C H A P T E R

23

Demographic Event History Analysis

DANIEL COURGEAU AND EVA LELIÈVRE
Institut national d'études démographiques (INED), Paris, France

As was seen in Chapter 10, traditional longitudinal analysis examines the occurrence of a single demographic event, and that event alone, by eliminating the effect of all other phenomena that are treated as interfering events. This form of analysis was developed before large amounts of individual-level data became available through specific surveys and before major advances in probability theory and statistical techniques. Longitudinal analysis studies the phenomena of interest isolated in their *pure state* and uses primarily civil registration or population register data to calculate series of probabilities or rates by age or by durations between two successive transitions. Obviously, this approach is purely descriptive.

The growth of survey-taking in the 1980s (Groupe R. A. B., 1999) and developments in probability and statistical techniques (Cox, 1972; Kalbfleish and Prentice, 1980; Courgeau and Lelièvre, 1989, 1992, 2001; Andersen *et al*., 1993; Lelièvre and Bringé, 1998) equipped demographers to switch their attention from aggregate-level data (the numbers of individuals experiencing an event in a particular year) to individual-level data (an individual with certain characteristics experiences various events at different times and in different contexts over the length of his or her life course). The result is an explanatory approach that attempts to disentangle the mass of factors to identify the interdependencies between phenomena and assess the influence on behavior over time of individual characteristics.

Studying these complex individual histories or trajectories requires the elaboration of an analytical and methodologic framework that generalizes the ideas outlined in the previous paragraphs.

I. FROM LONGITUDINAL ANALYSIS TO EVENT HISTORY ANALYSIS

Traditional longitudinal analysis is based on a double postulate: The demographer studies the occurrence of a single event in a population that remains *homogeneous* for as long as that phenomenon is present and from which the effect of the interfering phenomena, considered to be *independent* of the phenomenon being studied is eliminated (Henry, 1959).

Satisfaction of these conditions is the prerequisite for calculating the conventional demographic rates and probabilities, which are taken to express the risk of experiencing the event in question assuming no losses from observation or interfering events. Chapter 9 showed that the results thus obtained need to be treated with a degree of skepticism. We will now go further and show that these postulates have to be changed if the behaviors are to be analyzed in their true complexity.

Let us begin by considering the condition of independence. Empirical observation and simple reasoning tell us that there is not usually independence between processes (Henry, 1959). It is, for example, reasonable to expect migration to a metropolitan area to alter the fertility behavior of a female migrant from a rural region.

Accordingly, we need to discard this hypothesis and introduce the effects of the other phenomena into the analysis, not by excluding the individuals who experience them as in the traditional analysis, but by assessing their influence on the processes being studied. Also note that when analysis is limited to a single event, the restrictive assumption of independence makes it impossible to study losses from observation due to competing events, as with cause-specific mortality, for example. Nor can we work on a subgroup of the population that can be entered by several different events (as in the case of direct entry to a population of manual workers at the end of schooling or indirect entry after a series of previous experiences in the workplace). It would be unfortunate to forgo these analyses simply because they imply a heterogeneity of the cohort under investigation.

Consider now the condition of homogeneity. It is obvious from the sheer diversity of individuals that no human group will be homogeneous (Henry, 1959). One way of dealing with this heterogeneity is to break down the population being studied into subgroups, so that the population in each can be considered as homogeneous. The resulting multistate table will have an ever-larger number of cells, each containing fewer individuals. Large random errors result, as was seen in Chapter 8, and we can never be certain that all the factors of heterogeneity have been taken into account. Decomposition into subgroups is therefore an unsatisfactory solution and another way must be found to incorporate heterogeneity.

A further restriction is that for a population to remain homogeneous *over time*, the assumption must be made that new members immediately acquire the same characteristics and adopt the same behavior as existing members. Concomitantly, the same applies to those who move out of this subpopulation and join another, whose behavior they are assumed to adopt, while losing all recollection of their past conditions. This amounts to making a Markovian assumption, whereby experience of the present state is totally independent of the individuals' history. Also, the individuals remaining in the subpopulation are assumed to have a behavior that evolves in the same way over time. This assumption again bears little relation to what is actually observed, since real-life individuals obviously do remember their past and this does influence their future behavior.

