Journal of Interdisciplinary History, L:3 (Winter, 2020), 325–362.

50 years

The 50th Year: Special Essay 6

George C. Alter

The Evolution of Models in Historical Demography

In its first issue, the *JIH* published Goubert's account of a new approach to historical demography that originated in France in the 1950s. This "demographic revolution" brought together historians, statisticians, and demographers who used new methods to create a "scientific demographic history." This article examines the development of historical demography from the 1950s to the present by examining the development of its methods. It is not intended as a rigorous or comprehensive intellectual history of the field but as an appreciation of the extraordinary resourcefulness of the scholars who revolutionized the practice of historical demography during this period. It focuses on a few influential methodological developments, necessarily omitting many important substantive contributions.[1]

From a methodological perspective, historical demography underwent two important transitions with important links between them. The "classic" period, from the late 1950s to the mid-1980s saw an explosion of interest in historical demography, an expansion of research that involved the application of new

George C. Alter is Professor of History; Research Professor, Population Studies Center; and Research Professor, Inter-university Consortium for Political and Social Research, University of Michigan. He is co-author, with Richard D. Gonzalez, of "Responsible Practices for Data Sharing," *American Psychologist*, LXXIII (2018), 146–156; co-author, with Myron P. Gutmann, Susan Hautaniemi Leonard, and Emily R. Merchant, of "Introduction: Longitudinal Analysis of Historical-Demographic Data," *Journal of Interdisciplinary History*, XLII (2012), 503–517.

© 2019 by the Massachusetts Institute of Technology and The Journal of Interdisciplinary History, Inc., https://doi.org/10.1162/jinh_a_01445

---

[1] Pierre Goubert, "Historical Demography and the Reinterpretation of Early Modern French History: A Research Review," *Journal of Interdisciplinary History*, I (1970), 37–48. For a more comprehensive overview of historical demography, see also Antoinette Fauve-Chamoux, Ioan Bolovan, and Sølvi Sogner (eds.), *A Global History of Historical Demography: Half a Century of Interdisciplinarity* (Bern, 2016), the highly informative fiftieth-anniversary issue of *Annales de Démographie historique*; Fabrice Boudjaaba, Vincent Gourdon, Michel Oris, Isabelle Robin, and Marion Trévisi, "50 ans De Démographie Historique: Bilan Historiographique D'une Discipline En Renouvellement," *Annales de Démographie historique*, 129 (2015), 7–8; Myron P. Gutmann and Emily Klancher Merchant, "Historical Demography," in Dudley Poston (ed.), *Handbook of Population* (New York, 2019), 669–695.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

methods rooted in mathematical demography. These tools assume that regularities in demographic processes can be observed and modeled, and they create new ways to analyze historical sources. At that time, data collection was an important component of almost all research projects. Historical data were scarce and expensive to acquire, and research was characterized by an intense evaluation of sources for quality and potential biases. We benefit today from databases started during that period.

In the 1980s, historical demography began a reorientation from reconstructing trends to studying socioeconomic differentials in demographic behaviors. The most important contributions of the earlier generation had been aggregate-level reconstructions of trends in population sizes, fertility, and mortality, which were often interpreted within a framework derived from Thomas Malthus and Demographic Transition Theory. The next generation of research paid much more attention to individual-level analysis presented in a statistical rather than a demographic framework. This work emphasized differences within societies due to socioeconomic status, ethnicity, and community and to life-course models highlighting the importance of gender and age. New methods took advantage of changes in computer technology, advances in multivariate statistical models, and the emergence of large data collections.

The statistical methods used today and the demographic models used in the classical period have an underlying continuity. Although they are expressed in estimated coefficients and standard errors, statistical models are also based on simplifying assumptions about the phenomena being described. Event history analysis in particular makes the same assumptions about regularities in demographic behavior as do demographic projection and stable population theory. Unfortunately, this underlying continuity is not always understood, and some of the important lessons of the classic generation have been forgotten. I show why those principles are still important.

DEMOGRAPHIC MODELS IN CLASSIC HISTORICAL DEMOGRAPHY From the mid-1950s to the mid-1980s, historical demography was completely transformed by new methods that traced the demographic history of Europe in unprecedented detail and resulted in new ways of thinking about that history. These methods were

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

new applications of techniques of mathematical demography that had been perfected in the early twentieth century. Louis Henry and Ansley Coale, who were central figures in both historical demography and contemporary demography, also published textbooks about demographic methods. Four methods supported important findings during this pivotal period—the singulate mean age at marriage, family reconstitution, the European Fertility Project indices, and back projection.

Singulate Mean Age at Marriage In 1956, Hajnal identified a distinctive European pattern of marriage characterized by high average ages at marriage and high proportions never marrying. He showed that this pattern divided Europe into two zones; the "European marriage pattern" dominated northwest of a line from Trieste to St. Petersburg. Hajnal also argued that late marriage had been common in northwest Europe for at least two centuries, a point that he expanded in later writing. His evidence for this discovery relied on a statistic that he had described earlier, the singulate mean age at marriage (SMAM), which could be easily calculated from census data. Although the calculation of SMAM appears simple, Hajnal explicated its derivation from the life table and stable population theory.[2]

By the 1970s, the northwest European marriage pattern had become a central concept in historical demography. Laslett called attention to the correlation between the geography of late marriage and the prevalence of nuclear family households. Beginning with Le Play, sociological theorists had often argued that the modern pattern of small households was a product of the Industrial Revolution, which promoted individualism and weakened patriarchal authority. Laslett, however, showed that small families were common in England and elsewhere in northwest Europe for at least two centuries before the Industrial Revolution, suggesting that causation went from the culture of small families to individualism and economic innovation rather than the reverse. Wrigley described the importance of late marriage in a Malthusian world of limited resources and stagnant technology. Malthus had assumed a static economy in which population growth inevitably leads to

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[2] John Hajnal, "European Marriage Patterns in Perspective," in David V. Glass and David Edward Charles Eversley (eds.), *Population in History: Essays in Historical Demography* (Chicago, 1965), 101-143; *idem*, "Two Kinds of Preindustrial Household Formation System," *Population and Development Review*, VIII (1982), 449–494; *idem*, "Age at Marriage and Proportions Marrying," *Population Studies*, VII (1953), 111–136.

poverty; he envisioned only two ways to bring population and resources into balance—the positive check of high mortality or the preventive restraint of late marriage. Wrigley and Schofield argued that the positive check was no longer important in seventeenth-century England, which had come to rely on the preventive check and migration to cities and colonies.[3]

The remarkable achievement of SMAM is that it estimates an average age at marriage without any data on marriages, using only population counts by marital status and age from a single census. Hajnal's innovation involved seeing the similarity between proportions never married by age and the proportion of survivors in a life table. A life table describes the history of a hypothetical cohort of births, who experience a set of age-specific probabilities of dying. Starting with an arbitrary number of births, we compute the number who die in each year of life (d_x), the number of survivors at each birthday (l_x), and the number of "person-years" of life lived at each age (L_x). The expectation of life at birth (e_0, or average age at death) is computed by summing "person-years" of life and dividing by the number of births. Stable population theory shows that the life table can also be interpreted as a stationary population in which the number of births exactly equals the number of deaths. Under this interpretation, the L_x column can be understood as the number of people alive in the interval between age x and $x+1$. Hajnal draws an analogy between the proportion never married at each age in a census and the proportion of people surviving at each age in a stationary population. In his hypothetical stationary population, the number of never-married people "born" at age fifteen exactly equals the number who will marry (that is, "die") by age fifty.[4]

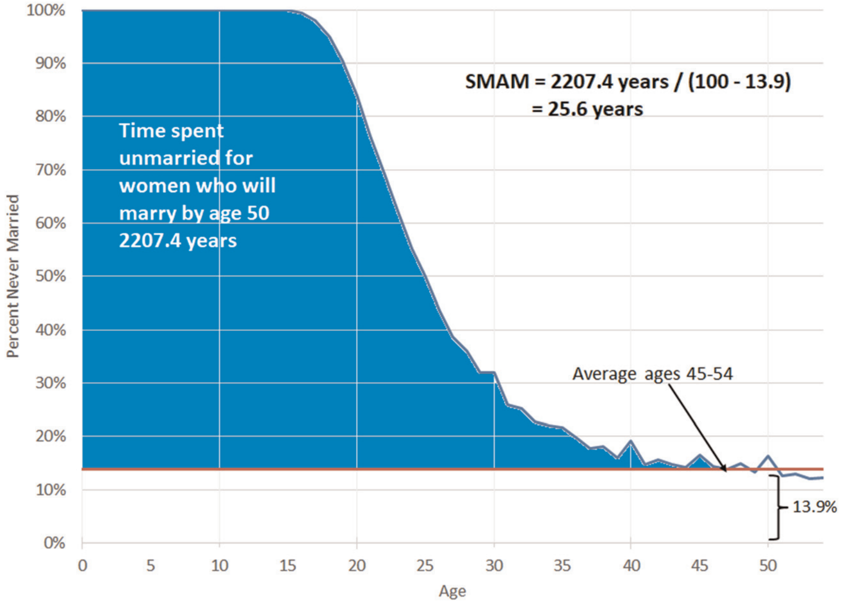
Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[3] Peter Laslett, "Characteristics of the Western Family Considered over Time," *Journal of Family History*, II (1977), 89–115; Frédéric Le Play, *L'Organisation de la Famille Selon Le Vrai Modèle Signalé par l'Histoire de Toutes les Races et de Tous les Temps* (Paris, 1895); Laslett, "Size and Structure of the Household in England over Three Centuries," *Population Studies*, XXIII (1969), 199–223; *idem*, "The Comparative History of Household and Family," *Journal of Social History*, IV (1970), 75–87; E. Anthony Wrigley, *Population and History* (New York, 1969); Thomas R. Malthus, *An Essay on the Principle of Population, as It Affects the Future Improvement of Society. With Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers* (London, 1798); Wrigley and Roger S. Schofield, *The Population History of England, 1541–1871: A Reconstruction* (London, 1981); Schofield, "Through a Glass Darkly: The Population History of England as an Experiment in History," *Journal of Interdisciplinary History*, XV (1985), 571–593.

