Introduction: Longitudinal Analysis of Historical-Demographic Data  The field of historical demography embraces the aspirations and challenges of both demography and history. It analyzes the vital processes that unfold within individual life courses—particularly fertility, nuptiality, migration, and mortality—while tracking aggregate changes in those processes over time. It identifies the social, cultural, economic, and political correlates of individual demographic outcomes and historical-

George C. Alter is Professor of History and Research Professor, Population Studies Center, University of Michigan; Director, Inter-university Consortium for Political and Social Research; and Director, Population Institute for Research and Training, Indiana University. He is the author of, with Muriel Neven and Michel Oris, “Economic Change and Differential Fertility in Rural Eastern Belgium, 1812 to 1875,” in Noriko Tsuya et al. (ed.), Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700–1900 (Cambridge, Mass., 2010), 195–216; with Isabelle Devos and Alison Kvetko, “Completing Life Histories with Imputed Exit Dates: A Method for Historical Data from Passive Registration Systems,” Population, LXIV (2009), 293–318.

Myron P. Gutmann is Professor of History, and Information and Research Professor of Population Studies, University of Michigan; Assistant Director, National Science Foundation. He is the author of Towards the Modern Economy: Early Industry in Europe, 1500–1800 (New York, 1988); editor of, with Glenn D. Deane, Emily R. Merchant, and Kenneth M. Sylvester, Navigating Time and Space in Population Studies (New York, 2011).


The work reported in these articles was supported by grant Number R25 HD040525 from The Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). Myron P. Gutmann’s work was supported, in part, by the National Science Foundation of the United States. Any opinions, findings, conclusions, or recommendations expressed in this material belong to the authors exclusively and do not necessarily reflect the views of the National Science Foundation or the NICHD.

© 2012 by the Massachusetts Institute of Technology and The Journal of Interdisciplinary History, Inc.
demographic transitions, seeking explanations that are universal and particular, both macro and micro. For the last fifty years, these pursuits have promoted and been facilitated by a burgeoning of individual life-history datasets covering broad spans of time and many parts of the globe. This valuable resource has allowed researchers to examine demographic processes within lives and across space and time. More recent developments in database software and statistical methods have promoted increasingly sophisticated analysis of these datasets, including linkage of families across generations and identification of complex causal mechanisms. The availability of these datasets and analytical capabilities has encouraged researchers to ask new questions and re-open old controversies about core demographic processes and historical transitions in them, while scholars in fields ranging from anthropology to epidemiology are extending the techniques of historical demography to ever more populations and research topics.

This special issue of the Journal of Interdisciplinary History presents six examples of recent forays into the longitudinal analysis of historical-demographic data, written by participants of summer workshops held in 2006 and 2007 at the Inter-university Consortium for Political and Social Research (ICPSR), the University of Michigan. These studies focus on Utah (United States), Sart (Belgium), Bologna (Italy), and Madrid (Spain) in the nineteenth and early twentieth centuries, attending to the particularities of these settings and discussing the wider applicability of their findings.


2 This intensive four-week course introduces current and potential users of the growing number of individual-level historical-longitudinal databases to the theoretical and substantive demographic questions that can be approached with this type of data; the specialized techniques used to create these datasets from surviving administrative, genealogical, and ecclesiastical records; and the subtle challenges that historical data present to statistical analysis. Participants approach important issues in historical demography by developing research projects that use multivariate longitudinal analysis to exploit the rich capabilities of large individual-level life-history datasets from Europe, Asia, and North America. It began in 2006 and has been held biennially since 2007.
standing goals of historical demography—understanding and accounting for the demographic transition that occurred in Europe and North America between the mid-eighteenth and mid-twentieth centuries. As a description of the historical experience of Europe and North America, demographic transition, in its broadest terms, refers to a secular decline in mortality, accompanied and followed by a decline in fertility. The result includes a shift from a high-pressure equilibrium with a younger age structure to a low-pressure equilibrium with an older age structure, generating dramatic but temporary population growth. Analyses of cross-sectional aggregate historical data reveal changes in mortality and fertility rates over time and identify the structural correlates of varying levels of mortality and fertility. More recent multivariate longitudinal analyses of individual-level life-history data have added considerable nuance to our understanding of pre- and posttransitional demographic regimes and of the transition itself by revealing variations in the experience of transition and assessing the individual- and household-level determinants of such demographic processes as migration, marriage, childbirth, and death.  

