

**Longitudinal Research in Social Science:
Some Theoretical Challenges**

Thomas K. Burch
Population Studies Centre
University of Western Ontario
London, Ontario, Canada

Abstract

Every advance carries with it potential problems, and longitudinal analysis is no exception. This paper focuses on the problems related to the massive amounts of data generated by longitudinal surveys. It is argued that a proliferation of data may be to the good but it will not necessarily lead to better scientific knowledge. Most demographers think the logical positivist way that theory arises out of empirical generalisations, but massive empirical investigations have only led to disappointing theoretical outcomes in demography. This paper discusses one way out of this impasse - to adopt a different view of theory, a model-based view of science. Theoretical models based on empirical generalisation should become the main representational device in science.

Résumé

Chaque innovation inclue des problèmes possibles, et l'analyse longitudinale ne fait pas exception. Nous attirons attention sur la surabondance de données dans les enquêtes longitudinales. Cette abondance contient des avantages mais elle ne résulte pas nécessairement en une meilleure connaissance scientifique. Beaucoup de démographes semblent penser que la théorie sort des généralisations empiriques, mais l'abondance de données en démographie nous donne souvent des sorties théoriques limitées. Nous proposons ici une autre façon, en adoptant une différente vue de la théorie qui commence avec une science à base de modèles. Les modèles théoriques à base de généralisations empiriques devraient être la principale façon de représenter en science.

Key words: Logical positivism, abstract modelling, the Semantic Model, middle range theory

Introduction

The increasing reliance on longitudinal surveys and analysis in social science represents a major step toward scientific maturity. It shifts our focus to dynamics and process, away from a preoccupation with cross-sectional relationships and equilibrium assumptions. The examination of individual sequences or pathways, defined by states and behavioural transitions among them, moves the analysis closer to concern with mechanisms as well as outcomes. Observations at different times allow for surer inferences regarding causality, especially in the case of subjective factors – attitudes, intentions, motives, etc. – now measured before rather than simultaneously with or after a behaviour to be explained.

But every advance carries with it problems or risks, and modern longitudinal analysis is no exception. One well-known problem is the high costs of longitudinal observation compared to one-time surveys, including the opportunity costs and loss of flexibility involved in long-term commitments of money and other resources. Longitudinal depth on a few areas of human behaviour may be purchased at the expense of breadth of coverage of a wider range of topics. Another set of problems arises from the difficulties of trying to observe the same human beings over extended periods – biases from selective drop-out, effects of observation on respondents, and even the analytic

distractions caused by inconsistent answers from one round to another. In Canada, where almost all national social and demographic surveys are carried out by Statistics Canada, scientific researchers face the dilemma that the richer and more detailed the longitudinal data become, the greater the restrictions on free access, due to Statistics Canada's concerns with privacy and confidentiality.

But the focus of this note is on a different set of potential problems, those related to the sheer mass of data generated by longitudinal surveys, and the more than proportionate increase in the number of statistical analyses possible with such data. In terms of description of the social landscape, such proliferation of data may be all to the good. But it will not necessarily or automatically lead to better scientific knowledge. In particular, the sheer amount of detail may lead to discouragement with respect to theory development. In an older view of theory still very much alive in social demographic circles [the logical positivist program] the more data the better, since theory arises out of empirical generalisations. But a number of contemporary philosophers of science and a few methodologically inclined social scientists question the idea that theory can or must arise from empirical generalisations. And the recent history of demography provides some interesting examples of disappointing theoretical outcomes from massive empirical investigations.