[4] Samuel H. Preston, Patrick Heuveline, and Michel Guillot, *Demography: Measuring and Modeling Population Processes* (Malden, Mass., 2001), xiii, 291.

Fig. 1 Proportion of Women Who Were Single, by Age in France, 1851



SOURCE Social, Demographic, and Educational Data for France, 1801–1897, distributed by Inter-university Consortium for Political and Social Research, University of Michigan, Ann Arbor (February 16, 1992), available at https://doi.org/10.3886/ICPSR00048.v1.

This relationship is illustrated in Figure 1 by the proportions of French women married in 1853. The proportion of never married decreases with age, because those who marry cannot return to the status of never married. If this collection were a true cohort—that is, the same group of women at each age—the curve would only decrease, but it is not. Women at ages thirty, forty, and fifty in 1853 were born in 1823, 1813, and 1803, respectively. Hajnal assumed that they all married at the same rate, but discrepancies are possible, as is evident from the older ages in Figure 1. SMAM also assumes that no first marriages occur after age fifty, but this figure is an approximation, too.

If the survival curve in Figure 1 were depicting mortality, it would fall to zero around age 100, but this curve will never reach zero. Everyone dies, but some people never marry. To compute an average age at marriage, Hajnal had to subtract those who will never marry from the starting cohort. He estimated this proportion from those who are unmarried at ages forty-five to fifty-four. Then,

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

he computed an adjusted survival curve describing those who had not yet married as a proportion of those who will eventually marry.

Hajnal interpreted the area under the curve in Figure 1 as average person-years lived by never-married women. Since no marriages occur under age fifteen, every woman contributes fifteen years to the number of years lived before marriage. Above age fifteen, some women are married, and the contributions at these ages are scaled downward to the proportion of women who remain unmarried. The horizontal line at 13.9 percent in Figure 1 is the proportion of never-married women at ages forty-five to fifty-five. The area between the curve and this line is 2207.4 years, which needs to be divided by the proportion who ever married (100–13.9) to get SMAM, average years lived before marriage.

By recognizing the analogy between the life table and the age distribution of never-married women, Hajnal was able to use census data to estimate a cohort measure (average age at marriage) from cross-sectional data. Since it is not a true cohort measure, SMAM is affected by trends in marriage ages that can be highly sensitive to migration in local populations. Thus, Hajnal provided historical demography with both a powerful new tool and an important pattern in need of explanation.

Family Reconstitution Sogner dates the founding of historical demography as a discipline to Louis Henry's presentation of family reconstitution at the meeting of the International Congress for Historical Sciences in 1960 (Stockholm). The historical study of populations was not new in 1960, but Henry added a new level of scientific rigor. For the first time, historical demographers could estimate demographic rates with the same precision as contemporary demographers; they could apply this method to the enormous quantity of parish registers available in archives and churches across Europe, as well as to similar records around the world. The Fleury and Henry manual provided a detailed plan for conducting these studies, which was replicated in hundreds of monographs, published and unpublished. Family reconstitution set off a wave of new research across Europe, North and South America, and Japan.[5]

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[5] Sogner, "Historical Demography in Norway 1960–2010," in Fauve-Chamoux, Bolovan, and *idem* (eds.), *Global History of Historical Demography*, 499–512; Michel Fleury and Henry, *Des Registres Paroissiaux à l'histoire de la Population: Manuel de Dépouillement et d'exploitation de l'état Civil Ancien* (Paris, 1956), 142–144; René Le Mée, "De la Naissance de la Démographie

Rosental argues that Henry's contribution goes beyond the method of family reconstitution. All the components used in family reconstitution were already available before Henry became involved. In fact, Swedish demographer Hannes Hyrenius had invented the family-reconstitution form (see below) more than a decade earlier. But Henry provided a conceptual framework that placed historical research at the center of demographic research. In monographs and manuals, Henry and his co-authors demonstrated a broad range of techniques for reconstructing demographic patterns, including, though by no means limited to, family reconstitution. Henry's primary interest in using historical data to inform demographic models was clearly stated, but he made alliances with archivists, historians, and "passionate amateurs" who could use family reconstitution for their own purposes.[6]

Henry actively sought partnerships with other disciplines. At the end of his study of the leading families of Geneva he wrote: "*Much closer collaboration between demographers and historians is also necessary.* The demographer left to himself can furnish only results; he can certainly, as we have done, indicate interpretations suggested by the relationships that he observes among diverse phenomena. It remains to be seen whether these interpretations accord with what is already known in political, economic, and social history and, if not, to propose others. This task belongs to historians." Henry was offering historians not only scientific precision but also a new set of tools and rigorous tests for evaluating the quality of historical sources. Even Goubert, who had responded sharply to Henry's criticism of his own work, eventually endorsed family reconstitution.[7]

Henry was successful both in obtaining funding for his projects and in asserting his leadership within France and international organizations. By 1958, he had launched the *Enquête Henry*, a demographic database of randomly selected French parishes. He fully

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Historique à l'enquête Henry *Population (French Edition)*," L (1995), 1475–1487; Fauve-Chamoux, Bolovan, and Sogner (eds.), *Global History of Historical Demography*.

[6] Paul-André Rosental, "The Novelty of an Old Genre: Louis Henry and the Founding of Historical Demography," *Population*, LVIII (2003), 103–136; Michel Terrisse, "Aux Origines de la Méthode de Reconstitution des Familles: *Les Suédois d'estonie* de Hannes Hyrenius," *Population* (French Edition), XXX (1975), 143–155; Rosental, *L'intelligence Démographique* (Paris, 2003).

[7] Henry, *Anciennes Familles Genevoises; Étude Démographique: XVIe-XXe Siècle* (Paris, 1956), 232 (Alter translation, emphasis in original).

reconstituted forty parishes, and collected a larger sample of baptisms, burials, and marriages for aggregate analysis without record linkage to be employed in a reconstruction of the population of France by age and sex back to 1740. The results of the family reconstitutions, which were published in four regional studies, reveal substantial regional differences in the level of fertility during the early eighteenth century.[8]

Henry turned to historical data because he could not find contemporary data with the level of detail that he required for estimating and projecting fertility. As demographers were struggling to understand the baby boom, Henry started a series of publications aimed at more sophisticated ways to measure fertility. For example, his models distinguish between the effects of increasing secondary sterility and decreasing fecundity on the level of fertility at older ages. Early in this project he saw the need for data about "natural fertility" (fertility "of a population making no conscious effort towards birth control") that would allow him to distinguish the role of choice from the effects of biology. He mined the first family-reconstitution studies of Geneva and Crulai for empirical results shedding light on natural fertility. As data accumulated, he recognized that the level of fertility varied considerably among populations that were not practicing family limitation, and that cultural and environmental conditions were heavily involved.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[8] *Idem*, "Une Richesse Démographique En Friche: Les Registres Paroissiaux," *Population* (French Edition), VIII (1953), 281–290; Goubert, "Une Richesse Historique en Cours d'Exploitation: les Registres Paroissiaux," *Annales: Histoire, Sciences Sociales*, IX (1954), 83–93; Rosental, "Thirteen Years of Thinking: From Population History to Historical Demography (France 1945–1958)," *Population*, LI (1996), 1211–1238; Fauve-Chamoux, "Historical Demography and International Network Developments (1928–2010)," in *idem*, Bolovan, and Sogner (eds.), *Global History of Historical Demography*, 15–66; Isabelle Séguy et al., *La Population de la France de 1670 à 1829: l'enquête Louis Henry et Ses Données* (Paris, 2001), xvi, 208; Henry and Yves Blayo, "La Population de la France de 1740 à 1860," *Population* (French Edition), XXX (1975), 71–122; Blayo, "Mouvement Naturel de la Population Française de 1740 à 1829," *ibid.*, 15-64; Henry, "Fécondité des Mariages dans le Quart Sud-Ouest de la France de 1720 à 1829 (I)," *Annales*, XXVII (1972), 612–640; *idem*, "La Fécondité des Mariages dans le Quart Sud-Ouest de la France, de 1720 à 1829 (Suite)," *ibid.*, 977–1023; Jacques Houdaille and *idem*,"Fécondité des Mariages dans le Quart Nord-Ouest de la France de 1670 à 1829," *Population*, XXVIII (1973), 873–924; Houdaille, "La Fécondité des Mariages de 1670 à 1829 dans le Quart Nord-Est de la France," *Annales de Démographie historique* (1976), 341–391; Henry, "Fécondité des Mariages dans le Quart Sud-Est de la France de 1670 à 1829," *Population*, XXXIII (1978), 855–883; Alain Bideau and Jean-Pierre Bardet, "Une Géographie Très Contrastée," in Jacques Dupâquier, Alfred Sauvy, and Emmanuel Le Roy Ladurie (eds.), *Histoire De La Population Française* (Paris, 1988), 364–372.

