elsewhere, since they continue: '...or, more commonly, to build techniques for estimating basic demographic indices for populations with limited or defective data' (p. 484).

Two Kinds of Transition Theory One of the problems with the theory of demographic transition is that we have never quite agreed on precisely what it is (McNicoll 1992). In keeping with what has gone before, I would suggest that there are two kinds of transition theory. The phenomenological version simply states that a large, sustained decline in mortality will be followed, after a time lag, by a sustained decline in fertility, resulting in an intervening period of rapid population growth. A more fundamental version would include the determinants of mortality and fertility decline – modernisation, economic growth, secularisation, individualism, technological developments in medicine and fertility control, and so forth.

Either version of transition theory can be stated as an abstract model. In the former case, the model would assume a population in dynamic equilibrium, with constant mortality and fertility (and no migration). An assumed pattern of mortality decline is followed, after a delay, by fertility decline. There is rapid growth in the intervening period, and slower growth when a new equilibrium is established. Such a model is true by construction, 'true in the way that a definition is true.' Empirically it can be used to characterize the modern demographic history of many –

though certainly not all – human populations. For others, for example France or Hungary, a different model is needed.⁹

A more behavioral version of transition theory can also be stated as an abstract model. Mortality decline is defined as a function of development, with subsequent fertility decline a function both of mortality decline and of development. The approach would be similar to population biology's definition of the logistic model, in which mortality and fertility are functions of population density. The key difference is that in a transition model, mortality declines with development and population growth, whereas in the logistic model it rises (in both, fertility declines). The link between mortality decline and fertility decline might be explained in terms of pressure at the individual, family, or community level, because of larger numbers of surviving children. Again, such a transition model would provide an approximate but accurate description of the demographic history of many nations, along with a behavioral explanation for that history. With closer specification and real data inputs, it could provide a better approximation of the history of a particular nation. Probably no one specification could provide a close fit to the history of all nations, since this history did not occur 'in a vacuum' or in controlled experimental conditions. It was this historical fact that led Coale to conclude that the only generalization to emerge from the vast historical studies of European fertility decline was that fertility would decline when a population was, to borrow Lesthaeghe's paraphrase, 'ready, willing, and able' (Lesthaeghe and Vanderhoeft 1997). The postulated mechanism linking mortality and fertility decline, of course, does not flow from an empirical generalization, which is precisely why it is a theoretical explanation.

It would be easy to multiply examples, drawn from the demographic literature both old and new. But these suffice to show that, whether in the realm of formal demography or of behavioral demography – as traditionally conceived – we can view our models as formal models, models that are true in the way that definitions are true. In either sub-field, our models are abstract representations of the real world, inspired by empirical observation; epistemologically, they are of one piece. The models of formal demography are not just measurement techniques. They have a theoretical character.¹⁰ The models of behavioral demography need not be rejected because they do not fit all the facts, so long as they fit some relevant

⁹The literature contains a few examples of formalization of transition theory (see, for example, Keyfitz 1985, pp. 23ff). But none has become standard or widely used or cited.

¹⁰The idea that some demographic measurement techniques are models is not novel. Newell (1988) for example has a chapter entitled 'Introduction to Demographic Models.' He distinguishes normative and descriptive models, and mentions the total fertility rate and the life table as examples of the former. He comments: 'These normative models so dominate formal demography that it is not often they are actually thought of as models; yet it should always be remembered that a move from ASFR's to a TFR, or from ASDR's to a life table, is a move from reality to a model' (p. 118). Newell does not take the further step, advocated here, of viewing such a model or a collection of related models as theory.

cases well enough to be useful for one or another purpose. In short, demography has more good theory than commonly recognized.

1.3 Demography Reconsidered

The model-based view of demography has many further implications for the way we think of the field, and the way we present it to others, notably our students. I highlight five, some of them recapitulated from above.

1. We need to become more comfortable with the idea of several different models for the same phenomenon. Logical empiricism pushed toward the view that in empirical tests, one model would emerge as a winner, with the others being falsified and rejected. I have always suspected that this idea reflects a deep, even subconscious, monotheistic belief. Early scientific thinking often was explicitly theological (note recent publicity about Newton's theological speculations). God created the universe, implanting in it certain laws. Science's job was to find them. And since there is only one God, laws of nature will be unique. This led to what Teller (2001) has called 'the perfect model model' of science.