A way out of this theoretical impasse lies in the adoption of a different view of theory than has been commonplace in demography. Philosophers of science have called it the 'semantic' view of theory; Ronald Giere, an American philosopher, thinks that term doesn't mean much to the non-philosopher, and prefers the term 'model-based view' of science (1999). The key to this approach is that models [including theoretical models, or 'theories' viewed as a collection of related models], rather than 'scientific laws' based on empirical generalisation, become the main representational device in science. And criteria of 'truth' or 'validity' are replaced with those of closeness of fit of model to some specific reality, with adequate fit judged in terms of some well-defined purpose.¹

In Section 2, I review some examples of demographic research in which extensive empirical study has had disappointing theoretical returns. In Section 3, I sketch the 'model-based' view of science. Section 4 suggests how this alternative approach to science might help us find a new use for many older – and in some cases rejected – theoretical models, now seen as applicable to longitudinal analysis and the study of behavioural sequences.

On the Theoretical Returns to Data

It is obvious that empirical science needs ample data, or else it would not be empirical science as opposed to armchair speculation. The observable phenomena to be explained or interpreted must be described. But, against

Positivism, the facts do not speak for themselves. Indeed, there is no such thing as a purely empirical fact, a brute empirical datum unsullied by theory or perspective. All human knowledge involves abstraction from concrete reality. It is a simplified version of reality, imposed by a human knower.

Data also are needed to help in the process of scientific judgement that tells us whether a model is a good one, or, which among several models seems the best available. Again, the role of data is not definitive. It has long been clear that all the data in the world cannot prove a theory, given the asymmetry of conditional logic. The modern concept of *under-determination* points as well to the difficulty, perhaps even the impossibility, of definitive disproof, as advocated by Popper. The logical connections between theory and data are seldom tight enough. ‘Verification’ of a theory or hypothesis is more a matter of scientific consensus, a judgement of informed persons that it is the best available, in general or for specific purposes.

In demography, a field noted for its concern with empirical data and with techniques for generating and analysing them, discussions of the inherent limits of empirical work are not commonplace. An interesting exception is a 1975 paper by Keyfitz entitled ‘How do we know the facts of demography?’ He 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.’ [p.267]

Earlier, a classic paper by the biophysicist John Platt (1964) had spoken in similar terms. He tells an anecdote of an early (1958) conference on molecular biology, at which theoretical modellers were criticised by empiricists. 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.” [quoted by Platt, p.348] An empirical 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.

Platt also points out a potential trap set by the mountains of data provided by our current longitudinal surveys. He is speaking of biology, but the description fits social demography as well: “Biology, with its vast informational detail and complexity, is a ‘high-information’ field, where years and decades can easily be wasted on the usual type of ‘low-information’ observations or experiments if one does not think carefully in advance about what the most important and conclusive experiments would be.” [p.349]

Given a few longitudinal surveys, the opportunities for statistical analysis are almost limitless. Researchers need some other criteria than mere possibility to guide their choice of topics. Given the costs of longitudinal surveys, the topics covered will usually be dictated by notions of social importance. But for scientific analysis of the data thus produced, theory is the relevant guide – how will an analysis advance scientific theory? Even massive data collection and statistical analyses are not enough, as the following examples make clear.

Coale on Fertility Transitions

By all accounts, Ansley Coale is one of the most versatile, creative, and influential demographers of the late 20th century. 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 modelling of demographic behaviour [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. And 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 issues, and typically have proven useful in further empirical research by others. He seldom if ever did mathematics for the sake of mathematics.

But like most demographers, he seems not to have been very self-conscious and explicit about the methodology of demography [logic and epistemology] as opposed to technique. Implicitly, there is some ambivalence in his work regarding the proper roles of data, models, and theory.

The *leitmotiv* of his career is formal mathematical modelling of demographic dynamics, popularising and extending the work of Lotka and other early pioneers. This work relies on formally true relationships in highly abstract population models, with the stable model as central. Generalisations emerged from the models rather than from empirical research.

His monograph with Hoover [Coale and Hoover, 1958] also involved abstract modelling rather than broad empirical research, although now the issues are not purely formal and mathematical, but behavioural and contingent. The core of this work was 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. And there is some discussion – but no systematic empirical research – of the wider applicability of the analysis. It was

what today would be called a ‘large-scale simulation,’ hampered, no doubt, by the limitations of early computers.