1. Changing Paradigm

These problems highlight the need for a new paradigm that will enable us to analyze interdependent processes affecting heterogeneous populations. This requires using a different unit of analysis—instead of the single isolated event that was considered previously, analysis now focuses on the entire individual life course, which is treated as a complex stochastic process. The following postulate can then be stated: The trajectory an individual follows over his or her lifetime moves through a succession of different states, and the position occupied at any given moment depends on the individual's life course to date, all that he or she has experienced or acquired, contextual constraints, and the working of personal choice (Courgeau and Lelièvre, 1996). With this change in perspective, the bases of event history analysis can be formulated in terms of individual processes.

The first point to note is that losses from observation (also known as *censored information*) are no longer problematic because the date of the survey or—if population registers are being used—of the study, is not linked in any way to the individual's life. The condition of independence between loss from observation and the events being studied is thus satisfied. The sampling is said to be noninformative, and these censored observations can be taken into account when estimating the hazard rates (Courgeau and Lelièvre, 1989, 1992, 2001). Selection bias can arise because only the individuals who survive and who are present when a retrospective survey is conducted can be questioned. Such bias can be measured using data from prospective surveys or population registers and varies depending on the population being studied. It should be adjusted for, whenever possible (Hoem, 1985). However, it can be assumed to be of limited importance if the event being studied does not occur in an elderly population or one subject to high levels of emigration.

This approach makes it possible to study the dependence between the processes under study and the other phenomena, which are now considered not as interfering events but as interactive events. In the next chapter we will see how this can be achieved by developing multistate event history models. When working on competing events, we can use cumulative transition intensities (sum over time of the hazard rates of experiencing the events being studied), which, contrary to survival probabilities, are suitable for valid comparison. They can be analyzed as approximately independent processes when the actual risks are dependent (Aalen, 1976).

Last, the heterogeneity of individuals must be introduced to improve our understanding of their behavior. Here we assume that behavior patterns are not innate but change over the life course in response to all that the person experiences and acquires with time. It is clear that traditional longitudinal analysis, whereby the initial population is broken down into more homogeneous subpopulations, offers no solution to this problem because it fails to allow for continuous change in this accumulated knowledge and experience. By contrast, the introduction of methods that generalize regression techniques offers a more satisfactory way of including these characteristics, which can then be considered in their time-dependent dimension.

Describing the full complexity of these behavior patterns calls for detailed observation of the different phenomena at work over the life of the individual. Traditional observation using data from civil registration, census, and survey sources is not adequate for this task; more sophisticated data-gathering tools are required to capture the multiple and interlinked chronologies of the different processes. Of these instruments, we prefer the life history survey, which assembles data on the events or *milestones* in the family, occupational, mobility histories that make up a person's life. These critical transitions enable us to identify the dynamic of the interacting processes through comparison of the various histories. Thus, a simple birth history does not constitute a life history account, because it does not include the other events that are related to it. In the same manner, a record of only the *high points* in an individual trajectory, even when it extends over several life domains, cannot form a basis for an event history analysis because it fails to observe the temporal continuity of each process.

Although realistic in scope, a data-collection operation that enables the subject to situate his or her history in the broader temporal context enhances reporting quality as regards consistency between the different calendars, by favoring a better general recall of events in different domains relative to each other (see Groupe R. A. B., 1999).

The event history techniques for handling the data raise highly complex statistical problems that have been overcome largely through advances in the theory of martingales, stochastic integration and counting processes (Anderson *et al*., 1993). A detailed exposition of these methods is not possible here, but after outlining the basic principles and main results, we will show examples illustrating their application in different areas of demographic research.

II. STATISTICAL FORMALIZATION

An initial formalization, referred to as *nonparametric*, enables us to illustrate the interactions between phenomena. Imagine a population comprising n groups between which transitions occur over time. Let $T_{i,j}$ be the positive random variables corresponding to the times of these transitions. If we have $P_i(t)$ individuals in state i just before state t and if $M_{i,j}(t)$ transitions are observed between states i and j at time t, the instantaneous transition rate between these states can be defined in the following form (Johansen, 1983):

$$h_{i,j}(t) = \lim_{dt \to 0} \frac{P(T_{i,j} < t + dt \mid T_{i,j} \geq t)}{dt} = \frac{M_{i,j}(t)}{P_i(t)}$$

The corresponding cumulative transition intensity is then written:

$$H_{i,j}(t) = \int_0^t h_{i,j}(u)du = \sum_{\theta} \frac{M_{i,j}(\theta)}{P_i(\theta)}$$

These transition intensities (or hazards) can be estimated from observations and compared with each other.