The model-based view prefers to think of a pantheon, or to change the metaphor, a toolkit of related models, with different characteristics and serving different purposes. With respect to population growth, for example, one can point to: the exponential model; the logistic model; transition theory; the stable model; the standard projection model. Which is the true model of population growth? The question makes no sense. With respect to fertility, similarly, one can point to: Becker's microeconomic model; the Easterlin-Crimmins socio-economic model; Friedman et al. (1994) uncertainty model; the social capillarity model; Davis's multiphasic model; Coale's model of the three preconditions; transition theory; Lesthaeghe on secular individualism; the newer models on 'social influence' and diffusion. Which is the true model of fertility? If we retain the notion of truth at all, then surely the most that can be said is that each model incorporates some element of truth. None is complete nonsense, such as the idea that fertility decline has been imposed on us by Martians to prevent our depletion of the world's resources before they can get to them.

This toolkit approach to scientific theories does not imply that all models are equal. Some models be better approximations of a wider variety of cases, or useful for a wider variety of purposes. Such models will naturally tend to be used more often. But the lesser models also will be used on occasion. It is not prudent to discard them.

2. As noted earlier, at the theoretical level the sharp traditional distinction between formal and behavioral demography is discarded. All theories or theoretical models become formal in the sense outlined above. In Lotka's phrase, all theory is 'analytic theory' (1939). The body of work we generally regard as

demographic techniques or methods can still be thought of as techniques. But much of it also can be thought of as theoretical models of population dynamics – substantive models of how populations or cohorts behave, often under idealized conditions. Past practice in this regard is inconsistent. The stable model is commonly referred to as ‘stable theory.’ But the cohort-component projection model is classified as a technique, and many demographers would object to its classification as theory. Yet both models represent the development of population size and structure in the face of assumed inputs. It is hard to see why one is theory and one is not – unless one can argue for a valid distinction based on the level of mathematics involved.

Reinterpreted models from formal demography are behavioral in a limited sense of dealing with the behavior of aggregates, without explicit reference to motivation, values, norms, and decision making (see McNicoll’s 1992 reference to ‘the limitless depths’ of human behavior). But they are behavioral nonetheless, in the same sense that Newton’s law speaks to the behavior of falling bodies.

But surely, it may be objected, the classic distinction between necessary and contingent relationships (see Lotka’s distinction between ‘analytic’ and ‘statistical’ demography) is valid. I reply with a distinction. There is contingency in our empirical observations. But we construct theoretical models in a way that contingency is left behind. The statement ‘natural increase equals births minus deaths’ states a necessary relationship; the empirical statement ‘high rates of female labour-force participation are associated with very low fertility’ is contingent, and not universally true. But the assumption that it is true can be incorporated in a model. It is then true by construction, true ‘in the way that a definition is true.’ This is essentially the message of Keyfitz’s (1975) paper on ‘How do we know the facts of demography?’

In the teaching of demography, I have come to see the formal-behavioral distinction as arbitrary. Consider a lecture on the determinants of population growth, based on a series of models. Successively we show that growth depends on numbers of births, deaths, and migrants (in and out). We then show that the number of births depends on a rate or set of rates interacting with population size and age-structure. A student asks: ‘But what determines the rates?’ The conventional answer might be: ‘This course only deals with formal demography and techniques. To consider that question you must take another course.’¹¹ A model-based approach would simply move on to the next set of theoretical models, those dealing with the determinants of fertility, for example, Easterlin-Crimmins or Coale. I have never encountered a physics or chemistry text that made such a sharp distinction between formal and behavioral physics or chemistry, or between substance and method.

¹¹Another manifestation of this distinction is seen in the common practice in introductory textbooks of relegating technical demography to appendices, and in demographic methods texts of omitting all but cursory discussions of behavioral models and theory.

3. Many computer simulations in demography may be viewed as theory or as tools for theoretical analysis. For many social and behavioral theorists, manipulation of a numerical model with a computer is ‘number crunching,’ not theory. For many empiricists, it is theory in the bad sense of armchair speculation, yielding numbers that are made up. The model-based view would say rather that a simulation is an abstract model of a real demographic system, and can be manipulated to yield insight into how that system works, or applied to real world systems to explain or predict. There is no inherent difference between this use of a simulation model and the classic uses of stable population theory to clarify dynamic interrelations among fertility, mortality, population growth, and age structure. An interesting question is why demographers have generally been more receptive to stable analyses than computer simulations.

This view of simulation is not unknown in demography. It has been expounded and illustrated over the years by researchers such as Hammel and Wachter, to give but one example. In fact, a computer model is just another kind of model, written in a new kind of language. And the model-based view of science does not differ greatly from the mainstream tradition of mathematical modelling, in which a model is constructed for a specific purpose and its performance judged explicitly with regard to that purpose. Mathematical modellers have generally not viewed their work as *theory*, however, although in many contexts, it can be so viewed.