The central point is that the general propositions that emerged from this project were based on the model, not on empirical data. Indeed, some of the critics of Coale-Hoover [Kuznets, Easterlin] criticised it precisely on the grounds that comparative empirical research showed no strong or regular relationship between population growth rates and economic development.

When Coale turned his attention to fertility transitions, the orientation was more empirical. In one of his earliest papers on the European fertility project (1965), he presented his indirectly standardised 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 a radical logical positivism. Speaking of the decline of marital fertility, he comments: “There are few, if any, universally valid generalizations about the circumstances under which neoMalthusian fertility reduction occurs.” [p.5] After a list of frequently hypothesised 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 (1973) Coale is prepared to deal with what would usually be called theoretical issues, in his IUSSP paper on “The demographic transition reconsidered.” But the emphasis is still on the search for universal empirical propositions, and, interestingly, Coale never uses the word *theory*, either in reference to Notestein’s work or his own ideas [the word does not appear anywhere in the paper]. It is difficult to know whether he was thinking that universal empirical propositions could become the foundation of new theory, or whether he was reverting to a Pearsonian view that social science could at most aspire to finding correlations rather than theoretical laws. In short, it is difficult to know just what his avoidance of the word *theory* means.

But the paper eventually produces some very broad generalisations 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 ‘causal’ 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]. Coale suggests that one or more of the three preconditions can exist in the absence of modernisation.

Coale acknowledges many good points about 'the idea of the transition' [Notestein's transition theory] but faults it finally 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 seem to consider 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 operationalised.²

Coale's last summary 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 rolling logistic model to characterise pre-modern or even pre-historical population dynamics [pp. 3-4]. 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.

On transition theory, Coale seems to have given up the hopes expressed in earlier papers that the project would arrive at 'universal empirical generalisations.' The three preconditions are not even mentioned. One long paragraph [p.24] summarises the causes of transitional mortality decline in broad-brush language that would not have passed muster according to his own standards in the 1973 paper. Ultimately, he writes of the fertility transition in terms not so different from those of Notestein forty years earlier, with reference to 'typical' patterns of transition and some exceptions [p.25]. There is no attempt to develop or quantify the 'idea of transition' beyond the presentation of

empirical measurements, their time trends, and intercorrelations. 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.

This is a long story, but it makes an important point: a massive twenty-year project with substantial resources and collaboration by a large number of first-rate demographers did not result in a substantial improvement in theory. Clearly, the project greatly increased our detailed knowledge of historical fertility declines in Europe, and it clarified some empirical relationships – for example, between delayed marriage and fertility control within marriage, or, the timing of mortality and fertility declines. It even suggested new directions that research and theory might follow, notably by looking into matters of culture and of diffusion. But no new, modified or corrected version of transition theory emerged, nor did any alternative theoretical models, at least not in any well-developed form. Theory did not flow from the detailed data, which revealed no universal empirical generalisations. Nor were there substantial efforts to build theory in response to the data [Lesthaeghe is the main, perhaps the only, important exception].

Hobcraft on Comparative Fertility Surveys

John Hobcraft tells a similar story with respect to the large number of comparative fertility and family planning surveys conducted under the aegis of WFS and its successors. Entitled “Moving beyond elaborate description: towards understanding choice about parenthood” (2000), the paper argues that “the results [of these 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 WFS, have never had a serious comparative analysis capacity (beyond the mainly descriptive Comparative Studies).” He adds that “a profound shift of emphasis is required in order to make real progress.” [p.1]

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 standardised 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 behaviour, but it is also arguable that we shall never remedy the problem without better agreement and testing of comparable information.” [p.2] Hobcraft would seem to agree with Griffith Feeney, who earlier had noted that the surveys in question had a lot of data but not necessarily the right data for testing or developing important theoretical ideas (1994).

