New Methodological Approaches in the Social Sciences. An Overview

Daniel Courgeau

Population: An English Selection, Vol. 10, No. 1, New Methodological Approaches in the Social Sciences (1998), 1-9.

Stable URL:
http://links.jstor.org/sici?sici=1169-1018%281998%292%3A10%3A1%3C1%3ANMAITS%3E2.0.CO%3B2-B

Population: An English Selection is currently published by Institut National d'Études Démographiques.

NEW METHODOLOGICAL APPROACHES IN THE SOCIAL SCIENCES
An overview

Daniel COURGEAU*

The social sciences have traditionally developed independently of each other, elaborating their own objects of study and their own methods. At present, however, we are witnessing an increasing convergence of approach at the same time as an extension of their fields of application. Methods derived from the same type of statistical model are now being employed in biometrics, demography, economics, human geography, educational science, sociology, and so forth. Naturally it is not purely fortuitous if these models are found to perform well in different fields; it is because they reflect with increasing accuracy the complexity of the social phenomena at work.

Event history analysis, for example, originated in the work of probabilists and statisticians to produce a correct formulation of the occurrence over time of very general processes that integrated the effect of various characteristics which were themselves also subject to change over time (Cox, 1972; Kalbfleisch and Prentice, 1980; Cox and Oakes, 1984; Andersen *et al.*, 1993). With these methods it became possible to explore such disparate social phenomena as changes in employment, migrations, marriage, education, ethnic conflicts, changes in political regimes, amongst others. The result is that this approach has been generalized and adapted to the problems encountered in sociology (Coleman, 1980; Tuma and Hannan, 1984; Blossfeld *et al.*, 1989; 1992), in econometrics (Lancaster, 1990), in demography (Courgeau and Lelièvre, 1989; 1992) and in many other social sciences.

An innovation of this kind has the potential to challenge the validity of established paradigms. Under the classic paradigm in demography, analysis was of a single phenomenon that was considered to be independent of all others, and took place in a sub-population which was required to remain homogeneous (Henry, 1959). This paradigm was acceptable so long as analysis was of simple data sets (from vital registration, for example)

* Institut national d'études démographiques.

Population: An English Selection, special issue *New Methodological Approaches in the Social Sciences*, 1998, 1-9.

but became increasingly untenable with the development of detailed surveys, such as the World Fertility Survey, which provided a wealth of individual-level information. It was no longer possible simply to ignore the interactions between different demographic phenomena or the heterogeneity of populations. These difficulties can be avoided by using event history analysis, since it postulates that "throughout his life an individual follows a complex itinerary, which at any given moment depends on his life-course to date and on the information he has acquired in the past" (Courgeau and Lelièvre, 1997). This new paradigm opens the way for analysis of the interactions between the different events that occur in the lifetime of an individual. It also makes it possible to identify the different forms of individual behaviour which exist in a heterogeneous population.

Our purpose here is not to give a detailed presentation of an approach which is now fully integrated into the social sciences (Keilman, 1993), but to present in this special issue of *Population* some of the new methodological developments which may, or may not, prove fruitful in the future.

From the individual to the group

First, a major and recurrent problem in the social sciences concerns the emergence of behaviour specific to groups made up of similar members. There are two contradictory approaches to this problem: one treats the aggregated behaviour as the sum or average of the individual behaviours and holds that there is no need to introduce a specific group effect, while the other holds that the group effect is preponderant and that the individual behaviours are the reflections of the underlying overall causes and as such have no intrinsic interest for the social scientist (Alexander *et al.*, 1987). To go beyond these approaches attention must focus on establishing the existence of aggregated behaviour and linking it to the behaviour of the members of the group. Efforts are needed to identify the processes which lead to the formation of highly adaptive aggregates and the nature of the boundaries that demarcate them, and to explain how the interaction between members of the group generates behaviours that transcend those of the component agents (Holland, 1995).

These problems are explored in two of the articles presented here, using approaches which are different but not contradictory, since they are related to the size of the group observed.