In 1961, Henry issued a more nuanced definition of natural fertility: "We can term as natural the fertility that exists or has existed in the absence of deliberate birth control. The adjective 'natural' is admittedly not ideal but we prefer it to 'physiological' since the factors affecting natural fertility are not solely physiological: social factors may also play a part—sexual taboos for example, during lactation. Some of these factors may result in a reduction of fertility but this cannot be considered a form of birth control. Control can be said to exist when the behavior of the couple is bound to the number of children already born and is modified when this number reaches a maximum which the couple does not want to exceed: it is not the case for a taboo concerning lactation, which is independent of the number of children already born."[9]

Thus, for Henry, the opposite of natural fertility is the resort to birth control to terminate childbearing after reaching a target family size. The reconstitution of families in the Genevan bourgeoisie provided an opportunity to document this pattern. Henry found increasing evidence of fertility control starting with couples married between 1650 and 1699. Beginning with this cohort, age-specific fertility rates started to fall, the average age at last birth decreased, and completed family sizes decreased. The family reconstitutions strongly suggested that couples had begun aiming for small families. Fertility rates at younger ages remained constant, as did the length of the interval between the first and second birth. Birth intervals became wider at higher parities, but, as Henry noted, mostly due to the last and next-to-last intervals. Henry attributed these longer birth intervals to failed attempts to stop in an era when birth control was often ineffective. Moreover, Henry noted that the onset of fertility decline coincided with increased out-migration and reductions in marriage, which suggested economic adversity.[10]

Natural fertility, which attracted criticism almost immediately, is one of the most contentious concepts in historical demography. However, much of the controversy surrounding it stems from how to explain the transition to low fertility, and not from the

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[9] Henry, "Some Data on Natural Fertility," *Eugenics Quarterly*, VIII (1961), 81–91. See also Rosental, "Novelty of an Old Genre," 103–136; Henry (ed. and trans. Mindel C. Sheps and Evelyne Lapierre-Adamcyk), *On the Measurement of Human Fertility: Selected Writings of Louis Henry* (New York, 1972), xix, 228; *idem*, "Fondements Théoriques des Mesures de la Fécondité Naturelle," *Revue de l'Institut International de Statistique*, XXI (1953), 135–151.

[10] Henry, *Anciennes Familles Genevoises*, 71-110, 178–179.

questions that motivated Henry initially. As explanations shifted from socioeconomic to cultural (see below), the "natural fertility hypothesis" became part of the debate about what initiates family limitation. One side of the debate espouses a strict interpretation of "natural fertility," which implies that fertility decline depended upon the spread of knowledge and attitudes favorable to birth control. The other side maintains that European couples already knew how to control fertility; the key change was a new preference for small families.[11]

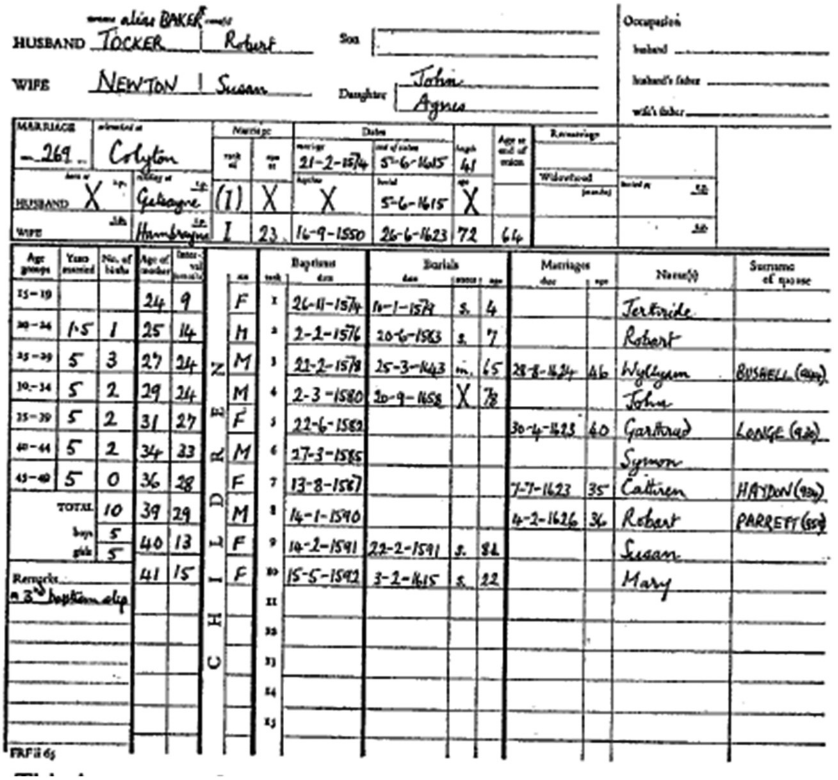
Family reconstitution solved a well-known problem with the parish registers: How do we use counts of births, deaths, and marriages when we do not know the size of the population? The standard way to compute demographic rates is to divide the number of events (births, deaths, marriages, or migrants) by the average number of people who could have experienced these events. The usual denominator is the mid-year population estimated from a census. Henry's solution was to use a different denominator, person-years of experience, a basic feature of the life table. A life table can be interpreted as the history of a hypothetical cohort of births. The age-specific death rate in the life table, m_x, is the ratio of deaths between ages x and $x+n$ and the number of person-years lived by the cohort in this age interval. For contemporary populations, demographers approximate m_x by dividing counts of deaths by the mid-year population at ages x to $x+n$. Henry recognized that person-years lived could be estimated directly from the parish registers without a census. If we can link a woman's date of birth/baptism to her date of death/burial, we can compute the number of person-years that she lived in each age of her life. Thus, we can compute age-specific mortality, fertility, and nuptiality rates without a census.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[11] Goubert, "Une Richesse Historique"; Gutmann and Merchant, "Historical Demography"; John E. Knodel and Etienne van de Walle, "Lessons from the Past: Policy Implications of Historical Fertility Studies," *Population and Development Review*, V (1979), 217–245; Coale, "The Demographic Transition Reconsidered," *International Population Conference, Liege*, I (1973), 53–72; Angus McLaren, *Reproductive Rituals: The Perception of Fertility in England from the Sixteenth Century to the Nineteenth Century* (New York, 1984); Tommy Bengtsson and Martin Dribe, "Deliberate Control in a Natural Fertility Population: Southern Sweden, 1766–1864," *Demography*, XLIII (2006), 727-746; Alter, "Theories of Fertility Decline: A Non-Specialist's Guide to the Current Debate on European Fertility Decline," in Louise Tilly, John R. Gillis, and David Levine (eds.), *The European Experience of Declining Fertility, 1850–1970* (Cambridge, Mass., 1992), 13-27; Timothy W. Guinnane, "The Historical Fertility Transition: A Guide for Economists," *Journal of Economic Literature*, XLIX (2011), 589–614; Gosta Carlsson, "Decline of Fertility—Innovation or Adjustment Process," *Population Studies—a Journal of Demography*, XX (1966), 149–174.

Fig. 2 Family Reconstitution Form



alias BAKER

HUSBAND TOCKER | Robert

WIFE NEWTON | Susan

Son

Daughter John
Agnes

Occupation
husband
husband's father
wife's father

MARRIAGE 269	Colyton			21-2-1574	5-6-1615	41		Remarriage
HUSBAND	X	[illegible]	(1)	X	X	5-6-1615	X	Widowhood
WIFE		[illegible]	I	23	16-9-1550	26-6-1623	72	64

Age groups	Years married	No. of births
15–19		
20–24	1·5	1
25–29	5	3
30–34	5	2
35–39	5	2
40–44	5	2
45–49	5	0
TOTAL		10
boys		5
girls		5

Remarks
n 3rd baptism slip

CHILDREN

Age of mother	Interval (months)	sex	rank	Baptisms date	Burials date		age	Marriages date	age	Name(s)	Surname of spouse
24	9	F	1	26-11-1574	16-1-1578	s.	4			Tertrude	
25	14	M	2	2-2-1576	20-6-1583	s.	7			Robert	
27	24	M	3	22-2-1578	25-3-1643	m.	65	28-8-1624	46	Wyllyam	BUSHELL ([illegible])
29	24	M	4	2-3-1580	20-9-1658	X	78			John	
31	27	F	5	22-6-1582				30-4-1623	40	Garthrud	LONGE ([illegible])
34	33	M	6	27-3-1585						Symon	
36	28	F	7	13-8-1587				7-7-1623	35	Cattren	HAYDON ([illegible])
39	29	M	8	14-1-1590				4-2-1626	36	Robert	PARRETT ([illegible])
40	13	F	9	14-2-1591	22-2-1591	s.	8d			Susan	
41	15	F	10	15-5-1592	3-2-1615	s.	22			Mary	
			11								
			12								
			13								
			14								
			15								

FRF ii 65

SOURCE E. Anthony Wrigley, David Edward Charles Eversley, and Peter Laslett, *An Introduction to English Historical Demography, from the Sixteenth to the Nineteenth Century* (New York, 1966), 126.