Hobcraft’s remedy: “...a serious attempt to reach agreement on the proximate real determinants of fertility (as opposed to the intermediate proximate

determinants) and on how to incorporate measures of these into surveys.” [p.2] He calls for a broad multidisciplinary framework, and greater attention to variables relating to individual decision making and to community-level variables. The aim is to develop ‘global’ regression models, that is, models whose structure applies to all or most populations, even if coefficients may vary.

In short, Hobcraft’s remedy would call for more and better data and for more sophisticated regression models. He recognises the need for theory to help define what data are needed, but the central thrust is toward a more elaborate empirical research program.

Another View of Science

The logical positivist view of empirical science, adhered to by many demographers and other empirically oriented social scientists, seems to be on the wane. It has always been criticised by non-quantitative sociologists, for many of whom ‘positivist’ is a derogatory term. Economics, with its heavy emphasis on theory and abstract models, followed a different version of the positivist program, although it too worried about universal generalisations and Popperian falsification [a key work is Friedman, 1953]. Many social scientists, including some economists, have criticised economics’ abstract models, based on concepts like equilibrium, rational maximisation, perfect markets, and so forth, as ‘having little to do with reality.’ But the stature and influence of economics should give one pause. It has something to do with the discipline’s ability to provide coherent explanations of economic phenomena, instead of just saying how complex they are, and how the data provide a mixed and unclear picture.

But there are alternatives to the logical positivist view, often framed as a negative reaction to it. They are to be found at various times in various disciplines. They differ in detail but share three common elements: 1] they reject the view that theory can or must be built on universal empirical generalisations; 2] they retain an emphasis on rigorous abstract theory as essential to scientific thinking; 3] there is a strong emphasis on judging theory or a theoretical model in terms of the purposes for which it is being used – theory is not empirically true or false [it must be logically true, or true by definition], but close enough or not to some real-world system to be adequate for explanation or prediction.

Meehan on the System Paradigm of Explanation

An early and remarkable statement of this view appears in a small [125 pages] work by the political scientist Eugene Meehan published in 1968. Meehan notes that physics, arguably the most powerful science of all, also does not seem to follow the logical positivist program; nor did it ever. In a powerful critique of the logical positivist ‘covering law’ approach to science and what he terms the

‘deductive paradigm’ of explanation, he asks why physical scientists didn’t attack the logical positivism of such post-WWII philosophers of science such as Nagel and Hempel, whose ideas permeated much of social science, especially sociology. His answer: ‘The deductive paradigm of explanation has not aroused the ire of the [physical] scientists primarily because the scientists have ignored it’ [p.6]. They did not need philosophers of science to tell them how to work, and in any case, the philosopher’s reconstruction of how scientists work apparently is not particularly accurate.

If the covering law approach has not been used in physics, where there are any number of ‘laws’ or universal empirical generalisations, argues Meehan, then certainly it will be frustrating and fruitless in social and behavioural science, where such universal empirical generalisations are rare if not virtually non-existent.

Meehan proposes an alternative approach to explanation that essentially relies on abstract models [Meehan uses the word *system*], with explanation defined as the logical deduction of a phenomenon to be explained from the model. The logical validity of an explanation has nothing to do with data, as in the covering law approach where explanation consists of deriving the *explanandum* from general propositions based on empirical generalisations. The empirical question is whether the explanation is useful for understanding, predicting, and perhaps controlling events in the real world, and this depends on how closely the model fits some real world situation. Meehan uses, or perhaps misuses, the mathematical term *isomorphism*, but he clearly views fit as a matter of degree.

Meehan also argues that purpose should be central in the judgement of how well a system [model] fits the real world, in contrast to logical positivism, where the criterion is purely the logical connection between theory and data, with no particular regard to purpose. Finally, he is suspicious of attempts to explain whole classes of social phenomena on the grounds that such classes often are formed without respect to the theory at hand. A demographic example would be the standard categories of marital status, based on legal-moral-political considerations, not with an eye to scientific explanation. In contemporary language, one might speak of ‘unobserved heterogeneity’ in such classes.