E. Lelièvre, C. Bonvalet and X. Bry examine groups of limited size (family, household, contact circle). In this case the analysis is able to treat at a detailed level the structure of the group and the interactions between its different members. However, although a family or household can be accurately defined at a given point in time when using a cross-sectional analysis, these notions are of limited use in a longitudinal or event history study (Keilman and Keyfitz, 1985). In this case it is better to focus on what the authors refer to as the *contact circle* ('entourage'). This is defined by tracking a key individual

throughout his life and is taken to comprise the members of the various households to which they have belonged and the relatives and friends with whom they have not cohabited but have maintained close relationships. The contact circle thus defined can then be followed, from the point of first contact of each member with the key individual. Correct observation of this new entity requires innovative survey designs. The behaviour of the contact circle will then have to be modelled, by making hypotheses about the relationships between the different members. The event history models required will be much more complex than the usual models, but can be expected to increase our understanding of the behaviour not just of individuals in isolation but of groups such as the contact circle.

D. Courgeau and B. Baccaïni adopt a different approach for examining groups that are less structured (pupils in a class, for example) or much larger, such as the inhabitants of a region, and for which a structure cannot realistically be introduced. The solution here is to employ models which do not individualize the members of a given level of aggregation but in which several levels corresponding either to hierarchies or to cross classifications can be examined. *Multilevel models* satisfy these requirements. These treat individuals as having behaviour which depends on their own characteristics, on the aggregated characteristics of the different zones they inhabit and on the random variables corresponding to these zones. But these models still suffer from the problems associated with unobserved heterogeneity, even when this is independent of the observed characteristics, and with development of an authentically multilevel event history analysis. That said, they are suitable for application in many social science fields where they can provide answers to important questions (Goldstein, 1995).

Events and the words to describe them

The *analysis of textual data* is a rapidly expanding domain which has many applications in the social sciences and provides new techniques for the analysis of complex biographical histories. Although these methods were originally developed for the purpose of stylistic analysis (Lebart and Salem, 1994), they have been found to have close affinities with event history analysis: they work with the itineraries formed by word sequences or stages that are structured through time, like personal life histories. The itineraries involved are much more complex, however, due to the diversity of words in a language and to the great variety of the sequences considered. It follows that this is primarily a descriptive analysis, which seeks to identify the course of the successive states an individual has occupied. However, it can also contribute to an understanding of the itinerary followed, by introducing diverse characteristics of the respondents and by classifying the itineraries according to various typologies.

In her presentation of these methods of analysis, F. Guérin-Pace indicates how they can be applied to very diverse fields of the social sciences.

Originally, of course, these methods were not intended for use in demography. Only recently have the social sciences begun to be applied in three important domains: analysis of responses to open questions, analysis of narratives and discourse, and above all the analysis of complex itineraries. In the *Population, living space and environment* survey conducted by INED, the meanings attached to the word 'environment' are found to be revealing of the different interpretations that can be given to a concept that is ostensibly straightforward but that each invests with a particular significance. By introducing the characteristics of the respondents these differences of meaning can be linked to different individual profiles. The study of narratives, particularly by using methods of classification, illustrates and explores the evolution of discourse. Lastly, the tracking of complex professional itineraries, using data from the *Labour Force survey* conducted by INSEE, shows how a textual analysis can replace an event history analysis when the number of different itineraries observed is very high (nearly 300 in the case studied). Unlike event history analysis, however, this approach is more descriptive than explanatory.

Models for the evolution of populations

Another area in which event history analysis has found new applications is that of population projections. The use of *microsimulation techniques* in this field is presented by E. Van Imhoff and W. Post. With these methods it becomes possible to make a highly effective use of the results of an event history analysis, by introducing the probabilities of experiencing the various events according to the state of the individual and the time elapsed since the other events of interest: fertility, for example, can be studied in relation to age, duration of marriage, and birth order. The process being simulated is then random and is used not only to estimate the expected values of the different characteristics of the populations over time but also to give an indication of their standard errors. Here too, by integrating a wide range of socio-economic events, these simulations can produce models which are more complete than simple population projections. However, the projections obtained remain within the classic framework of linear models.