In his ‘liberal view of models,’ Keyfitz noted, ‘...they may be algebraic, arithmetical, computer simulation, or verbal’ (1971, p. 575). His example of a verbal model is ‘demographic transition’ theory. And as his later paper (1975) strongly suggests, there is no essential difference between a model and a theory.

4. To qualify as theoretical models of the kind I am advocating, many of our ‘behavioral’ models must be defined more clearly and rigorously (Burch 1996). This is necessary for discovering logical implications of the models. It is not enough to work with highly discursive models which ‘give us a feel for what’s going on.’ Such models, as is well known, can be used to explain or predict almost anything, and therefore explain or predict nothing.

The need to derive definite implications, incidentally, is why the ‘probabilistic finesse’ – the reliance on probable generalizations rather than universal ones – is not able to save the logical empiricist approach to theory. A chain or other combination of several probabilistic empirical statements yields implications of at best low probability (by the multiplicative rule). In the model-based approach, within the model itself the inferences are certain, allowing only for some stochastic elements. Uncertainty comes when the inferences are applied to the real world, since the fit is never perfect. But this is a matter of scientific judgement, not just logic. To quote Giere, ‘It is enough that the premises confer some appropriate degree of ‘probability’ or ‘rational warrant’ to the conclusion’ (1988, p. 11).

5. The model-based view is comfortable in dealing with unique phenomena. In the logical empiricist model of science, theory is based on generalization across many cases of a phenomenon considered as a class, e.g., national fertility transitions. Explanation of a particular case is achieved by ‘subsuming’ it under some general theoretical propositions about the class.

There are at least two problems with this approach. First, there may be no class. Some phenomena are unique in the literal sense of the word – there is only one case. The origin of the universe and the evolution of the human species are examples. Secondly, even if there appears to be a class, often it will have been defined for extra-scientific purposes (e.g., national and provincial political boundaries). Theoretical considerations do not guide the definition. But in this case, there is no reason to assume that the classes are homogeneous with respect to characteristics of interest. To try to find a general model of such a class may involve trying to represent systems that differ in ways that are centrally important.

The model-based view of science, by contrast, has no problem with constructing models to deal with unique events. It offers an alternative to giving up on theory in the face of the tremendous variability of real-world phenomena, which seems – perhaps rightly – to defy generalization. To some, especially historians and anthropologists, this variability means that one can only resort to detailed or ‘thick’ description. But as has been seen above, we can have theory without generalization.

1.4 Conclusion

But what difference does it make? Why should we take the trouble to change our way of looking at our discipline, and the way we present it to students, policy makers, and the public. Why bother to revamp our textbooks and our courses, as might be necessary were the model-based view to be taken seriously?

My first answer is that the model-based view should be taken seriously because it makes more sense than the views to which we are accustomed. It accords with an emerging mainstream in philosophy of science, and with what scientists in some of the most successful fields think about their work and teach their students. Without being inconsistent, of course, I cannot argue that this is *the true* approach to science, only that it is a liberating and fruitful one.

A second answer is that this liberal view of theory and of models enhances the stature of demography as a science, an autonomous, well-balanced scientific discipline, with a large body of good theory, as well as of techniques, data, and empirical findings.

Thirdly, other than some mental effort, there is little downside risk in doing demography in a new key. There need be no wholesale abandonment of what we currently think and do. We can still think of the life table as a measurement tool, while beginning to think of it also as a theoretical model of cohort survival. We still will spend a great deal of time and effort on the statistical analysis of census, vital

statistics, and survey data, providing descriptive studies of demographic trends. Techniques will be refined, and descriptive studies will continue apace, both detailed descriptions of individual cases and attempts to generalize where possible.

Collectively, more time and effort will be spent on the construction of new models and on the rigorous statement and systematisation of those we already have (theoretical synthesis). Not everyone is likely to become a theorist or model builder. But, one hopes, empirical research will be better informed by explicitly and rigorously crafted models, used to design the research, not just heuristically in the introductory sections of papers, or for ad hoc interpretation in the concluding sections.

The danger of an uncontrolled profusion of models seems unlikely in a discipline so closely wedded to empirical data. But the model-based approach itself guards against this danger, with its great emphasis on purpose: every model is built for some clear purpose, and is judged accordingly. One doesn't model for the sake of modelling. If a model doesn't yield insights into basic principles or fit some important empirical case – some of the data – then it may be abandoned, or better, reworked or refined.

What is at issue is a balance between empirical observation and theory, in a complete science. Demography has been exceptionally strong at empirical observation, and has one of the largest bodies of reliable data of any of the human sciences. But what does it mean? How can it be organized and presented to others? That is the role of theory and theoretical models.