Consider the relationship between family formation and urbanization (Courgeau 1987, 1989). The fertility of French women born between 1911 and 1935 is compared according to their residential migrations between metropolitan areas (Paris, Lyon, and Marseille) and less urbanized nonmetropolitan areas. After second births, migration to the metropolitan areas is accompanied by a sharp reduction in fertility—approximately one third—whereas out-migration from the metropolitan areas increases it to the same extent, multiplying the cumulative intensity by 1.4 or even 2. However, migration to metropolitan areas is by women whose fertility before the move was already as low as that prevailing in the urban areas, resulting in a *selection* process. Conversely, migration to the less urbanized regions attracts women whose fertility before moving was the same as that of other women in the metropolitan areas who did not move. In this instance the reduction of the migrants' fertility leads to a process of behavioral *adaptation*.

An additional finding is that the birth of successive children influences women's residential migration, which combined with the previous results produces a *reciprocal dependence*. The propensity to move to a metropolitan area is reduced after each successive birth, whereas the same births have the opposite effect on the migration risks of those leaving metropolitan areas.

Complex dependencies can be identified using this analysis. For example, one process may be observed to influence another whereas there is no effect operating in the other direction. In this case we speak of *unilateral dependence* or *local dependence*. For instance, if we observe the interactions between marriage and departure from farming of men originally employed in this sector, leaving farming is seen to produce a sharp rise in their chances of getting married, whereas marrying while still in farming does not affect their chances of changing employment. We might also observe a *total independence* between these events, in the case that neither has any influence on the other. However, this case is seldom encountered in practice, thus illustrating the need for an event history analysis.

Analysis of this type can be applied to more complex situations, involving trivariate and multivariate relationships, for example. Thus we could study the interaction between fertility and separation in different types of union (cohabitation or marriage) or the interaction between two events with multiple occurrences (fertility and residential migrations).

To develop the analysis further and examine the effect of multiple characteristics on more than one process, we have to adopt more restrictive assumptions about the behavior being studied, by use of parametric or semiparametric proportional hazard models, for example.

Consider an individual k who makes a transition from state i to state j at time t. The instantaneous hazard is expressed in the following general form, which is a generalization of the Cox model (1972):

$$h_{i,j}(t/Z_k(t)) = h^0_{i,j}(t)\exp(\beta_{i,j}Z_k(t))$$

where the first term is the instantaneous transition rate from state i to state j, at time t, under the standard conditions, $Z_k(t) = 0$. The second term of the equation incorporates the individual's characteristics, which may be time dependent, in the form of a column vector $Z_k(t)$ and a vector of $\beta_{i,j}$ parameters. When the individual characteristics are binary, this exponential expresses their proportional effect on the baseline hazard.

When $h^0_{i,j}(t)$ is considered as a function whose form is fully specified in advance, the model is described as *parametric*. Many possible functions can be used for this purpose (Pareto, Gompertz, Weibull, log-normal, etc.) depending on the phenomenon being investigated. Their parameters and the values of $\beta_{i,j}$ can be estimated by the maximum likelihood method. However, this formulation implies the selection of a function, which is not always fully adapted to the phenomenon being studied.

For this reason it is preferable to leave this baseline hazard in a non parametric form. The result is a *semiparametric* model (Cox, 1972), whose parameters $\beta_{i,j}$ can be estimated using partial likelihood methods, which can then be used to estimate the baseline hazard rates $h^0_{i,j}(t)$.

For example, looking again at third births and the influence on them of residential migration to or from metropolitan areas, the effect of many different characteristics can be demonstrated. Thus the fact of having a farming father influences this birth: before migration the effect is to delay third births while after migration the effect disappears. Other variables are not affected by this migration. For example, the more siblings a women has, the more likely she is to have a third child herself, irrespective of whether she lives in a metropolitan area or not, or has moved or not. When all these effects have been incorporated, moving to metropolitan areas remains associated with a reduced chance of having a third child, whereas moving out of metropolitan areas continues to be associated with a raised fertility, thus confirming the persistence of what was observed when the characteristics were not introduced.