Keyfitz on the Fruitfulness of Abstract Modelling

As mentioned above, Keyfitz has expressed similar views in one of the few extant papers specifically on scientific methodology -- as opposed to technique - - written by a demographer (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.” [p.267]

Longitudinal Research in Social Science: Some Theoretical Challenges

Rather, much of our most solid and important demographic knowledge is derived from work with theory or models.

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 issue of the effect of marriage delay on completed fertility, that of promotion in organisations, and the effects of development on population growth – all questions involving behavioural models on which there is less consensus than on the stable model used to solve the problems on age structure.

The important point is that Keyfitz attributes our firm answers to these issues to work with theory or models. With respect to growth and proportions over 65, he notes: "This simple introductory example shows how uncertain our knowledge would be if analytical tools like the stable model were not available. One can imagine extensive research projects for describing the various extraneous factors, methodological controversies, and schools of opinion, some perhaps taking the view that the relation was really different for different races or different continents. One who has been through the theory would no sooner say that the underlying relation between growth and age compositions is different for continents than he would say that the laws of thermodynamics differ from country to country." [p.273]

It is important to note also that Keyfitz does not make a sharp distinction between formal models [e.g., the stable model] and behavioural 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].

Note the parallel between Meehan's notion of isomorphism, and Keyfitz's concern with 'how closely this constructed machine resembles the one operated by nature.' Neither asks whether theories or theoretical models are empirically true, but whether they fit some particular piece[s] of the real world.

The Semantic or Model-Based View of Science

As noted earlier, the logical positivist view of science has dominated social science, including demography, in the latter half of the 20th century. According to this view, theory – a summary of what is known in a field – must be based on valid empirical generalisations 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 generalisation, etc. – 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.

Contemporary philosophy of science has increasingly challenged this view, arguing that the classic logical positivist view of Nagel (1961) or of Hempel (1965) 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 idealisations of certain features of an indefinitely complex real world. Nor are they so much ‘discovered’ in nature as constructed by the human mind. 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. Without specific reference to it, both Meehan and Keyfitz espouse what has come to be known as the *semantic* school of the philosophy of science.

A leading representative of this school, Ronald 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]³

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 [compare point b to Keyfitz’s phrase ‘degree of realism’ and Meehan’s

notion of isomorphism]. 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, or a three-body 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 a particular 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 used to loose inference in explanation that its explicit statement is necessary.⁴

Note that in this account of science, all models are *formally* true [assuming, of course, no logical errors or internal contradictions], that is, true by definition. The empirical question then becomes one not of empirical truth or validity, but whether a valid model applies to a particular empirical case.

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, with respect to calculation of the average annual growth rate over a period is also applicable to every real-world population. With respect to describing a population's growth trajectory, however, the exponential growth formula applies more or less to some populations, but is not at all applicable to others.

A behavioural model such as the theory of demographic transition can be stated in such a way that it is rigorous and formally true – although in fact there have been few attempts to do so. Its applicability to the real world is another question, and has been a matter of debate ever since transition theory first appeared. But it is worth noting, in terms of Giere's criteria of applicability, that it correctly represents a large number of actual cases of mortality/fertility decline, at least in qualitative terms.⁵

In my reading of Giere's and Cartwright's 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 a certain purpose. In science, that purpose often will be explanation or prediction as opposed to practice. And in some schools of computer modelling, the emphasis is on less abstract models, trying to capture more and 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. Theories or models can be more or less complex; they can be quantitative or qualitative; they can be stated in words, mathematical formulas, or – as is increasingly the case – in the form of computer simulations. But their epistemological status and the way they are used for prediction or explanation are fundamentally the same.

Note that this position does not agree with the view in some sociological and cultural studies circles 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.’ Nor does it have anything to do with ‘critical theory,’ in which the analyst judges some portion of the real world in terms of values, as good or bad.