Nancy Cartwright writes of theory: 'Explanations [at least the high-level explanations of theoretical science...] organize, briefly and efficiently, the unwieldy, and perhaps unlearnable, mass of highly detailed knowledge that we have of phenomena' (1983, p. 87). A large dose of the right kind of theoretical thinking could help us all digest the vast body of demographic information. Teller (2001) speaks of 'humanly accessible understanding.'

Cartwright continues in the above quote: 'But organizing power has nothing to do with the truth.' Truth, if only of a relative sort, lies in our facts (The largest national population in the world is that of China), and in our empirical generalizations such as they are (Low fertility in the modern world tends to be associated with high levels of socio-economic development).

With a better appreciation of demography's large fund of theoretical models, we can have the best of both worlds: truth in our empirical observations, and, in Giere's words, 'realism without truth' in our models. But models don't have to be true to be useful. Sufficient realism for the purpose at hand supports understanding, explanation, prediction, and policy guidance. And, to give Keyfitz the last word – 'no models, no understanding' (1975, p. 275).

References

- Baylis, W. E. (1994). *Theoretical methods in the physical sciences: An introduction to problem solving using maple V*. Boston: Birkhauser.
- Billari, F. C., & Prskawetz, A. (Eds.). (2003). *Agent-based computational demography: Using simulation to improve our understanding of demographic behavior*. Heidelberg: Physica-Verlag.
- Billari, F. C., Prskawetz, A., Fürnkranz, J. (2003). On the cultural evolution of age-at-marriage norms. In Billari & Prskawetz (pp. 139–158).
- Burch, T. K. (1993). Theory, computers, and the parameterisation of demographic behavior. *IUSSP International Population Conference, Montreal*. Liege: International Union for the Scientific Study of Population., 3, 377–388.
- Burch, T. K. (1996). Icons, strawmen and precision: reflections on demographic theories of fertility decline. *The Sociological Quarterly*, 37, 59–81.
- Carey, J. R. (2002). The importance of including biodemography in the demography curriculum. *Genus*. [Special Issue: *Teaching Demography in the Early 21st Century*, edited by G. Caselli], 58, 189–199.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.
- Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. New York: Cambridge University Press.
- Coale, A. J. (1977). The development of new models of nuptiality and fertility. *Population [numero special]*, 32, 131–152.
- Coale, A. J., & McNeil, D. J. (1972). The distribution by age of the frequency of first marriage in a female cohort. *Journal of the American Statistical Association*, 67, 743–749.
- Coale, A. J., & Trussell, J. (1996). The use of demographic models. *Population Studies*, 50, 469–484.
- Ekeland, I. (1988). *Mathematics and the unexpected*. Chicago: University of Chicago Press.
- Franck, R. (Ed.). (2002). *The explanatory power of models: Bridging the gap between empirical and theoretical research in the social sciences*. Boston: Kluwer Academic Publications.
- Friedman, D., Hechter, M., & Kanmazawa, S. (1994). A theory of the value of children. *Demography*, 31, 375–402.
- Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.
- Giere, R. N. (1999). *Science without laws*. Chicago: University of Chicago Press.
- Gould, H., & Tobochnik, J. (1996). *An introduction to computer simulation methods: Applications to physical systems*. Reading: Addison-Wesley Publishing.
- 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.
- Keyfitz, N. (1985). *Applied mathematical demography* (2nd ed.). New York: Spring-Verlag.
- Lesthaeghe, R., & Vanderhoeft, C. (1997). *Ready, willing and able: A conceptualization of transitions to new behavioral forms* (IPD Working Paper #8). Brussel: Vrije Universiteit.
- Lotka, A. J. (1939). *Theorie Analytique des Associations Biologiques, Deuxieme Partie*. Paris: Hermann & Cie.
- McNicol, 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.
- Newell, C. (1988). *Methods and models in demography*. New York: The Guilford Press.
- Newton, R. (1997). *The truth of science: Physical theories and reality*. Cambridge, MA: Harvard University Press.

- Skvoretz, J. (1998). Theoretical models: Sociology's missing links. In A. Sica (Ed.), *What is social theory: The philosophical debates* (pp. 238–252). Oxford: Blackwell Publishers Ltd.
- Teller, P. (2001). Twilight of the perfect model model. *Erkenntnis*, 55, 393–415.
- Todd, P.M., & Billari, F.C. (2003). Population-wide marriage patterns produced by individual mate-search heuristics. In Billari, & Prskawetz (pp. 117–138).
- Vance, R. B. (1952). Is theory for demographers? *Social Forces*, 31, 9–13.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