Because, as was seen in Chapter 21, we can never be sure that all the variables with an effect on the phenomena of interest have been introduced, it is important to know how this unobserved heterogeneity affects the estimated risks.

If this unobserved heterogeneity is independent of the characteristics introduced, Bretagnole and Huber-Carol (1988) have established that when a proportional hazards model is used the result is to underestimate the absolute value of the effect of these characteristics, whereas in a linear regression it has no effect on these parameters. This means that an effect that appeared significant will remain so when the initially nonobserved characteristics are introduced. Conversely, an effect which initially appeared not significant, may become significant when the previously unobserved characteristics are introduced.

When the explanatory factors, observed or not, are correlated with each other, the choice of characteristics to introduce into the model is of crucial importance. This can be illustrated by an example. When studying duration of residence at a particular dwelling, age is usually taken to be a characteristic strongly correlated with the risk of moving. But it is reasonable to think that this characteristic is itself linked to other characteristics with a more fundamental role in the migration process, such as the stage of the individual's family life at the start of residence, the tenure status of the present dwelling, and so on. Many such characteristics can be introduced, after age, using data from the *Triple Biography* survey conducted by INED (Courgeau, 1985). Not only does this greatly improve the performance of the model for explaining these durations of residence, but more important, the age effect decreases sharply and, for some of the cohorts observed, even disappears completely. Age thus appears merely a proxy variable, attributable to the fact that it is strongly correlated with the onset of critical transitions in the life course.

The proportional hazards model is merely one of several alternative models that can be used. Others include the linear and logistic models, models introducing time-varying estimated $\beta_{i,j}$ parameters and so forth (Gill, 1992). Some of these models present the advantage of dispensing with some of the earlier restrictions relating to unobserved heterogeneity.

In the case of a proportional hazards model it is important to verify with a nonparametric model whether the effect of the various risks considered separately does in fact have a multiplicative effect on the hazard rates.

III. APPLICATIONS OF THESE METHODS IN VARIOUS RESEARCH FIELDS

These techniques have been applied in many research fields, ranging from econometrics to biostatistics and epidemiology. In this section, we give an overview of their demographic applications in several demographic areas. The studies are categorized by their primary field of application, although it should be emphasized that their methodology addresses the interaction between several fields. Attention will be drawn to the originality of the results obtained with these methods.

1. Mortality

Mortality analysis was an important field of application in the early development of these methods, hence the name *survival analysis*, by which they are sometimes known. Since the original article by Cox (1972), the spread of these techniques in medical research has been reviewed (see, for example, Andersen, 1991). But their use outside of clinical trials and medical epidemiology has not developed to the same extent, particularly in demography. One reason for this concerns the type of data needed for their implementation. But we begin by looking at two examples of studies based on survey data.

Use of proportional hazards models to analyze the effects of childbearing on the occurrence of Hodgkin's disease (Kravdal and Hansen, 1996) is a perfect illustration of the most important use of this method for prognostic purposes. An earlier article by these authors, using a more conventional methodology, concluded that childbearing conferred a protective effect against Hodgkin disease. However, that protective effect is called into question in this article (Kravdal and Hansen, 1996) by the more relevant results obtained using survival models.

Guan Guo and Lawrence Grummer-Strawn (1993) analyze twin survival using data on multiple births from 26 Demographic and Health Surveys (DHS). This identifies the factors of vulnerability for twins and indicates a general increase in the effects of the variables measured for twins compared with those for other children. The model used is derived from the bivariate model and takes into account the difference between the survival chances of the two categories of children. This is a major advantage of these models.

In contrast to traditional demographic mortality analysis, the two previous studies do not use civil registration data. That source supplies too few individual characteristics for it to be suitable for in-depth event history analysis. A linkage is sometimes possible with census data, which provide more detailed information. Thus several authors (Sahli et al., 1995; Wunsch *et al*., 1996) have studied mortality in relation to characteristics from three successive censuses linked with Norway's register of deaths. These analyses take the form of a series of logistic regressions on time-ordered sequences, without introducing duration. This procedure, while entirely legitimate and effective, differs from the event history family of models by not including duration, which is a fundamental dimension of these models. However, it will be seen shortly that duration can be introduced for the analysis of interval-censored life history information of this kind.