‘Social Mechanisms’: A Return to Middle-Range Theory

Within the discipline of sociology, a perennial school of thought has emphasised the development and use of ‘abstract analytic theory.’ But in recent decades, their influence seems to have been outweighed by quantitative empirical analyses, analyses that often as not are only loosely linked to theory. The ideas of this school resonate with those of Meehan, Keyfitz, and the ‘model-based’ school of philosophy of science. A recent collection of essays, *Social Mechanisms: An Analytical Approach to Social Theory* [Hedström and Swedberg, 1998], suggests a possible comeback.

In their introductory essay, Hedström and Swedberg 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 an approach that would 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]⁶ Mertonian middle-range theory, in their view now out of favour, is seen as an appropriate middle ground between pure description and the search for social laws.

The search for mechanisms [or underlying processes] is seen as different from 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 or structural equation 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 can be seen as a systematic set of statements that provide a plausible account of how *I* and *O* are linked to one another [compare Meehan’s ‘system’].

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 empirical validity: “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.” [p.15]

Theory for Longitudinal Analyses

The last fifty years of demography give warning that large-scale and detailed data analyses may have less than proportionate theoretical returns to the effort and resources devoted to them. The case studies described earlier can be interpreted in this sense, although not all would agree with this interpretation. But if it is at all close to the mark, then the recent enthusiasm for longitudinal surveys needs tempering. It is not that longitudinal data should not be collected or analysed, but that the planning of surveys and analysis needs to be informed with a different spirit and a different view of science. It would be unfortunate if we were to assume once again that ‘the facts speak for themselves,’ or that ‘theory will flow from the data.’

Keyfitz, Meehan, some contemporary philosophers, and the ‘social mechanisms’ approach all agree on a different and more fruitful way to proceed. But it requires a change in ways of thinking that may be hard for many demographers and other empirical social scientists. At a very simple level, it requires more attention to theory and to the development of theoretical models, even if this seems to be at the expense of data collection and statistical analysis.

Demography needs to take theory seriously – something it has hardly ever done – at least if it wants to aspire to status as a science rather than a set of techniques.

But our understanding of theory-data relationships also needs to change. We need to see that theory and data occupy different planes, as it were, even if these planes intersect, as they must in an empirical as opposed to a purely speculative or formal discipline. Theory and the development of theoretical models need to be granted provisional autonomy. We need to learn not to reject models because they are ‘oversimplified’ or because they do not agree with a particular data set, or even several data sets. Sensible computer simulations need to be seen, not as attempts to ‘make up data,’ but as instructive efforts to represent the inner workings of complex systems. Otherwise theory development is cut off at the very start, and the models never really get off the ground.

The autonomy for theory development remains provisional because at some point it must be demonstrated that a given theory or model is useful for explanation or prediction with respect to some real-world phenomenon. It is not a case of theory development as mental recreation, with total disregard for empirical data.

But apart from general methodological orientations, are there any particular theoretical directions that might be particularly valuable for longitudinal analyses? Certainly the ‘social mechanisms’ approach sketched above would be a prime candidate. It urges us to dig deeper, below our multistate analyses to get at the mechanisms and mental processes that move a person from one state to another. Clearly, at the individual level, these will be decision mechanisms, including the decision to imitate others or to conform to fad, fashion, and cultural norms [recall the discussion above of Hobcraft’s emphasis on decision-making variables in future fertility surveys]. To the extent that longitudinal data sets do not directly measure such mental or psychological variables – and there will invariably be some gaps – explanation can only be done by theory. Multivariate statistical models, even dynamic models, will not suffice. Too many relevant factors will remain in the black box, with no sense of their role in the transformation of measured inputs into measured outputs.