The Fleury and Henry manual presented simple paper forms and step-by-step procedures for performing a family-reconstitution study. All baptisms, marriages, and burials are transcribed to color-coded slips of paper and then sorted and assembled into families. A family begins with a marriage; it includes a husband, wife, and children. When all the events associated with a family have been assembled, they are transferred to a family reconstitution form (FRF), as shown in Figure 2. Information about the husband and wife is at the top of the form. Children are listed in order in the central columns below their parents. A table on the left is divided into age groups for recording the mother's time at risk and number of births for computing age-specific fertility rates.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Family reconstitution began as a manual process and resisted computerization for a long time.[12]

Henry and his co-authors were sensitive to the matter of data quality. Acutely aware of the limitations of the parish registers, they went to great lengths to evaluate their sources. The first chapter in Henry's *Manuel de démographie historique* is "Vérification des données." His technical manuals offer tools for detecting and estimating missing events. Fertility rates in each of the regional reports on the *Enquête Henry* are adjusted to correct for births that were not reported as baptisms.[13]

The central problem in family reconstitution is the handling of incomplete life histories that arises because the parish registers did not record migration. Since searching for people over a wider area is difficult and expensive, almost all family reconstitution studies focus on one district, usually a single parish. Even the histories of families who moved to an adjacent parish are incomplete. We can know that a person was present in the community only when an event (baptism, marriage, or burial) was recorded for them or a close family member. The family migrated sometime after the last recorded event, but we have no way of knowing how long they were present in the parish before moving away.

Because Fleury and Henry recognized that incomplete histories could bias the analysis of family reconstitution data, they introduced strict rules for selecting which families should be included. Family histories qualify for an analysis only when the event ending a history is unrelated to the subject of the analysis. If an analysis is about fertility, family histories that end with a birth or death of a child must be excluded, because they will overestimate the level of fertility. Statisticians now call this principle "noninformative censoring."[14]

We can illustrate this problem by imagining a population consisting of couples with only two fertility patterns. Suppose that

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[12] Fleury and Henry, *Des Registres Paroissiaux.*

[13] Henry, *Manuel de Démographie Historique* (Paris, 1967), xii, 148; *idem*, *Techniques d'analyse en Démographie Historique* (Paris, 1980), 177; Henry and Alain Blum, *Techniques d'analyse en Démographie Historique* (Paris, 1988); Henry, "Fécondité des Mariages dans le Quart Sud-Ouest (I)"; *idem*, "La Fécondité des Mariages dans le Quart Sud-Ouest (Suite)"; Houdaille and *idem*, "Fécondité des Mariages dans le Quart Nord-Ouest"; Houdaille, "La Fécondité des Mariages dans le Quart Nord-Est"; Henry, "Fécondité Des Mariages dans le Quart Sud-Est."

[14] John D. Kalbfleisch and Ross L. Prentice, *The Statistical Analysis of Failure Time Data* (New York, 1980), 195–196.

some couples have a birth every three years, whereas others have a birth every five years. Assume also the same proportion of out-migrants in each group, and a timing of migration that is random with respect to childbearing. The last event that we observe in the migrant families will be a birth or the death of a child, because the husband and wife will die somewhere else. Under these assumptions, the average time between the last observed birth and the migration date will be longer in the families with long birth intervals than among those with short birth intervals. By construction, the maximum time between last birth and migration will be five years for couples with longer birth intervals instead of three years for the short interval couples. This time span should be included in the denominator of the fertility rate, because the couple could have had a birth in the study area during this time. However, we cannot measure the time between last birth and migration, because migration was not recorded in the parish registers. If we use the date of the last observed birth to close family histories, we will be excluding more time at risk from the low fertility couples than from the high fertility couples. Consequently, the fertility rates that we compute will be too high.

Unfortunately, this important principle was asserted but not explained in the seminal texts on family reconstitution. Fleury and Henry wrote, "As the date of the end of union holds an essential place in the study of fertility, it is only usable when it is known independently of any document, death or marriage in particular, concerning the children; not respecting this rule favors the most fertile families and leads to an over-estimation of fertility."[15]

Since the Henry manuals were never translated into English, the most detailed account of family reconstitution came from Wrigley, Eversley, and Laslett, who offered this version of the same point: "[I]t is clear that evidence from this FRF could not be used in the study of marital fertility. The continued residence of the family in the parish is known only because of the baptism and burial of children. The burial of the parents is not recorded in the register. If they had been childless nothing would have been known about them after their marriage and they would not have been included in any study of marital fertility. If families whose

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[15] Fleury and Henry, *Nouveau Manuel de Dépouillement et d'Exploitation de l'État Civil Ancien* (Paris, 1965), 183 (Alter's translation).

residence in the parish is known only from entries about their children in the baptism and burial registers are included in the calculation of marital fertility rates, the rates which result will be higher than the true rates."[16]

Even if it was not fully explained, the prohibition of informative censoring was embedded in the rules and procedures for dating the end of observation of a family. The period of observation for the analysis of fertility is usually closed by the date of death of the spouse that died first. A census or tax document showing that a couple was present in the study area may be sufficient to close observation when the couple died somewhere else. However, when the last document pertaining to the family is a baptism, child death, or marriage of a child, the end of observation is considered "open," and the family is not eligible for computing fertility rates. The computation of age-specific fertility rates also requires both the date of marriage and the age of the mother, which is usually obtained by linking to her baptism. These dates will not be available if the couple migrated into the study area after marrying somewhere else. More than half of the families in a parish are usually excluded from fertility analysis because they lack the dates required to begin or end observation. Thus, family reconstitution describes the sedentary population who spent their entire lives in a single parish.[17]

THE EUROPEAN FERTILITY PROJECT The European Fertility Project under the direction of Ansley Coale was one of the first "big data" projects in historical demography. Conceived as a test of demographic transition theory, the European Fertility Project was designed to measure the impact of economic development on fertility decline

---

[16] Wrigley, Eversley, and Laslett, *An Introduction to English Historical Demography, from the Sixteenth to the Nineteenth Century* (New York, 1966), 148.

[17] Henry, *Techniques d'Analyse*, 67–69; Steven Ruggles, "The Limitations of English Family Reconstitution: English Population History from Family Reconstitution 1580–1837," *Continuity and Change*, XIV (1999), 105–130; Wrigley, "The Effect of Migration on the Estimation of Marriage Age in Family Reconstitution Studies," *Population Studies—a Journal of Demography*, XLVIII (1994), 81–97; Levine, "Reliability of Parochial Registration and Representativeness of Family Reconstitution," *ibid.*, XXX (1976), 107–122; Ruggles, "Migration, Marriage, and Mortality—Correcting Sources of Bias in English Family Reconstitutions," *ibid.*, XLVI (1992), 507-522; Alice B. Kasakoff and John W. Adams, "The Effect of Migration on Ages at Vital Events—a Critique of Family Reconstitution in Historical Demography," *European Journal of Population*, XI (1995), 199–242.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

in nineteenth-century Europe. Under the influence of books like Erlich's *The Population Bomb*, concern about the rapid growth of population in Africa and Asia had spread from academic and policy circles to the general public. Coale's earlier work had shown how rapid population growth could inhibit economic development. Demographic transition theory stressed the impact of industrialization and urbanization on the costs of children and on attitudes about large families. Demographers began to explore whether fertility decline was possible without economic development.[18]

Coale and his associates set out to map patterns of economic development and fertility by consulting the abundant censuses and vital registration available for Europe from the middle of the nineteenth century. Most countries' annual counts of births, marriages, and deaths, as well as decennial counts of population by age and sex at the provincial and often the district level, were readily available in libraries on both sides of the Atlantic. The European Fertility Project collected and digitized these data for Europe as a whole, from Ireland to Russia, publishing books and articles about fourteen countries and describing the implications of the project in an influential summary volume.[19]

The findings of the European Fertility Project had far-reaching ramifications in contemporary as well as historical demography. Maps, like Figure 3, dramatically undermined the assumption in demographic transition theory that industrialization, urbanization, and demographic change moved together. France stands out clearly in Figure 3 as the early leader in fertility decline. The fertility transition in England, the first industrialized country in the world, lagged behind France by at least half a century. Demeny anticipated this finding in the early years of the project, showing that fertility decline in Hungary, one of the least industrial areas in nineteenth-century Europe, was almost simultaneous with the decline in England. The distinctive imprint of French borders in Figure 3 also points to the unexpected importance of national and linguistic boundaries