2. Fertility

Fertility analysis has produced numerous surveys suitable for event history analysis. In a review article, Adrian Raftery *et al*. (1996) give a critical appraisal of the methods used to analyze the World Fertility Survey (WFS) data. With the exception of descriptive analyses, using conventional tabular presentations of fertility rates and linear regression analyses in which the explanatory variable is the number of live births, the authors indicate that although these data give detailed, reproductive histories combined with calendars of contraceptive use they have seldom been subject to event history analysis. After outlining the practical difficulties involved in conducting such an analysis and the innovative nature of the indicators obtained, they present an analytical strategy for event history modelling of WFS data, with a view to increase the use of this strategy. Using the number of live births as a regression variable encounters a certain number of problems, which can only be resolved within the much more appropriate framework of event history analysis. The first problem concerns incomplete information due to censoring at the time of the survey, something that cannot be allowed for in a conventional regression analysis. The second problem relates to the impossibility of distinguishing cohort and period effects. Finally, if the influence of explanatory variables varies with parity, these effects are *diluted* in a conventional regression because individual variables cannot be modified according to exposure duration (i.e., cannot be time-dependent). The authors then apply the proposed procedure to WFS survey data for Iran, achieving results that could not be obtained with the other procedures.

In an article on extramarital fertility, Eva Lelièvre (1993) reviews the methodologic problems encountered in such a study and how they can be overcome by means of event history analysis. Particular attention is given to the question of the validity of analyses where the results are conditioned by knowledge of the future (i.e., by the outcome of the process being studied). In the case of extramarital fertility, the processing of the period of cohabitation is determinant. If, for the evaluation of fertility differentials, unions are distinguished by whether they lead to a marriage or not, the rates are in effect conditioned relative to the future. These difficulties are avoided with an event history analysis that allows for past events and introduces union transformations at the time they occur.

3. Union Formation and Dissolution

Many analysts have come to favor event history analysis for the study of divorce, because the introduction of time-dependent variables is necessary to grasp the nature of the mechanisms at work. This also makes it easier to distinguish between period, age, and cohort effects.

Michael Bracher *et al*. (1993) have carried out a systematic examination of hypotheses advanced to explain the risk of marriage dissolution in Australia, using a Cox model. In constructing their model, they consider successively the introduction of historic time (characterized by period and cohort aspects) and biographic time (age at marriage, age and duration of marital union), the relevant characteristics. It is clear that only event history methods (and the data for running them) are suitable for conducting such an analysis, which is applied to empirical data.

Working on the plentiful and very precise data from the Swedish Population Register, Hoem (1997) has called into question the effects of educational attainment on divorce and has shown (as is only possible with event history models) that these effects are time-variant. These findings run counter to the generally accepted hypothesis (in the English-speaking and Scandinavian countries) of a decrease in the risk of divorce as educational attainment increases. This hypothesis was already at odds with the opposite or nonexistent effect of educational attainment observed in other countries, and thus appears very questionable.

In his study on the place of children in the history of couples, Laurent Toulemon (1994), responding to an earlier critical comment by Adrian Raftery *et al.* (1996), proposes simple indicators, drawn from event history analysis, to present the variations in the risk of marriage breakdown by number of children and age of last child. This study concludes that increased risks of marri-age breakdown are partly independent of fertility behavior.

4. Internal Migration

The French *Echantillon Démographique Permanent* (Demographic Longitudinal Survey), which links civil registration and census data since 1968, makes possible a new approach to interval-censored event histories. In this data file, the residence of individuals is recorded only at the time of censuses and family events. The dates observed are therefore those before and after the migration event to be studied.

Two simplifying assumptions have to be made to use these interval-censored data. The first is that only one of the migrations being studied can occur between two observation times. The second is that the observation scheme is non informative, i.e. that observation of individual trajectories occurs independently of their progression, at random points in time. The first of these conditions determines the choice of observation, using both census and civil registration data, to obtain the optimum estimation. For the second condition, it can be seen that while it is fully satisfied by census observation, civil registration observation can be a source of bias, since there is no reason to expect migrations to be independent of family events.