Historical demographers have used longitudinal data and methods to examine the unevenness of demographic transition within societies, asking, for example, which families continued to experience infant death as mortality declined at the societal level, what characteristics made some couples more likely to limit their fertility in advance of widespread fertility decline, and how mortality and fertility were related at the household level. Scholars have also utilized longitudinal fertility data to contest and complicate Henry’s early description of the fertility transition as a shift from “natural fertility,” non-parity-specific fertility behavior, to “controlled fertility,” the cessation of childbearing when an ideal number of offspring is achieved. Drawing on birth-interval data, recent studies have accounted for differences in fertility levels among natural-fertility populations, identified deliberate fertility

---

control prior to fertility transition, examined the role of increased birth spacing in effecting fertility decline, and analyzed changes in the determinants of fertility between the pre- and posttransitional periods.  

But change in number of children born is not the whole story: Mason has argued that the fertility transition also included a shift from postnatal to prenatal methods of controlling family size and composition. Historical demographers have tested this assertion by using life-history data to explore the dynamics of infanticide and child abandonment, identifying which children were most at risk and examining how the presence of other household members increased or decreased their vulnerability. Such work has demonstrated that the demographic transition was multiple rather than singular: Demographic regimes and changes in them varied not only by region but also by class and even gender within regions. The articles that follow this introduction use life-history data and longitudinal methods to address many of these questions, demonstrating the subtle concerns that these data and methods are particularly suited to answering and suggesting the wealth of knowledge yet to be revealed through the rigorous analysis of historical sources.  


5 On the shift from postnatal to prenatal fertility control, see Karen Oppenheim Mason, “Explaining Fertility Transitions,” Demography, XXXIV (1997), 443–454. For analysis of postnatal control in pre-transition populations, see Tsuya et al., Prudence and Pressure. The
HISTORICAL DATA  Life-history databases provide researchers with data that literally took lifetimes and generations to collect. Most are compiled from historical sources originally created for other purposes—typically ecclesiastical, administrative, and genealogical—often using such specialized techniques as family reconstitution and back projection. Analyzing these sources therefore poses unique challenges and requires detailed knowledge about the past societies and institutions that produced them and the methods used to transform them into individual-level longitudinal databases. The comprehensiveness and quality of historical data depend upon both the administrative capabilities of the institutions that produced the original documents and the goals and needs motivating them to collect information. The nature of the data, in turn, determines how researchers can use them. For example, databases produced through the reconstitution of families from records of baptisms, marriages, and burials provide precise dates of vital events but no observation between events; those produced through the linkage of frequent censuses offer more observations of individuals but not exact dates of vital events. Continuous registers of population, kept in several European and Asian localities, typically combine the strengths of vital registers and frequent censuses but are rarely available for the long swaths of time necessary to analyze historical transitions and intergenerational processes.  

Nonetheless, if used with the appropriate caution, each type of database can yield important information about historical-demographic regimes that both challenge and refine demographic-transition theory. Moreover, as more and more such databases have become available, researchers have been able to supple-

---

ment knowledge gained from one type of source with information from other types, and to draw cross-cultural comparisons by analyzing analogous data from different societies.