A focus on decision making suggests we take a fresh look at large bodies of contemporary and older theoretical work. These would include microeconomics, including more recent work on sequential decision making, the large psychological literature on decision making, and social and psychological exchange theory. Hedström and Swedberg, in their book on social mechanisms, hearken back to the Columbia school of ‘middle-range’ theory, notably Robert Merton [see especially 1957]. One senses that his essays on ‘social structure and anomie’ and on reference group behaviour could help greatly in the understanding of outcomes of longitudinal analysis – life-cycle changes, life-courses, and sequences. But this work is now old, and is not necessarily

Longitudinal Research in Social Science: Some Theoretical Challenges

adequate in its original form. At least two things are required by way of development: a) restatement of the some of the underlying models in more rigorous form, perhaps as computer models: b) an introduction of more dynamics into the models, to deal with process and changes over individual lifetimes or large segments thereof. Merton was a 'structural functionalist,' and although there are dynamic elements in his work, the focus is on equilibrium and comparative statics.

Dynamic models suitable for looking at event-history data on individuals will emphasise often changing relations to authority figures, social norms, and culturally prescribed goals. In Mertonian terms, some deviants become conformists, some conformists become innovators or ritualists, and so forth (1957, p.140). People learn new ways of relating to life circumstances and the larger social structure. And this learning often lies beneath the kinds of changes of state captured in longitudinal analysis.⁷

A neglected work by Hanneman (1988) is full of good ideas for dealing with these kinds of models, using a systems approach to theory building. In Ch.6, for example, he illustrates 'goal-referencing' systems, but also goal-setting systems, in which the actor changes goals over time, based on prior experience. In order to handle the complexity involved in even simple versions of such dynamic systems, due to things like feedback and delay, Hanneman advocates the construction of models using systems dynamics software. What appears to have been a poor reception of the book may have been due in part to its title – with the phrase 'computer-assisted theory building' putting off both social theorists uninterested in computers, and 'hard' computer modellers who did not particularly equate their work with theory building, or whose mathematical and computer programming skills went well beyond the level represented by this particular work.⁸ But, Hanneman offers many potentially fruitful leads to the building of models relevant to longitudinal analysis.

These are only a few examples of bodies of work that come immediately to mind. And there is no shortage of good theory for interpreting and explaining longitudinal findings. But it must be read and used systematically, and adapted to this special purpose. It is not enough to lift a few ideas simply to give 'theoretical orientation' to an otherwise purely empirical description. [McNicoll (1992) has reminded us that the main use of microeconomics in demography has been 'heuristic'].

Conclusion

Science involves a balance between theory and data or observation. This is but a restatement of John Locke's view of human knowledge as a blend of experience and reflection on experience, and few would question it directly. But not a few demographers and other empirical social scientists have questioned it by their actions, through neglect of systematic work on theory. At the individual level, the choice to emphasise theory or empirical work is a matter of taste. But no discipline can claim to be a science without good theory. In one sense, theory is the most important element in science in that it summarises what we know about how the world works at some deeper level.

Theory does not just summarise empirical observations as the logical positivists would have it. Nor does theory flow automatically from the facts, systematically analysed using the most sophisticated statistical techniques. Theory must be constructed; it is the result of an act of the creative imagination of the scientist. It is a response to empirical observation but should not be limited by it, particularly in the early stages of theory formation. It could be argued that in demography, theory has been smothered by data.

Some of the ideas reviewed above suggest that the problem has not been one of simple neglect, but also one of a misunderstanding of the nature of theory, and its development and uses.

Longitudinal surveys hold great promise of scientific advance in social and behavioural science, including demography, if for no other reason than that they bring us closer to empirical measurement of process and mechanisms as these occur over time. That is, the resulting descriptions can be more realistic, they have opened some corners of the behavioural black box to direct observation. But they cannot eliminate the black box. There will be relevant phenomena resistant to direct observation. The language for studying and discussing these is theory, not statistics.

Acknowledgements::

This is a summary of remarks at the Workshop on Longitudinal Research in Social Science – A Canadian Focus, Population Studies Centre, University of Western Ontario, London, Canada, 25-27 October 1999. An earlier version was presented at a Workshop on Synthetic Biographies: State of the Art and Developments, organised jointly by the University of Groningen and the University of Pisa, San Miniato, Italy, 6-9 June 1999. Ongoing research on modelling and demographic theory has been supported by a grant from the Social Sciences and Humanities Research Council of Canada.