The test conducted with data from the Triple Biography survey, which were artificially censored for this purpose (Courgeau and Najim, 1995, 1996), established that bias was not large when working on changes of residence. The estimates are similar with census or civil registration data, or with both together, although when civil registration data alone are used significant overestimation of the hazards occurs related to the interactions between migration and fertility. On the other hand, when working on changes of region, serious errors arise due to return migration. A semiparametric analysis of residential migration after marriage has been conducted using the EDP data (Courgeau *et al*., 1998).

5. Occupational Histories

In the field of the study of individuals' occupational histories, Daniel Courgeau (1993) has shown how event history analysis can answer a perennial research question. The problem is that of how do we *allow for the interactions between occupational and migratory events, by introducing a continuous time during which these events occur?* Whereas traditional methods cannot deal with this problem, the general framework of event history analysis is effective for identifying the nature of the interactions in question and the heterogeneity of the population under consideration.

This general framework is also effective for handling far more complex problems, such as the interaction between the professional and migration behavior of the members of a couple (Courgeau and Meron, 1995). These interactions can be modelled, taking into account the effect of the couples' characteristics, which can be time-dependent, to obtain information about the close links between its residential mobility, occupation, and family life. In this way we are able to relate the three main role domains in the life of a couple.

This brief review shows the scope of application of event history analysis and indicated its value as an explanatory tool in demography.

IV. FUTURE DEVELOPMENTS

The event history approach is now widely accepted in the social sciences, including in demography (Keilman, 1993). We must nonetheless resist the temptation to explain a behavior by reference only to individual-level characteristics. The danger then is of committing the atomistic fallacy, because the analysis ignores the broader context in which human behavior occurs. Individuals actually operate in multiple domains (family and occupational, for example) and are linked to a geographic context (resident of a neighborhood, originating from a region, etc.), all of which influence their behavior.

On the other hand, when working at a given level of aggregation, only the conventional demographic rates can be related to the aggregate-level characteristics of the regions being considered, such as the regions' emigration rates to their percentages of farmers, labor force participation rates, and so forth. The danger here is of committing what is known as the *ecological fallacy*, which occurs when inferences about individual-level behavior are drawn from aggregate-level results.

Thus, we must find a way of integrating these different levels of aggregation in the same event history analysis. An initial approach is to examine the effect of both individual and aggregate characteristics on individual behavior, with the possibility of introducing a specific effect for each region at any given level of aggregation. This leads to contextual and multilevel event history models (see Chapter 24).

A second approach involves the analysis of more complex groups, when it is no longer possible to consider the group as forming a single entity but as composed of interacting individuals. This means that instead of studying the interactions between the processes that affect an individual, we study the interactions between individual processes affecting members of a group (Lelièvre *et al*., 1997, 1998). This shift from individuals to their entourage (or *contact circle*), whether for the purpose of data gathering or analysis, implies rethinking the units used for longitudinal observation.

Whereas an individual can be tracked over his or her entire lifetime, the practical impossibility of defining a household in longitudinal terms (Keilman and Keyfitz, 1988) has led to development of the concept of entourage or *contact circle* (Lelièvre and Bonvalet, 1994). It is still the individual who is tracked, but linked with all the members of the various households to which he or she has ever belonged, plus other close but not necessarily co-resident family members.

One approach involves an event history analysis that treats such groups as a composite individual who can be described by both collective and individual characteristics. But the events studied will be mainly collective in nature (change of membership, place of residence, etc.), whereas individual characteristics and the interactions between members are considered only implicitly. In these conditions, the techniques of event history analysis of the group considered as an individual can indeed be applied, on condition the definition of the group and the events it experiences are carefully specified.

Modelling on these lines is not perfect, however, because it overlooks the fact that the process being studied is multivariate, with each of the group trajectories resulting from the interacting trajectories of its members. Since the life histories of the individual members of the group will no longer be independent, a specific modelling strategy has to be developed. This problem has been explored from a statistical angle and these models are currently being adapted for application to demographic processes (Lelièvre *et al*., 1997, 1998).