Let us examine in greater detail how these linear models have been formulated in the past, by working at a more aggregated level, a country, for example, or a country divided into regions. If it is assumed that demographic behaviour is represented by age-specific probabilities which are assumed to be stable over time, by multiplying these probabilities by the corresponding initial populations, it is possible to produce a linear model for the evolution in their numbers. For example, combining age-specific fertility and mortality probabilities produces models which lead to stable or stationary populations (Keyfitz, 1968). If in addition we introduce age-specific emigration probabilities between the different regions, we still end up with

stable populations, first of all for all ages combined, then by each age group (Rogers, 1975). All other inputs to these models preserve their linear form and so continue to produce stable or stationary equilibrium solutions.

In these conditions it is reasonable to wonder if such models are realistic and if it is not necessary to introduce nonlinear variations. For example, the prey-predator system can only be correctly modelled by introducing a term of interaction that corresponds to the product of the competing populations at any point in time. Likewise, a nonlinear system that produces a better fit to empirical evidence, can also be obtained by introducing into a multiregional model an index of migration intensity, in the form of the ratio of the flow between two regions to the product of their respective populations.

The new properties introduced by such nonlinear models are presented here by D. Blanchet, who discusses the usefulness of some of them in demo-economics.

First of all, these models usually no longer lead to stable populations. Populations can vary cyclically over time but without ever reaching a stable state. Some of these populations may die out, while others may vary in an acyclical way. The most interesting, however, is that this dynamic can sometimes become chaotic, possessing no cyclical component. In these conditions it becomes hard to forecast the evolution of these populations over time, since very small deviations from the initial states can result in entirely different patterns of change. If this property was found to be confirmed by human populations it would imply that some social phenomena were intrinsically unpredictable, a result which is in contradiction with the classical theories (Keilman, 1992). D. Blanchet here offers some critical observations on this important problem. It must be noted, however, that the complexity of behaviour is such that although it appears to be chaotic, this cannot be attributed to erratic fluctuations in the parameters of the process, since the model itself is deterministic.

So now it is time to examine what happens when a random component is introduced into these models. This can be done by means of *viability theory* (Aubin, 1990) which is concerned with the evolution of nonlinear macro-systems in the absence of any determinism. The theory also assumes the presence of constraints that a system must either respect or cease to exist. It follows that the system must possess controls that it is able to operate in order to survive. This theory cannot be used to forecast the evolution of an initial system; instead it indicates the limits beyond which some forms of behaviour (the controls) have to be modified in order to keep the system viable. The extremely general nature of these principles means that the theory has applications in many social sciences: anthropology, demography, economics, sociology, etc.

The basic principles and main strengths of viability theory are presented here by N. Bonneuil, who successfully applies it to the social and economic problems facing the populations studied by the anthropologist

R. Barth. A social organization whose existence is subject to a number of constraints, such as the Basseri nomads of Persia or the Norwegian fishermen examined here, can maintain itself, not in a state of equilibrium but on a viable path, by means of a certain number of strategies, which are its controls. For the nomads this control takes the form of sedentarization; for the fishermen it is the decision to move to new fishing grounds. These are the strategies which enable these social groups to survive over time. However, this theory is not deterministic: these groups have a certain freedom of action, on condition that they remain inside the system's kernel of viability. So despite the great diversity of actual situations it is possible to identify the regularities which underpin, though do not determine, the existence of these social groups. The absence of determinism at the macro level in viability theory, corresponds to the absence of determinism at the individual level in event history analysis.

Space, networks, diffusion

To complete this special issue we need to consider the spatial context in which social phenomena occur. This question has already received indirect attention in the multilevel analyses, where the levels of aggregation can be made up of communes, regions or towns. It must now be introduced explicitly, by examining the structures and morphology that are produced by the societies and social networks for which it is the setting, as well as its hierarchies and discontinuities.