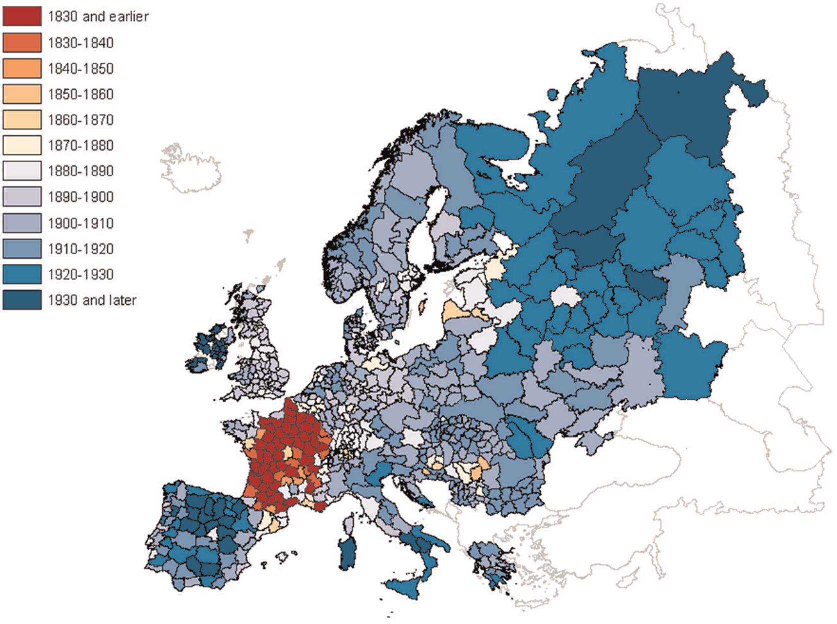
---

[18] Merchant, "Prediction and Control: Global Population, Population Science, and Population Politics in the Twentieth Century," (unpub. Ph.D. diss. Univ. of Michigan, 2015), 336–345; Paul R. Ehrlich, *The Population Bomb* (San Francisco, 1969); Coale and Edgar Malone Hoover, *Population Growth and Economic Development in Low-Income Countries: A Case Study of India's Prospects* (Princeton, 1958).

[19] Coale and Susan Cotts Watkins, *The Decline of Fertility in Europe: The Revised Proceedings of a Conference on the Princeton European Fertility Project* (Princeton, 1986).

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Fig. 3 The Timing of the Fertility Transition, by Region



SOURCE Ansley J. Coale and Susan Cotts Watkins, *The Decline of Fertility in Europe: The Revised Proceedings of a Conference on the Princeton European Fertility Project* (Princeton, 1986).

in fertility decline. Lesthaeghe showed that the timing of fertility decline among Flemish and French speakers varied systematically between pairs of villages on opposite sides of the Flemish–Walloon border. The importance of language had also been anticipated in a study of Spain by Leasure, whose mentor was Coale.[20]

The evidence produced by the European Fertility Project turned attention away from socioeconomic explanations of fertility decline toward the study of attitudes and culture. Coale did not dismiss the importance of socioeconomic factors, but he also emphasized knowledge and attitudes about birth control. Knodel and van de Walle suggested that the diffusion of information and attitudes favorable toward birth control played a key role in the fertility transition. These conclusions were important in the contemporary

---

[20] Paul Demeny, "Early Fertility Decline in Austria-Hungary: A Lesson in Demographic Transition," *Daedalus*, XCVII (1968), 502–522; Ron J. Lesthaeghe, *The Decline of Belgian Fertility, 1800–1970* (Princeton, 1977); J. William Leasure, "Factors Involved in the Decline of Fertility in Spain 1900–1950," *Population Studies*, XVI (1963), 271–285.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

debate between those who advocated family-planning programs in high-fertility countries and others who argued for the necessity of economic development.[21]

Figure 3 derives from new measures of fertility developed for the European Fertility Project. The data available to the project posed a familiar problem for demographic research. Aggregate measures, like the crude birth rate (births divided by total population), are sensitive to differences in age, sex, and marital-status distributions across populations. Places with early marriages and young populations are likely to have higher crude birth rates than places with late marriages and older populations. Demographers prefer to solve this problem by computing age-specific birth rates separately for married and unmarried women, which can then be summarized by a weighted average. These computations require tables of births by the age and marital status of mothers as well as census counts of all women by age and marital status. The European Fertility Project had detailed census counts, but national statistical offices rarely published births by age of mother in the nineteenth century. Coale turned to indirect standardization, a technique well known to demographers and actuaries.

Indirect standardization assumes that the age-specific rates of a population of interest are a constant proportion of the age-specific rates in a reference population, as represented by

$$f_a^i = I^i f_a^S \quad (1)$$

where

f_a^i is the age-specific fertility rate for age a in population i,

f_a^S is the age-specific fertility rate for age a in a standard population (S), and

I^i is a constant for all age groups.

Under this proportionality assumption, we can solve for a multiplier or index (I^i) that can be used to convert rates in the standard population to rates in the population of interest.[22]

---

[21] Coale, "Demographic Transition Reconsidered"; Dennis Hodgson, "Orthodoxy and Revisionism in American Demography," *Population and Development Review*, XIV (1988), 541-569; *idem*, "Demography as Social Science and Policy Science," *ibid.*, IX (1983), 1–34; Merchant, "Prediction and Control."

[22] Henry S. Shryock, Jacob S. Siegel, and Edward G. Stockwell, *The Methods and Materials of Demography* (New York, 1976), 285.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

The solution involves calculating the expected number of births that would have occurred in the standard population if it had the same age distribution as the population of interest. The formula is

$$I^i = \frac{\sum f_a^i w_a^i}{\sum f_a^S w_a^i}. \qquad (2)$$

where

w_a^i is the number of women at age *a* in population *i*,
$\Sigma f_a^i w_a^i$ is the total number of births to all women in population *i*, and
$\Sigma f_a^S w_a^i$ is the total number of births that women in population *i* would have had if they had the fertility rates in the standard (*S*) population.

Although they did not have age-specific fertility rates, the European Fertility Project did have the total number of births, which is the numerator in Equation 2, and the number of women at each age required to compute the denominator. The Project used age-specific fertility rates of the Hutterites, an American religious sect known for their high fertility, as the standard in these calculations. Furthermore, the Project ingeniously devised an equation separating the effects of marriage, illegitimate fertility, and fertility within marriage:

$$I_f = I_m \times I_g \times (1.0 - I_m) \times I_n \qquad (3)$$

where

I_f = index of overall fertility,
I_g = index of marital fertility (relative to the Hutterites),
I_h = index of non-marital fertility, and
I_m = index of marriage.

All of these indexes are scaled to Hutterite fertility, including the illegitimacy and marriage measures. An index of marital fertility (I_g) of .7 implies that married women were having children at 70 percent the rate of Hutterite married women. The index of marriage (I_m) is not a pure measure of marriage patterns because age groups are weighted by their level of fertility in the Hutterite population. An

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

index of marriage (I_m) of .6 implies that marriage patterns reduced the potential fertility of the population by 40 percent.

As we will see, the proportionality assumption embedded in the European Fertility Project indexes continues to play a central role in historical demography. It is based on confidence that demographic processes follow patterns determined by biological and social regularities. Coale was one of the foremost developers of demographic measurement techniques based on this assumption. Although the methods used by the European Fertility Project to date fertility transitions have come under fire, the European Fertility Project fertility indexes enabled comparisons of fertility and marriage across time and space on an unprecedented scale.[23]

BACK PROJECTION In 1985, the *JIH* published a special issue examining the implications of Wrigley and Schofield's *Population History of England*, which represented much more than an incremental addition to knowledge about the size of the English population. Wrigley and Schofield presented annual estimates of the full range of demographic indicators for population size, age structure, mortality, fertility, and nuptiality. Underlying this work was a powerful new technique, back projection, with a fully articulated demographic model. Back projection is an extension of inverse projection, a method invented by Lee that employed well-known demographic tools in a new way.[24]

As mentioned above, Wrigley, who had been describing demographic history in terms of the Malthusian model since the 1960s, drew particular attention to the potential role of late marriage in moderating the rate of population growth and the recurrence of subsistence crises. *The Population History of England* provided evidence that population growth reduced the standard of living, as Malthus

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[23] Coale, Demeny, and Barbara Vaughan, *Regional Model Life Tables and Stable Populations* (New York, 1983); Coale and Donald R. McNeil, "Distribution by Age of Frequency of First Marriage in a Female Cohort," *Journal of the American Statistical Association*, LXVII (1972), 743–749; Coale and T. James Trussell, "Model Fertility Schedules—Variations in Age Structure of Childbearing in Human Populations," *Population Index*, XL (1974), 185–258; Guinnane, Barbara S. Okun, and Trussell, "What Do We Know about the Timing of Fertility Transitions in Europe," *Demography*, XXXI (1994), 1–20.