**LONGITUDINAL ANALYSIS** Life-history datasets include continuously or periodically updated information for each individual over time, and therefore require specifically longitudinal analytical methods. Each article in this issue employs one of two multivariate longitudinal regression models: Cox proportional hazards and discrete-time event history. These methods, which originated in engineering and biomedical research in the 1960s and 1970s, were adopted by social scientists for use with duration or time-to-event data in the early 1980s. The dependent variable is whether or when a particular event occurs, such as childbirth or death; results indicate the effects of independent variables on the instantaneous probability of experiencing the event in question (the hazard rate) across the period of analysis. The baseline hazard reflects the hazard rate for an individual possessing values of 0 for all of the independent variables as a function of the time at risk for the event in question (analysis time). Although values of independent variables can change during the period of analysis, their effects are assumed to be constant and in proportion to the baseline hazard.\(^7\)

The shape of the hazard function is not specified in advance when using the Cox model, allowing the baseline hazard to vary continuously. Cox models are therefore well suited to datasets based on continuous observations in which the exact dates of vital events are known, such as those constructed from family reconstitution or population registers. These models are less appropriate for datasets compiled from periodic observations, which can reveal only that an event occurred (or did not occur) between one observation and the next. For this type of data, discrete-time event-history analysis divides the time under study into “spells” or “episodes” equal to the period between observations, with the assumption that the baseline hazard remains constant within each spell. Results of both Cox and discrete-time models indicate

which individual attributes increase or decrease the hazard rate at any given time, and quantify the relative risk associated with that attribute.

Longitudinal analysis establishes causal relationships because the state of attributes is specified prior to the occurrence of the event in question. In the Cox model, the values of covariates that change over time—such as age, occupation, or parity—may be updated continuously; in discrete-time models they are updated at the beginning of each spell. Databases rich with individual attributes that are monitored continuously or frequently can produce fine-grained analyses of the relationships between these attributes and vital events. Statistical software readily handles such analysis, but complex methods of data management are often required to translate data from the relational databases in which they are stored to the rectangular format required by statistical packages. The complexity increases with the number of time-varying covariates, especially when those covariates refer to other individuals in a household. Another advantage and challenge of longitudinal models is their ability to accommodate incomplete records and individuals for whom the event in question does not occur during the period of analysis. Valuable information could be lost if, for example, only individuals who died under observation were included in mortality analysis or if women who died, moved away from the observation area, or were widowed or divorced before reaching the end of childbearing age were excluded from fertility analysis. In longitudinal models, individuals can be included for the length of time during which they are under observation, and excluded, or censored, when observation ends, regardless of whether the event in question has occurred. Use of these observations allows researchers to retain information, maintain sample sizes, and avoid sample-selection bias.

8 The editors of this issue are currently working with representatives of life-history databases throughout the world to develop a standard format for storing these types of data that will allow for the development of a common set of programs to rectangularize data for various forms of demographic analysis. Such standardization will also promote comparative analysis across databases. Alter, Kees Mandemakers, and Gutmann, “Defining and Distributing Longitudinal Historical Data in a General Way through an Intermediate Structure,” Historical Social Research, XXXIV (2009), 78–114. Successful censoring requires knowledge of the date at which observation ended. This date must be determined with care to avoid introducing new biases. Death clearly ends observations and, in data derived from population registers, dates of migration are often known. Determining the end of observation is much more challenging in family-reconstitution datasets, which include only births, marriages, and deaths that occurred.
The articles presented in this issue successfully manage the challenges of longitudinal analysis and fully utilize the capabilities of life-history data to answer important questions about demographic processes and historical change: “Offspring Sex Preference” by Nora Bohnert, Hilde Leikny Jåstad, Jessica Vechbanyongratana, and Evelien Walhout and “Is Sibling Rivalry Fatal?” by Rebecca Kippen and Sarah Walters examine the influence of family composition on fertility and mortality. “Intergenerational Transmission of Reproductive Behavior” by Julia Jennings, Allison Sullivan, and J. David Hacker and “Migrants and Diffusion of Low Marital Fertility” by Mathew Creighton, Christa Matthys, and Luciana Quaranta analyze the intergenerational transmission and spatial diffusion of fertility behavior. “The Determinants of Reproductive Behavior” by Rosella Rettaroli and Francesco Scalone and “Releasing Mother’s Burdens” by Bárbara Revuelta Eugercios explore family formation in times of social and economic transformation.