It is clear that with the exception of human geography and epidemiology the other social sciences have traditionally treated space as a secondary element of only limited interest for the study of human society. While the temporal dimension has been fully integrated into the analyses of demographers, economists and sociologists, they accord little or no attention to the spatial dimension. P. Gould (1993) has shown recently how even epidemiologists may suppress any consideration of the spatial dimension of the AIDS pandemic and consider it as a purely temporal phenomenon, something he believes would be a fundamental error. The role of space is a priority for research, not just for the phenomena in which it is directly involved, such as internal and international migration and the spread of epidemics, but for the entire range of social phenomena.

In their article, J.-P. Boquet-Appel and L. Jakobi explore the *diffusion of the demographic transition*, something that has usually been considered as an essentially social process, descending through society from the elites to the unqualified rural proletariat. However, if this diffusion occurred on a step by step basis within Europe it should be possible to trace its progress on geographical maps. This has been made possible by the Princeton Project, which has gathered data at a detailed geographical level. Such an analysis is not as straightforward as it appears, however, since there are numerous pitfalls attached to tracing the evolution over time in spatial data

of this kind. The techniques used by the authors to examine this problem are drawn in part from geostatistics and in part from epidemiology. They confirm that contraception does seem to have spread in Great Britain by a contagion-like process, albeit over a much shorter period of time than in France. The task now is to test this hypothesis at the level of Europe as a whole and to identify the characteristics which are valid for such a large spatial context. This project puts the frictional role of space back at the heart of the diffusion of social processes and clearly illustrates the dangers of excluding the spatial dimension from the study of human societies.

We must now turn to consider the problems associated with the distribution of population over a particular territory. The aim is to identify the origins and forms taken in different cultures of the rules that govern this spatial distribution. Towns and cities are especially interesting structures to study in this respect. Notwithstanding the large and constant flows of people, companies and public services, to which they are subject, and although they experience continuous change in their buildings, channels and means of communications, and surface areas, these urban structures can be stable over very long periods (Derycke *et al.* 1996). In addition, although towns and cities are the products of civilizations and cultures that present great diversity, they are a universal form of geographical organization present across the globe.

P. Frankhauser turns his attention to this form and to the conditions of its persistence over long periods of time. It is time to move away from the old analogy between urban systems and physical systems, based on the use of gravity models and entropy-maximizing models (Wilson, 1967). These are too deterministic and are being replaced by a number of models which emphasize the way in which local decisions that are unplanned and uncoordinated can give rise to coordinated global patterns, which define the size and shape of the towns and cities under observation. In this perspective towns appear as self-organizing structures which emerge from local actions. This leads to *fractal clusters*, whose evolution over time can be modelled. With these models it becomes possible to understand the development not only of the single-centre cities associated with the industrial revolution but the poly-centric urban areas of today. At this point we can bring our discussion full circle and return to the problems of aggregation raised at the start of this introduction: can an aggregated structure such as a town or city be considered independently of the individuals who make up its population? The extension of the urban models presented here may provide an answer.

*
* *

This brief overview of new approaches in the social sciences makes no claim to exhaustivity. It has shown the extreme vitality of research in this area and the rich variety of methods being applied. The development of these approaches has not occurred in a random fashion, however, and

in this presentation we have tried to show the links that exist between them. We have identified a number of underlying assumptions, which can now be enumerated.

First, the individuals and groups that come under examination are *heterogeneous*, in terms both of their past history and of their possible courses of action. The same diversity is found in individual itineraries, in the word sequences of a text, and so forth.

Second, the action of these individuals and groups can only be understood as an *interaction* between events experienced in the past or experienced by other members of the group, and the events that we wish to study.

Third, it is necessary to consider the effects of *aggregation*, which are responsible for forms of behaviour which are not the sum of the behaviour of units when aggregated. If the fact of belonging to a group can have effects on the behaviour of individuals, it is reasonable to wonder if the group as such does acquire a degree of independence vis-à-vis its members.