[24] Schofield and Wrigley, "Population and Economy: From the Traditional to the Modern World," *Journal of Interdisciplinary History*, XV (1985), 561–569; *idem*, *Population History of England*; Ronald D. Lee, "Estimating Series of Vital Rates and Age Structures from Baptisms and Burials: A New Technique, with Applications to Pre-Industrial England," *Population Studies*, XXVIII (1974), 495–512.

claimed, but it argued that poorer living conditions reduced marriage in a way that dampened the cycle of growth and crisis. Moreover, Wrigley and Schofield added a new element to the history of modern population growth. Earlier accounts attributed the beginning of modern population growth in the eighteenth century to a decrease in mortality, partly due to the absence of bubonic plague. The back-projection estimates revealed an increase in fertility during the eighteenth century as well. Wrigley and Schofield argued that rising fertility stemmed from an increase in marriage, which began as a Malthusian response to rising wages. However, the early development of manufacturing in England loosened Malthusian constraints and sustained higher marriage and fertility rates into the nineteenth century. Weir demonstrated that this increase in marriage was due to a decrease in permanent celibacy, not younger ages at marriage.[25]

In demographic terms, back projection operates in the opposite direction of the European Fertility Project indexes and SMAM. Whereas Coale and Hajnal used population counts from censuses to estimate events (births and marriages), Wrigley and Schofield used events (births and deaths) to estimate population counts for a time before censuses were conducted. The conventional approach to population projection is the "cohort component" method. A projection begins with a baseline population divided into subgroups by age and sex. In each period, the subgroups are multiplied by selected mortality, fertility, and out-migration rates to calculate numbers of deaths, births, and migrants, which determine the population at the beginning of the next period. Thus, a projection consists of a known starting population and a set of hypothetical rates. Lee inverted this procedure by using numbers of births, deaths, and migrants to estimate rates. Lee's inverse projection finds a set of rates that produces the observed number of events.[26]

At its heart, inverse projection uses a simplifying assumption about demographic rates that is a variant of indirect standardization. Lee reduced the universe of possible demographic rates by

---

[25] Wrigley and Schofield, *Population History of England*, 438–443, 473–476; David R. Weir, "Rather Never than Late: Celibacy and Age at Marriage in English Cohort Fertility, 1541–1871," *Journal of Family History*, IX (1984), 340–354.

[26] Preston, Heuveline, and Guillot, *Demography*, 119-127; Lee, "Inverse Projection and Back Projection: A Critical Appraisal, and Comparative Results for England, 1539 to 1871," *Population Studies*, XXXIX (1985), 233–248.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

considering sets of rates that could be identified by a single parameter. For fertility, he reasoned that the age-specific fertility rates in any time period are all proportional to a set of standard rates, which is the same assumption employed by the European Fertility Project shown above in Equation 1. For mortality, Lee used a model in which rates are related by a linear model where

$$q_{i,t} = q_i + k_t d_i \tag{4}$$

where

$q_{i,t}$ is the death rate in age group i at time t,
q_i and d_i are constants depending only on age, and
k_t is a constant depending only on time period.

This model has the same property as Equation 2. It is possible to solve for k_t by using the total number of deaths, the constants (q_i and d_i), and the number of people at each age. Thus, Lee turns a single number, total deaths, into a set of age-specific rates for each period, applying them to the starting population to estimate deaths at each age and the surviving population at the end of the period.[27]

As its name implies, back projection is a modification of inverse projection that operates backward in time. The assumptions in inverse projection are straightforward, but back projection requires additional assumptions described by Oeppen in the 1981 volume. Lee criticized the assumptions in back projection, but he reproduced most of the results in *The Population History of England* with inverse projection. Oeppen later developed a more general version of back projection that integrates more data in the calculations.[28]

The Cambridge Group dedicated the same attention to data quality as Henry. The first 154 pages of *The Population History of England 1541–1871* is devoted to evaluating data quality and correcting for problems. Not to be outdone, the Cambridge Group

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[27] Lee, "Estimating Series of Vital Rates and Age Structures."

[28] Wrigley and Schofield, *Population History of England*, 715–738; Lee, "Inverse Projection and Back Projection," 233–248; Jim Oeppen, "Back Projection and Inverse Projection: Members of a Wider Class of Constrained Projection Models," *Population Studies*, XLVII (1993), 245–267.

family reconstitution volume has five appendixes on data quality as well as numerous discussions in the text.[29]

MODELS IN HISTORICAL DEMOGRAPHY SINCE THE 1990S In the 1990s, historical demography entered a new phase that differed in important ways from the work of pioneers like Henry, Hajnal, Coale, Wrigley, and Schofield. The types of models in historical demography shifted from demographic to statistical. Whereas key features of earlier research can be linked to life tables and stable population theory, multiple regression became the dominant organizing framework. Regression was not new in quantitative historical research, but three developments allowed it to acquire new importance in historical demography—changing research questions, easier access to individual-level data, and new methods.

First, the focus of research in historical demography shifted away from long-run trends in national populations toward differences within populations. On one hand, this development was a tribute to the success of the previous generation of studies. Henry's team at the Institut national d'études démographiques (INED) had reconstructed population dynamics in France as far back as 1740; the Cambridge Group had taken England back to 1541. Although debate about the quality and the meaning of those estimates sometimes occurred, the results found wide acceptance. On the other hand, the study of differential experiences by class, race, and ethnicity has a long history in demography, and it was never absent from historical demography. In the 1980s and 1990s, studies focusing on socioeconomic differences became much more common. For example, Szreter assigns an important role to subnational "communication communities" that have both social and geographical dimensions.[30]

Second, large-scale databases and comparative projects became much more important. Family reconstitution created a template for village studies that even a solitary scholar could complete.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[29] Wrigley and Schofield, *Population History of England*; Wrigley, R. S. Davies, Oeppen, and Schofield, *English Population History from Family Reconstitution, 1580–1837* (New York, 1997).

[30] Henry and Blayo, "La Population de la France de 1740 à 1860," 71-122; Blayo, "La Proportion de Naissances Illégitimes en France de 1740 à 1829," *Population (French Edition)*, XXX (1975), 65–70; *idem*, "La Mortalité en France de 1740 à 1829," *ibid.*, 123–142; Wrigley and Schofield, *Population History of England*; Simon Szreter, *Fertility, Class and Gender in Britain, 1860–1940* (New York, 1996), 546–547; Eilidh Garrett, Alice Reid, and Szreter, *Changing Family Size in England and Wales: Place, Class, and Demography, 1891–1911* (New York, 2001).

The operations of transcribing the original documents and linking individuals to families were done on paper; computers came into play only for the final calculation of tables, if at all. Even the European Fertility Project was organized as a series of single-author national monographs. This routine began to change in the 1980s as new computer technology dramatically reduced the cost of collecting and analyzing historical data. Instead of transcribing documents to paper coding forms and then to punch cards, researchers could type their data directly into inexpensive personal computers. Flexible database software reduced the costs of managing data and allowed consistency checks to be built into data-entry programs. New technology accelerated a trend toward larger databases intended for broad research programs, such as the Programme de Recherche en Démographie Historique in Quebec, the Demographic Data Base at the University of Umeå, the Utah Population Database, the Historical Sample of the Netherlands, and the TRA Project in France. In 1999, the North Atlantic Population Project was formed to create comparative samples of historical population censuses based on the model established by the Integrated Public Use Microdata Series led by Ruggles in the United States. The new databases welcomed researchers studying a wide range of questions.[31]

An important subset of these databases focused on population registers, which have important advantages over parish registers for demographic analysis. The key difference between them is that population registers contain explicit reporting of migration. Population registers are not exempt from the principles that motivated Henry's rules for family reconstitution, but the availability of clear dates for entry and exit from observation allows the experiences of migrants to be analyzed. Population registers also tend to have more socioeconomic information than parish registers. Since they were usually maintained for administrative rather than religious purposes, population registers often have consistent reporting of occupations. In addition, the organization of population registers as lists of individuals within households allows for the dynamic study of co-residence. Ever since van de Walle demonstrated the advantages of population registers in studies of the Belgian village of La Hulpe, historical demographers have turned to population

---

[31] Ruggles et al., "The North Atlantic Population Project: Progress and Prospects," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, XLIV (2011), 1–6.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

registers or similar documents from Belgium, China, Italy, Japan, Korea, Netherlands, Russia, Sweden, and Taiwan.[32]

Third, technology for statistical analysis, especially event history analysis, was catching up to the needs of historical demographers interested in comparative studies. Tables of age-specific rates can be helpful in simple comparisons, but they quickly become unwieldy as the number of dimensions (time period, marital status, occupation, urban/rural residence, etc.) increases. Statisticians were developing regression models that could accommodate multiple explanatory variables in the 1970s, but such models did not become available in general purpose statistics packages, like SAS and Stata, until the late 1980s and 1990s. The Cox proportional-hazards model in particular allowed historical demographers to examine the effects of individual-level characteristics, like occupation and family size, on demographic events.[33]