CONTRIBUTIONS MADE BY THE ARTICLES IN THIS ISSUE The extent to which couples’ desire for male (or female) children led them to adopt fertility control has become an important issue in analysis of the fertility transition, arising first in studies of Asian societies where son preference is well established, and more recently examining possible son preference in agricultural households. Bohnert et al. expand the question to ask whether couples also preferred daughters or a balance of sons and daughters, to examine whether these preferences changed over time, and to assess whether the prevalence of sex-preferential fertility behavior increased with the onset of fertility transition and smaller families. Furthermore, they directly address the question of agriculture and religion as indicators of couples who may have been particularly motivated to produce sons. Both parity-specific control (stopping) and progression to another birth (spacing) are considered. They utilize data from

under observation. A missing date of death might imply migration, but the date of migration is unknown. As a pioneer of this method, Henry specified that an event related to the process in question cannot serve to mark the end of observation. In the case of fertility analysis, if the dates of child birth or child death were used as the end of observation, women with more children would be overrepresented in the analysis. Henry, “Some Data on Natural Fertility”; Gutmann and Alter, “Family Reconstitution as Event-History Analysis,” in Reher and Schofield (eds.), Old and New Methods in Historical Demography (New York, 1993), 159–177.
the Utah Resource for Genetic and Epidemiologic Research (RGE)—compiled by the Utah Genealogical Society in the 1970s from the family histories of the descendants of the Utah pioneers. To create the database, the three-generation family group sheets used by the Church of Jesus Christ of Latter-day Saints (LDS), which closely resemble the family-reconstitution forms developed by historical demographers, were linked into longer genealogies and supplemented with vital records kept by the state of Utah. The time depth of the RGE makes it a valuable source for analyzing change in demographic behavior over time.9

The authors draw on fertility histories of women born between 1850 and 1900 in order to span the pre- and posttransitional periods, dividing their analysis by mothers’ birth cohort. Comparisons of the parity-specific sex ratios of living children between couples who stopped at a given parity and couples who continued to bear children address the extent to which the sex of offspring influenced further childbearing; Cox regressions with independent variables indicating the sex mix of previous children assess the effect of this mix on birth spacing. Results for the first analysis indicate that women who gave birth to a daughter at higher parities were more likely to have another child; results for the second demonstrate that, also at higher parities, the likelihood of a subsequent birth was elevated for women whose previous children were mostly girls and reduced for women whose previous children were mostly boys, both relative to those with a balanced sex mix. Whereas the spacing pattern is evident even in the earliest cohorts, it emerges at lower parities in later cohorts, among whom the stopping pattern is also more pronounced. These findings suggest that the pretransitional preference for sons persisted across the transition, though the means of realizing it changed. Agriculturalists demonstrated a preference for larger families, regardless of sex; LDS members demonstrated a preference for sons similar to that of non-members. Bohnert et al.’s finding of son preference in a North American population breaks new ground.

Kippen and Walters also consider the influence of siblings on demographic outcomes, in this case the survival prospects of children under the age of five. They question whether having siblings near to one’s own age increased the risk of death, whether having much older siblings had a protective effect, and whether some families faced poorer survival prospects for all of their children. The theoretical underpinnings of these questions lie in the competition for resources, maternal depletion, and unobservable biological traits that elevate the risk to related children. Since differences in mortality for neonates, infants, and children are well documented, the authors analyze these age groups separately. Kippen and Walters’ questions are particularly suited to longitudinal analysis of life-history data because the characteristics of the offspring set vary over the course of parents’ marriages and children’s lives. The authors use data from the village of Sart in the Historical Database of the Liège Region—a database created by Alter and Michel Oris from population registers kept by eleven Belgian communes in the second half of the nineteenth century, supplemented with information from civil registers of births, marriages, and deaths. Beginning in 1846, Belgian law required municipalities to record the population by domicile and to update the information continuously in response to births, deaths, marriages, and migration. Sart began its population registration earlier, in 1812, providing an even deeper record of continuous individual-level data covering most of the nineteenth century. Because population registers list individuals as members of households, they allow scholars to analyze the effects of household composition on demographic outcomes.