From the standpoint of data collection strategies, all the events of union and reproductive histories are already included in an individual life history survey, and if the varying position of the individual within his or her contact circle is defined using these dimensions, no additional information is required. If data collection also includes the residential histories of children/parents (direct descent) when the respondent does not live with them, information on the role of kinship in individual decisions becomes available. By working on factual data, such as the location or occupation of a certain number of close friends and relatives, the existence of an influence can be established when there is residential or social proximity. Last, the structure of the individual's network, that is, the nature of his or her contact circle will be known for the entire life course of the individual. A survey along those lines, the *Biographies et entourage* (Event Histories and Contact Circle) survey, was conducted by the Institut national d'études démographiques at the end of 2000. Analysis of the results from this survey will represent a major advance for event history analysis.

References

AALEN Odd O., 1976. Nonparametric inference in connection with multiple decrement models, *Scandinavian Journal of Statistics*, no. 3, p. 15–27.

ANDERSEN Per Kragh, BORGAN Ørnulf, GILL Richard, KEIDING, 1993. *Statistical Models Based on Counting Processes*. New York, Springer Verlag, 768 p.

BRETAGNOLE Jean and HUBER-CAROL Catherine, 1988. Effects of omitting covariates in Cox's model for survival data, *Scandinavian Journal of Statistics*, n° 15, p. 125–138.

COURGEAU Daniel and LELIEVRE Éva, 1986.—Nuptialité et Agriculture, *Population*, vol. 41, n° 2, p. 303–326.

COURGEAU Daniel and LELIEVRE Éva, 1989.—*Analyse démographique des biographies*, Paris, coll. de l'INED, PUF, 268 p. (Also in English: *Event history analysis in demography*, Oxford, Clarendon Press, 1992, 226 p., and in Spanish: *Análisis demográfico de las biografías*, México, El Colegio deMéxico, 2001, 306 p).

COURGEAU Daniel and LELIEVRE Éva, 1996. Changement de paradigme en démographie, *Population*, n° 3, p. 645–654. (Also in English 1997. Changing paradigm in demography, *Population: English Selection*, vol 9, p. 1–10).

COURGEAU Daniel, 1985. Interaction between spatial mobility, family and career life-cycle: a French survey, *European Sociological Review*, vol. 1(2), p. 139–162.

COURGEAU Daniel, 1987. Constitution de la famille et urbanisation, *Population*, vol. 42(1), p. 57–82. (Also in English 1989. Family formation and urbanization, *Population: English Selection*, vol. 1, p. 123–146).

COX David R., 1972. Regression models and life tables (with discussion), *Journal of Royal Statistical Society*, vol. B34, p. 187–220.

GILL Richard, 1992. Multistate life-tables and regression models, *Mathematical Population Studies*, no. 3, p. 259–276.

Groupe de réflexion sur l'approche biographique, 1999. *Biographies d'enquêtes. Bilan de 14 collectes biographiques*, Paris, INED/PUF, 340 p. (Coll. Méthodes et savoir, no. 3)

HENRY Louis, 1959. D'un problème fondamental de l'analyse démographique. *Population*, vol. 14(1), p. 9–32.

HOEM Jan, 1985. Weighting, misclassification and other issues in the analysis of survey samples of life histories, *in:* James Heckman and Burton Singer (eds.), *Longitudinal analysis of labour market data*, p. 249–293. Cambridge, Cambridge University Press, 425 p.

KALBFLEISCH John D., PRENTICE Ross L., 1980. *The statistical analysis of failure time data*, New York, John Wiley & Sons, 322 p.

KEILMAN Nico and KEYFITZ Nathan, 1988. Reccurent issues in dynamic household modelling, *in*: Nico Keilman, Anton Kuijsten and Ad Vossen (eds.), *Modelling household formation and dissolution*, p. 254–285. Oxford, Clarendon Press, 298 p.

LELIÈVRE Èva, BONVALET Catherine and BRY Xavier, 1997. Analyse biographique des groupes, les avancées d'une recherche en cours, *in:* Daniel Courgeau (ed.), Nouvelles approches méthodologiques en démographie, *Population*, vol. 52(4), p. 803–830. (Also in English, 1998: Event history analysis of groups: the findings of an on-going research project, *Population: English Selection*, vol 10(1), p. 11–37).