End Notes:

1. Another leading philosopher of this school is Cartwright (1983, 1999). Among social scientists, Meehan (1968) was very early in suggesting a similar approach as an alternative to logical positivism or the covering-law approach.
2. Only in 1997 has a serious effort been made to do this. See Lesthaeghe and Vanderhoeft, 1997.
3. 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 all the more forcefully to other scientific disciplines, for example biology or the social and behavioural sciences.
4. 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 emphasised by Platt (1964).
5. 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 lags, slopes, etc., or by introducing probabilistic elements.
6. For a very different view of the status of 'laws' in physics, however, see Giere (1999) and Cartwright (1983, 1999).
7. In psychology, some attempts have been made to cast learning theory in the form of Markov models. See, for example, Wickens, 1982.
8. There are other issues. Hanneman's modelling package of choice was Dynamo, software associated with the MIT school of 'systems dynamics' and *The Limits to Growth*, and therefore discredited in the minds of many. But the naivté of some early uses of this software have not destroyed its potential usefulness, and, in any case there are other, scientifically more sophisticated packages available, such as Vensim and Modelmaker.

References:

- Cartwright, N. D. 1999. *The Dappled World: A Study of the Boundaries of Science*. New York: Cambridge University Press.
- Cartwright, N. D. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Coale, A. J. 1973. "The demographic transition," International Union for the Scientific Study of Population's *International Population Conference, Liege*. 1: 53-72.
- Coale, A. J. 1965. Factors associated with the development of low fertility: an historic summary. United Nations World Population Conference, Belgrade, Yugoslavia. WPC/WP/194.
- Coale, A. J. 1956. "The effects of changes in mortality and fertility on age composition," *Milbank Memorial Fund Quarterly* 34: 79-114.
- Coale, A. J. and Hoover, Edgar M. 1958. *Population Growth and Economic Development in Low-Income Countries*. Princeton: Princeton University Press.
- Coale, A. J. and Watkins, S. C. 1986. *The Decline of Fertility in Europe*. Princeton: Princeton University Press.
- Feeney, G.. 1994. "Fertility decline in East Asia," *Science* 266: 1518-1523.
- Friedman, M. 1953. *Essays in Positive Economics*. Chicago: The University of Chicago Press.
- Giere, R. N. 1999. *Science Without Laws*. Chicago: The University of Chicago Press.
- Hanneman, R. A. 1988. *Computer-Assisted Theory Building: Modeling Dynamic Social Systems*. Newbury Park, CA: Sage Publications.
- Hedström, P.; and Swedberg, R. 1998. *Social Mechanisms: An Analytic Approach to Social Theory*. Cambridge University Press.
- Hempel, C. G. 1965. *Aspects of Scientific Explanation*. New York: The Free Press.

Longitudinal Research in Social Science: Some Theoretical Challenges

- Hobcraft, J. 2000. Moving beyond elaborate description: towards understanding choices about parenthood. Paper presented at U.N. Commission for Europe, Fertility and Family Surveys Flagship Conference, Brussels, Belgium.
- Keyfitz, N. 1975. "How do we know the facts of demography?," *Population and Development Review* 1: 267-288.
- Lesthaeghe, R. and Vanderhoeft, C. 1997. Ready, willing and able: a conceptualization of transitions to new behavioral forms. IPD Working Paper, Interface Demography, Vrije Universiteit Brussel.
- 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, Ill.: The Dorsey Press.
- Merton, R. K. 1957. *Social Theory and Social Structure* [revised edition]. Glencoe, Ill.: The Free Press.
- Nagel, E. 1961. *The Structure of Science*. New York: Harcourt, Brace, and World.
- 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.
- Wickens, T. D. 1982. *Models for Behavior: Stochastic Processes in Psychology*. San Francisco: W.H. Freeman and Co.