Using this rich data set, the authors include the number and ages of siblings present in the household as time-varying covariates in a series of Cox proportional-hazards models. Their models also include the proportion of siblings dying in infancy as a way to assess whether infant mortality clustered in households even after controlling for household composition and other potentially confounding factors. Results confirm that for neonates, co-resident siblings under five years of age increased the index child’s mortality risk and that infant mortality clustered in families, even when

---

controlling for the length of time elapsed since the previous sibling was born and the presence of siblings old enough to supplement a mother’s care. Importantly, they also find that children under the age of one did not benefit from Sart’s nineteenth-century mortality transition. This study reveals important determinants of neonatal mortality, but, methodologically, it also points to the importance of segmenting analysis of infant mortality by age. Taken together, this article and that of Bohnert et al. make a compelling argument for the importance of family composition to a proper understanding of fertility and child mortality, while also attending to critical methodological issues.

Another important dimension of family context is the relationship between generations. Accordingly, intergenerational processes have received a great deal of attention recently in historical demography. Jennings et al. extend the “nature versus nurture” debate on the basis of intergenerational fertility transmission, making three important contributions: First, they examine the relationship not only between women’s fertility and that of their mothers but also between women’s fertility and that of their mothers-in-law; second, they test whether the strength of this relationship increased during the fertility transition; and third, they add the vital status of a woman’s mother and mother-in-law as time-varying covariates to assess the role of the extended family in fertility decisions.

The genealogical basis and long historical span of the rge (described above) make it ideal for Jennings et al.’s intergenerational analysis. Results from bivariate correlations and Cox proportional-hazards models demonstrate that correlations between a woman’s fertility and that of both her mother and her mother-in-law emerged at the end of the nineteenth century, providing evidence that the vital status of the mothers of both spouses affected fertility, controlling for age, cohort, and other demographic factors. The authors suggest that, although the behavior of a woman’s own mother had a stronger influence on her fertility, her mother-in-law’s ideals of the number and timing of births may have been transmitted directly or through her son, indicating the importance of husbands in conception and contraception. They emphasize that explanations of intergenerational fertility patterns must go beyond simple biological models of fecundity to include social norms, support mechanisms, and intrafamilial power dynamics.
The conclusions of this article argue that the transmission of new social norms from one generation to the next played an important role in the fertility transition, and particularly in its irreversibility.

Mechanisms for the spatial diffusion of low fertility in the nineteenth century have been more asserted than tested. In addition to the familial pathway explored in Jennings et al., migration is a logical theoretical mechanism for the movement of ideas through space. Although migration within Europe increased in both volume and distance during the same period when fertility declined, the question of whether movers and stayers experienced different fertility trajectories has received little attention in the historical-demographic literature. This lacuna may reflect the difficulty of tracking migration in family reconstitution and genealogical data. Creighton et al. begin to fill this gap through analysis of the Historical Database of the Liège Region (described above), based on records noting migration into and out of the village of Sart between 1812 and 1900.

The authors predict at the outset that because Sart was a rural commune that experienced a relatively late fertility transition, migrants would be found to have had lower fertility than natives. They further suggest that, if this fertility gap were produced by conscious family limitation, it would have opened only in the second half of the nineteenth century, when the fertility transition was underway in the surrounding areas. The authors test these ideas with Cox proportional-hazards models, controlling for a number of other factors that might have affected the likelihood of conception, including husband’s occupation, wife’s age, parity, breast-feeding, and the sex composition of children in the household. Acknowledging the role of both partners in making fertility decisions, the authors assess the effects of both wives’ and husbands’ migration status. Analysis confirms their theory only for dual-migrant couples. Although migrant status had no effect on fertility earlier in the century, after 1850, couples in which both husband and wife had migrated to Sart were less likely to conceive than were couples in which at least one partner was born in Sart. These findings suggest an important relationship between the movement of people and the movement of ideas, and a potentially important role for migration in the European fertility transition.