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

32 Gutmann and van de Walle, "New Sources for Social and Demographic History: The Belgian Population Registers," *Social Science History*, II (1978), 121–143; van de Walle, "Household Dynamics in a Belgian Village, 1847-1866," *Journal of Family History*, I (1976), 80–94; *idem* and Olivier Blanc, "Registre de Population et Démographie: La Hulpe," *Population et Famille*, XXXVI (1975), 113–128; Alter, *Family and the Female Life Course: The Women of Verviers, Belgium, 1849–1880* (Madison, 1988); James Z. Lee and Cameron Campbell, *Fate and Fortune in Rural China: Social Organization and Population Behavior in Liaoning, 1774–1873* (New York, 1997); David I. Kertzer, *Family Life in Central Italy, 1880–1910: Sharecropping, Wage Labor, and Coresidence* (New Brunswick, 1984), xvii, 250; Laurel L. Cornell and Hayami, "The Shumon-Aratame-Cho–Japan Population Registers," *Journal of Family History*, XI (1986), 311–328; Akira Hayami, *Kinsei Noson No Rekishi-Jinkogakuteki Kenkyu* (A Historical Demographic Study of Agricultural Villages in Early Modern Japan) (Tokyo, 1973); Thomas C. Smith, Robert Y. Eng, and Robert T. Lundy, *Nakahara: Family Farming and Population in a Japanese Village, 1717–1830* (Stanford, 1977); Sangkuk Lee and Wonjae Lee, "Strategizing Marriage: A Genealogical Analysis of Korean Marriage Networks," *Journal of Interdisciplinary History*, XLVIII (2017), 1–19; Angelique Janssens, *Family and Social Change: The Household as a Process in an Industrializing Community* (New York, 1993); Frans Van Poppel and Kees Mandemakers, "Differential Infant and Child Mortality in the Netherlands: First Results of the Historical Sample of the Population of the Netherlands," in Alain Bideau, Bertrand Desjardins, and Héctor Pérez Brignoli (eds.), *Infant and Child Mortality in the Past* (New York, 1997), 276–300; Theo Engelen, John Robert Shepherd, and Wen-Shan Yang (eds.), *Death at the Opposite Ends of the Eurasian Continent: Mortality Trends in Taiwan and the Netherlands, 1850–1945* (Amsterdam, 2011); Blum and Irina Troitskaja, "Mortality in Russia during the Eighteenth and Nineteenth Centuries: Local Assessments Based on the Revizii," *Population*, LI (1996), 303–328; Bengtsson, "The Vulnerable Child. Economic Insecurity and Child Mortality in Pre-Industrial Sweden. A Case Study of Västanfors, 1757–1850," *European Journal of Population*, XV (1999), 117–151; Anders Brändström and Jan Sundin, "Infant Mortality in a Changing Society: The Effects of Child Care in a Swedish Parish 1820–1894," in *idem* (eds.), *Tradition and Transition: Studies in Microdemography and Social Change* (Umeå, 1981), 67–104.

33 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 Eurasia Project on Population and Family History, which grew out of a 1994 meeting convened by Akira Hayami to discuss methods of analysis for historical population registers, illustrates all these trends. The outcome was a comparative study of demographic responses to economic stress in five countries (Belgium, China, Italy, Japan, and Sweden) that produced three comparative volumes and dozens of papers at international meetings and conferences. The questions addressed by the Eurasia Project were not new. In fact, they can be traced in a direct line back to Malthus. But these questions were posed in new ways.[34]

The shift from long-run trends to individual-level comparisons is evident in the Eurasia Project's focus on individuals within families. The statistical model favored in the Eurasia Project was an extension of the Cox proportional-hazards model to allow the inclusion of annual grain prices with time lags as a community-level variable. Individual (for example, age, sex, and relationship to household head) and household (size, occupation of household head, numbers of children, working-age adults, and the elderly) attributes interacted with prices to examine the effects of economic hardship on different family members.[35]

Many previous studies had shown that fertility, mortality, and marriage responded to fluctuations in harvests and food prices in pre-industrial populations. Thomas' pioneering work and Lee's chapter in the Cambridge Group back projection volume said as much. The Eurasia Project took demographic responses for granted, asking who in the family was most affected by hard times in different family systems? Malthusian theory predicts that mortality responses to prices should have been weaker in the "low pressure" European populations, and anthropological theories suggest stronger biases against females in the Asian families. But the results were

---

David R. Cox, "Regression Models and Life Tables," *Journal of the Royal Statistical Society*, LXXIV (1972), 187–220; *idem*, "Partial Likelihood," *Biometrika*, LXII (1975), 269–276.

[34] Bengtsson, Campbell, and Lee, *Life under Pressure: Mortality and Living Standards in Europe and Asia, 1700–1900* (Cambridge, Mass., 2004); Tsuya, Feng, Alter, and Lee, *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700–1900* (Cambridge, Mass., 2010); Christer Lundh and Satomi Kurosu, *Similarity in Difference: Marriage in Europe and Asia, 1700–1900* (Cambridge, Mass., 2014).

[35] Bengtsson, "Combined Time-Series and Life Event Analysis: The Impact of Economic Fluctuations and Air Temperature on Adult Mortality by Sex and Occupation in a Swedish Mining Parish, 1757–1850," in Reher and Schofield (eds.), *Old and New Methods in Historical Demography*, 239–258.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

rarely simple. Overall, death rates in the European communities were as susceptible to high prices as in the Chinese and Japanese case studies. However, status within households tended to be more important in the Asian cases. For example, high prices and competition with other members of the household (such as grandparents) were more harmful to girls (ages two to fourteen) than to boys in Asian families.[36]

The most important differences between Europe and Asia were in the areas of nuptiality and fertility. Malthus was correct in emphasizing the importance of late marriage in Europe, but he dismissed the idea that Asians were controlling family size by regulating their reproduction. Both the Chinese and Japanese families in the Eurasia Project were using infanticide in deliberate ways. Tsuya and Kurosu clearly show that Japanese couples not only limited family size but also selected the order in which boys and girls were born. As a result, the number of births per couple was similar between East and West, but population growth was lower in the Asian populations because of higher mortality. When population growth increased, Europeans often responded with migration before they adopted fertility control, as in our Belgian case study.[37]

Over the course of the Eurasia Project, the Cox model went from novelty to standard procedure in historical demography. The *JIH* supported the diffusion of event history analysis by publishing an issue devoted to papers resulting from the Inter-university Consortium for Political and Social Research (ICPSR) summer course on longitudinal analysis in historical demography created by Gutmann and myself.[38]

---

[36] Dorothy Swaine Thomas, *Social Aspects of the Business Cycle* (New York, 1927); *idem*, *Social and Economic Aspects of Swedish Population Movements, 1750–1933* (New York, 1941); Lee, "Short-Term Variation: Vital Rates, Prices, and Weather," in Wrigley and Schofield (eds.), *The Population History of England, 1541–1871: A Reconstruction*, 356-401; Patrick R. Galloway, "Basic Patterns in Annual Variations in Fertility, Nuptiality, Mortality, and Prices in Pre-Industrial Europe," *Population Studies: A Journal of Demography*, XLII (1988), 275–302; Bengtsson, Campbell, and Lee, *Life under Pressure*, 68–82, 341–343.

[37] Malthus, *Essay on the Principle of Population* (1798 and 1826); Tsuya and Satomi Kurosu, "Family, Household, Reproduction in Northeastern Japan, 1716 to 1870," in Tsuya, Feng, Alter, and Lee (eds.), *Prudence and Pressure*, 249–286; Michel Oris, Alter, and Paul Servais, "Prudence as Obstinate Resistance to Pressure: Marriage in Nineteenth-Century Rural Eastern Belgium," in Lundh and Kurosu (eds.), *Similarity in Difference*, 261–294.

[38] Jon Gjerde and Anne McCants, "Individual Life Chances, 1850–1910: A Norwegian-American Example," *Journal of Interdisciplinary History*, XXX (1999), 377–405; Kevin McQuillan, "Family Composition and Remarriage in Alsace, 1750–1850," *ibid.*, XXXIII (2003), 547–567; Alter, Gutmann, Susan Hautaniemi Leonard, and Emily R. Merchant, "Introduction: Longitudinal Analysis of Historical-Demographic Data," *ibid.*, XLII (2012), 503–517.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

EVENT HISTORY ANALYSIS AS A DEMOGRAPHIC MODEL Event history analysis is usually described in the same way as other regression models with tables of estimated coefficients for explanatory variables. Below the surface, however, event history analysis is built from the same tools that Hajnal, Henry, Coale, Wrigley, and Schofield employed. The debt of the Cox proportional-hazards model to the life table is explicit. Cox himself made the connection in the first of his two seminal articles, "Regression Models and Life Tables." Moreover, proportional-hazards models, like the Cox model, are also a form of indirect standardization.[39]

Event history models are usually expressed in terms of an event's hazard function. Hazard rates are conditional probabilities that an event will occur at a moment in time as measured from some starting point. For example, the risk of dying varies by time since birth (age). It is high immediately after birth, decreases to a minimum in late childhood, and then rises (sometimes irregularly) into old age. Since the risk is not constant, most event history models cannot be subsumed into simpler statistical models assuming constant risks, like a Poisson process or Markov chain. The hazard rate of a life table is a well-known quantity, the "force of mortality," which is designated by μ_x and related to the mortality rate by the equation

$$\mu_x = \lim_{n \to 0} {}_n m_x \tag{5}$$

in which ${}_n m_x$ is the age-specific death rate for ages x to $x+n$.[40]

Estimation procedures for event history models usually involve simplifying assumptions, the most common of which is proportional hazards. The Cox proportional-hazards model simplifies the model as

$$h(t|x_1, x_2, \ \ldots, x_k) = h_0(t)e^{(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)} \tag{6}$$

where

$h(t|x_1, x_2, \ldots, x_k)$ is the hazard rate at time t for an individual with explanatory variables $x_1, x_2, \ldots, x_k$,

---

[39] Gutmann and Alter, "Family Reconstitution as Event-History Analysis"; Cox, "Regression Models and Life Tables."

[40] Preston, Heuveline, and Guillot, *Demography*, 59.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

$h_o(t)$ is hazard function for a standard individual, and

$\beta_1, \beta_2, \ldots \beta_k$ are parameters describing the effects of the explanatory variables.