LELIÈVRE Èva, BONVALET Catherine, 1994. Construire les principes d'une analyse biographique du groupe familial, in *Ménages, familles, parentèles et solidarités dans les populations méditerranéennes*, Actes du séminaire international de l'AIDELF (1996), p. 103–116.

LELIÈVRE Èva, BRINGE Arnaud, 1999. *Manuel pratique pour l'analyse statistique des biographies/Practical guide to event history analysis using SAS, TDA, STATA*, Paris, INED/PUF, 378 p.

Mortality

ANDERSEN Per Kragh, 1991. Survival analysis 1982–1991: the second decade of the proportional hazards regression model, *Statistics in Medicine*, vol. 10, p. 1931–1941.

GUO Guan and GRUMMER-STRAWN Lawrence, 1993. Child mortality among twins in less developed countries, *Population Studies*, vol. 47, p. 495–510.

KRAVDAL Øystein and HANSEN Svein, 1996. The importance of childbearing for Hodgkin's disease: new evidence from incidence and mortality models, *International Journal of Epidemiology*, vol. 25, p. 737–743.

LOPEZ Alan, CASELLI Graziella and VALKONEN Tapani (eds.), 1995. *Adult mortality in developed countries.* Oxford, Clarendon Press, 361 p.

SALHI Mohamed, CASELLI Graziella, DUCHÊNE Josianne, EGIDI Viviana, SANTINI Antonio, THILTGÈS Évelyne and WUNSCH Guillaume, 1995. Assessing mortality differentials using life history: a method and applications, *in:* Alan Lopez, Graziella Caselli and Tapani Valkonen (eds.), *Adult mortality in developed countries: from description to explanation*, *p. 57–79*. Oxford, Clarendon Press, 361 p. (Coll. International Studies in Demography).

WUNSCH Guillaume, DUCHÊNE Josianne, THILTGES Évelyne and SALHI Mohamed, 1996. Socio-economic differences in mortality. A life course approach, *European Journal of Population*, vol. 12(2), p. 167–185.

Union Formation and Dissolution

BRACHER Michael, SANTOW Gigi, MORGAN Philip S. and TRUSSELL James, 1993. Marriage dissolution in Australia: Models and Explanations, *Population Studies*, vol. 47(3), p. 403–425.

HOEM Jan, 1997. Educational gradients in divorce risks in Sweden in recent decades, *Population Studies*, vol. 51, p. 19–27.

LERIDON Henri, 1990. Cohabitation, marriage, separation: an analysis of life histories in French cohorts from 1968 to 1985, *Population Studies*, vol. 44, p. 127–144.

TOULEMON Laurent, 1994.—La place des enfants dans l'histoire des couples, *Population*, vol. 6, p. 1321–1346.

Fertility

LELIÈVRE Eva, 1993. Extra-marital births occurring in cohabiting unions, *in:* Maire Nì Bhrolcháin (ed.), *New perspectives on fertility in Britain*, p. 109–120. Londres, OPCS, 169 p. (Coll. Studies on Medical and Population subjects no. 55).

KRAVDAL O and HANSEN S. (1993). Hodgkin's disease: The protective effect of childbearing. *Inrernational Journal of Cancer*, 55: 56–61.

RAFTERY Adrian, LEWIS Steven, AGHAGANIAN Akbar and KAHN Michael, 1996. Event history modeling of World Fertility Survey Data, *Math. Population Studies*, 2, p. 129–154.

Internal Migration

COURGEAU Daniel, LELIEVRE Eva, WOLBER Odile, 1998.—Reconstruire des trajectoires de mobilité résidentielle. Éléments d'une analyse biographique de l'EDP, *Économie et Statistique*, no. 316–317, p. 163–176.

COURGEAU Daniel, NAJIM Jamal, 1995. Analyse des biographies fragmentaires, vol. 50(1), p. 149–168. (Also in English 1996. Interval-censored event history analysis, *Population: An English Selection*, vol. 8, p. 191–208).

Work Careers

COURGEAU Daniel and MERON Monique, 1995. Mobilité résidentielle, activité et vie familiale des couples, *Economie et Statistique*, no. 290, p. 17–31.

COURGEAU Daniel, 1993. Nouvelle approche statistique des liens entre mobilité du travail et mobilité géographique, *Revue Economique*, vol. 44(4), p. 791–807.