The study by Creighton et al. points to the geographical unevenness of the fertility transition; that by Rettaroli and Scalone...
documents the socioeconomic unevenness of both pre-transition fertility and fertility decline in the rural hinterland of Bologna. Having reconstituted both individual life histories and household contexts in two parishes of Bologna between 1818 and 1900 from institutional records of the Catholic Church, the authors use discrete-time event-history analysis to assess the relative effects of biodemographic factors, household characteristics, and economic fluctuations on fertility prior to, and in the early years of, Bologna’s fertility transition.

The results of their analysis indicate that such biodemographic variables as mother’s age, time since last birth, and vital status of most recently born child were still the main determinants of fertility throughout this period. Nonetheless, controlling for these variables reveals independent effects of both socioeconomic status and fluctuations in grain prices, even in the earliest years of analysis. The authors also identify a monotonic decrease in fertility across the nineteenth century. However, analysis of interaction effects indicates that this fertility decline was not shared equally across social strata. The fertility of farmers and sharecroppers, who typically lived in multiple-family households and had the highest fertility overall, neither changed over time nor responded to short-term price oscillations. This article hints at a link between fertility decline in this region and the post-unification shift in agricultural production away from sharecropping, which privileged larger families, and toward wage labor, which privileged smaller ones. Rettaroli and Scalone’s breakthrough is the establishment of socioeconomic status as a determinant of fertility prior to the demographic transition.

Even as fertility declined throughout Europe, parents continued to employ postnatal strategies for controlling family size and composition. Revuelta Eugercios explores the complexities of abandonment, the most common European postnatal strategy. Relinquishing children to foundling hospitals had long been part of the family-building repertoire of poor parents, but the author demonstrates that in the early twentieth century, parents in Madrid and the surrounding areas began to utilize foundling hospitals as a means of temporary rather than permanent relief from the burden of caring for their children. Individual-level data from the Admission and Exits and Entries books of the Foundling Hospital of Madrid between 1890 and 1935 offer a unique opportunity to
examine how poor families adapted available maternal-welfare provisions to meet their needs in a moment when ideas of motherhood were changing and new social-reform legislation was emerging.

Revuelta Eugercios uses discrete-time event-history analysis to examine the correlates of this new use of the Foundling Hospital, marked by successful child retrieval. Results reveal two patterns: Among those born at the adjoining maternity hospital, legitimate children and the children of younger mothers were more likely to be retrieved, regardless of sex; among those who were brought to the foundling hospital after birth, male children and children who were older at the time of abandonment were more likely to be retrieved, as were children whose parents provided some kind of information to the institution at the time of abandonment. These patterns suggest that some of the women who availed themselves of the maternity hospital also utilized the foundling hospital for temporary child care as part of this welfare entitlement, and that some parents who brought their children to the foundling hospital viewed abandonment as a short-term survival strategy, leaving notes to indicate their retrieval intentions. Revuelta Eugercios’ analysis highlights the changing roles of institutions in demographic processes and parental initiatives in the use of institutions, as well as some evidence for son preference in early twentieth-century Madrid. Because of the high mortality rates in the foundling hospital, particularly among infants, parents likely recognized the risk involved in using the Foundling Hospital as a temporary refuge.

The six articles presented in this special issue of the Journal of Interdisciplinary History bring new substantive and methodological insights to the field of historical demography—revealing the responsiveness of pre-transition fertility to changing household and economic contexts, tracking the transmission of new fertility practices between generations and across space, exploring the unevenness of mortality and fertility decline within societies, and documenting the changing role of social institutions in family formation. They also suggest fruitful avenues for further research, indicating the vast and still largely untapped potential for longitudinal data and methods to provide additional knowledge about demographic processes, demographic change, and the connec-
tions between demographic change and social, economic, and political transformations. This collection demonstrates the immense value of training new generations of scholars in the longitudinal analysis of historical-demographic data, equipping them to extend this work in time and space as the number and size of life-history databases grows, and, as analytical methods continue to advance, inspiring them to seek new insights from the surviving traces of the past.