Fourth, attention must be given to the group behaviours which lead to *nonlinear* or *nondeterminstic* models, or to both at once. Some of these models may produce chaotic behaviour, whose validity must be verified. It also appears useful to find ways to increase understanding of the links between the behaviour of these groups and that of the individuals who compose them.

Fifth, both individual and aggregate level models will benefit from taking into account the various kinds of *flows* which occur in physical space or in a more complex space, such as a living space.

The task now is to continue research along these different paths, as well as along others not mentioned here, in order to identify the hypotheses that withstand empirical verification and those which must be discarded because they are inadequately confirmed or are contradicted by observations. This special issue of *Population* is intended to show the extent of the renewal that has affected the social sciences in recent years. Future research will determine which of these developments are of lasting value.

Daniel COURGEAU

ACKNOWLEDGEMENTS

We would like to thank Godfrey Rogers for his translation or revision of the articles in this special issue.

BIBLIOGRAPHY

ALEXANDER J., GIESEN B., MÜNCH R., SMELSER N., (1987), *The micro-macro link*, University of California Press, Berkeley.

ANDERSEN P., BORGAN O., GILL R., KEIDING N., (1993), *Statistical models based on counting processes*, Springer-Verlag, New York.

AUBIN J.-P., (1990), "A survey of viability theory", *SIAM Journal of Control and Optimization*, vol. 28, n° 4, pp. 749-788.

BLOSSFELD H.-P., HAMERLE A., MAYER K.-U., (1989), *Event history analysis*, Lawrence Erlbaum Associates, Hillsdale NJ, Londres.

COLEMAN J., (1981), *Longitudinal Data Analysis*, Basic Books, Inc., New York.

COURGEAU D., LELIÈVRE É., (1989), *Analyse démographique des biographies*, Éditions de l'Ined, Paris.

COURGEAU D., LELIÈVRE É., (1992), *Event history analysis in demography*, Clarendon Press, Oxford.

COURGEAU D., LELIÈVRE É., (1997), "Changing paradigm in demography", *Population. An English Selection*, n° 9, pp. 1-10.

COX D.-R., (1972), "Regression models and life tables", *Journal of the Royal Statistical Society*, B.34, pp. 187-220.

COX D.-R., OAKES D., (1984), *Analysis of survival data*, Chapman and Hall, London.

DERYCKE P.-H., HURIOT J.-M., PUMAIN D., (1996), *Penser la ville. Théories et modèles*, Anthropos, Éditions Economica, Paris.

GOLDSTEIN H., (1995), *Multilevel Statistical models*, Edward Arnold, London.

GOULD P., (1993), *The slow plague*, Blackwell Publishers, Cambridge, Massachussets.

HENRY L., (1959), "D'un problème fondamental de l'analyse démographique", *Population*, 1, pp. 9-32.

HOLLAND J., (1995), *Hidden order : how adaptation builds complexity*, Addison Wesley Publishing Company, Reading, Massachussets.

KALBFLEISCH J.-D., PRENTICE R.-L., (1980), *The Statistical analysis of failure time data*, John Wiley and sons, New York.

KEILMAN N., (1993), "Emerging issues in demographic methodology", *European population, II. Demographic dynamics*, Congresses and Colloquia, John Libbey/Ined, Paris.

KEILMAN N., KEYFITZ N., (1988), "Recurrent issues in dynamic household modelling", *Modelling household formation and dissolution*, Clarendon Press, Oxford.

KEYFITZ N., (1968), *Introduction to the mathematics of population*, Addison Wesley, Reading, Massachussets.

LANCASTER T., (1990), *The econometric analysis of transition data*, Cambridge University Press, Cambridge.

LEBART L., SALEM A., (1994), *Statistique textuelle*, Dunod, Paris.

ROGERS A., (1975), *Introduction to multi-regional mathematical demography*, John Wiley and sons, New York.

TUMA N., HANNAN M., (1984), *Social dynamics. Models and methods*, Academic Press, Orlando, Florida.

WILSON A.-G., (1967), "A statistical theory of spatial distribution models", *Transportation research*, 1, pp. 253-269.