The Cox model is a proportional-hazards model, because an explanatory variable has the same proportional effect on the hazard rate regardless of time (t). The proportional-hazards assumption in Equation 6 is the same as the proportionality assumption seen in indirect standardization given in Equation 1. Both models assume that a standard set of age-specific rates, f_a^S in Equation 1 and $h_o(t)$ in Equation 6, are modified proportionally by a case-specific index, I^i and $e^{(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_x x_k)}$. In other words, the Cox model is an elaborate form of indirect standardization.

Event history models are also similar to demographic models in the way that they focus on transition rates rather than average duration times. The expectation of life in a life table is equal to the average age at death. Following a cohort from birth until every subject has died works well for studies of fruit flies, but it is rarely feasible for human populations. But we can calculate death rates by age from vital registration and census data, treating these rates as the experience of a hypothetical cohort. The life table is an accounting tool for translating age-specific transition rates into expectations of life and numbers surviving at each age in this hypothetical cohort. Event history analysis faces the same challenge. How can we compute the average survival of cancer patients if some of them are still alive? The answer is to focus on hazard rates, which can be computed at each point for which we have data. Censored cases, patients who have not died, are used in the computation of the hazard rate until the time (duration) when they were last observed. Omitting censored cases results in an overestimation of the hazard rate and an underestimation of the time that survivors lived—the very bias against which Henry warned when he created the rules for family reconstitution.

LESSONS FORGOTTEN The apparent similarity between event history analysis and other regression models has had unfortunate consequences; some authors seem to be unaware of the lessons of classic studies in historical demography. Published fertility studies that include family histories terminated with births or child deaths violate Henry's rules for family reconstitution, producing biased results that are misinterpreted as new findings.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

In a series of articles, Van Bavel claimed to use event history analysis to detect evidence of birth spacing. Van Bavel is not alone in arguing that previous methods for detecting birth control do not reliably distinguish between stopping and spacing. He also criticizes Okun's simulation study for not testing for situations in which both stopping and spacing are present. But Van Bavel's alternative method for discovering birth spacing goes awry by focusing on closed birth intervals: "Secondly, a duration model could be developed for all closed birth intervals, representing the determinants of birth spacing and again including age and attained parity as two of the covariates." As he explained his intended analysis, "In this paper, the question asked is: given that another birth has occurred, what are the covariates of the length of the past birth interval?" Unfortunately, his chosen method, Cox regression, does not answer that question. In fact, Cox regression does not look at the lengths of birth intervals at all. It analyzes transition rates. By including only closed birth intervals, Van Bavel omitted time at risk that should have been included in those rates. Selecting intervals that end in a birth, the event of interest, is a clear example of informative censoring, which leads to biased results.[41]

Event history models are estimated from duration-specific transition rates that compare the number of events to the number of women at risk at a moment in time. The Cox model ignores time altogether by using the order of events without considering the amount of time between them. By computing duration-specific rates, event history models can include information from censored histories—histories that have not experienced the event.

---

[41] Jan Van Bavel and Jan Kok, "Birth Spacing in the Netherlands: The Effects of Family Composition, Occupation and Religion on Birth Intervals, 1820–1885," *European Journal of Population*, XX (2004), 119–140; *idem*, "The Role of Religion in the Dutch Fertility Transition: Starting, Spacing, and Stopping in the Heart of the Netherlands, 1845–1945," *Continuity and Change*, XX (2005), 247–263; *idem*, "A Mixed Effects Model of Birth Spacing for Pre-Transition Populations: Evidence of Deliberate Fertility Control from Nineteenth Century Netherlands," *History of the Family*, XV (2010), 125–138; Van Bavel, "Deliberate Birth Spacing before the Fertility Transition in Europe: Evidence from Nineteenth-Century Belgium," *Population Studies—a Journal of Demography*, LVIII (2004), 95–107; *idem*, "Detecting Stopping and Spacing Behaviour in Historical Demography: A Critical Review of Methods," *Population* (English Edition), LIX (2004), 117–128; Okun, "Evaluating Methods for Detecting Fertility-Control—Coale and Trussells Model and Cohort Parity Analysis," *Population Studies—a Journal of Demography*, XLVIII (1994), 193–222; Douglas L. Anderton and Lee L. Bean, "Birth Spacing and Fertility Limitation—a Behavioral-Analysis of a 19th-Century Frontier Population," *Demography*, XXII (1985), 169–183; Okun, "Distinguishing Stopping Behavior from Spacing Behavior with Indirect Methods," *Historical Methods*, XXVIII (1995), 85–96.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

If we are studying birth intervals, the denominator of the hazard rate should include all women who were at risk of a birth, including those whose birth intervals were censored by migration or sterility. This piece-wise approach to the hazard function, which is the way that life tables are constructed, underlies the computation of age-specific fertility rates in family reconstitution. When informative censoring is correlated with explanatory variables, such as age and birth order, these variables do not "control" for the effects of those variables; rather, the estimated coefficients of these variables reflect the biases in the data. Omitting incomplete birth intervals does not make the model estimate the lengths of completed birth intervals; it produces biased estimates of the hazard rates.[42]

We can see the consequences of violating the rules of family reconstitution in Table 1, which displays age-specific fertility rates computed from family reconstitutions based on six German village genealogies that Knodel collected. The first two columns, labeled "Henry rules," are computed according to the strict rules in the Henry manuals. Families are included only if we observe the death of at least one spouse or another event unrelated to childbearing, such as a census or divorce. The columns labeled "incomplete histories" refer only to families that are excluded by the Henry rules. "Incomplete histories" end with a birth or the death of a child, because the family moved somewhere else before one of the spouses died. The columns labeled "closed intervals" include all intervals ending in a birth from both the "Henry rules" and the "incomplete histories" columns.[43]

Violating the rules of family reconstitution has dramatic effects on the calculation of age-specific fertility rates, especially at older ages. The rates computed correctly show fertility decreasing with age. The fertility transition after 1875 reduces rates at older ages more than at younger ages. Both these features are absent when the rules are violated. As Henry warned, the inclusion of birth intervals with informative censoring overestimates birth rates. The bias is greatest at older ages, because more time at risk is lost due to incomplete intervals.

Event history methods, including the Cox proportional-hazards model (see Table 2), produce the same biased results as the simpler calculations in Table 1. When family histories end with informative

---

[42] Janet M. Box-Steffensmeier and Bradford S. Jones, *Event History Modeling: A Guide for Social Scientists* (New York, 2004), 51.

[43] Knodel, *Demographic Behavior in the Past: A Study of Fourteen German Village Populations in the Eighteenth and Nineteenth Centuries* (New York, 1988).

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Table 1 Age-Specific Fertility Rates in Six German Village Genealogies, Calculated in Three Ways, by Time Period

	HENRY RULES		INCOMPLETE HISTORIES		CLOSED INTERVALS	
Age	Before 1875	After 1875	Before 1875	After 1875	Before 1875	After 1875
15–19	0.400	0.383	0.596	0.711	0.443	0.589
20–24	0.454	0.428	0.559	0.472	0.498	0.539
25–29	0.396	0.356	0.472	0.384	0.454	0.471
30–34	0.343	0.263	0.431	0.341	0.424	0.420
35–39	0.267	0.192	0.435	0.372	0.415	0.444
40–44	0.137	0.078	0.520	0.449	0.499	0.527
45–49	0.019	0.008	0.463	0.567	0.784	0.928

SOURCE Alter, "From Data Scarcity to Data Abundance," distributed by the Inter-university Consortium for Political and Social Research, University of Michigan, Ann Arbor (March 29, 2019), available at https://doi.org/10.3886/E109127V1.

censoring, the estimated coefficients for age do not decrease as we know that they should. When only closed birth intervals enter the analysis, the estimated coefficients in Table 2 imply that the birth rates of women over age forty-five were the same as those of women aged twenty-five to twenty-nine. This bias will also affect other variables if they are correlated with the process creating informative censoring. Since birth order (parity) is highly correlated with age, we should expect birth rates and hazard rates computed from data with informative censoring to be biased too. Evidence of this bias is in the Cox regression estimates published by Van Bavel. In his 2004 study, the relative risk of a birth after age forty compared to that before age twenty-five is reported as 0.56 for the 1830 generation, 1.03 for the 1850 generation, and .94 for the 1864 generation, meaning that women older than forty were having children at the same rate as those younger than twenty-five. Van Bavel and Kok suggest that the absence of age effects in this model is due to the control of marital duration, which is an indicator of coital frequency, but Wood estimates that most of the decrease in fecundability after age thirty-five is due to physiological changes.[44]

When we understand that the Cox model compares duration-specific rates and not the lengths of completed intervals, Van Bavel's

---

[44] Van Bavel, "Deliberate Birth Spacing before the Fertility Transition in Europe," 103; *idem* and Kok, "Birth Spacing in the Netherlands," 131; James W. Wood, *Dynamics of Human Reproduction Biology, Biometry, Demography* (New York, 1994), 315.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Table 2 Relative Risks of a Birth in Six German Village Genealogies, by Age, from Cox Proportional-Hazards Regression, Calculated in Three Ways, by Time Period

	HENRY RULES		INCOMPLETE HISTORIES		CLOSED INTERVALS	
Age	Before 1875	After 1875	Before 1875	After 1875	Before 1875	After 1875
15–19	1.00	0.53	3.10	2.00	1.04	0.97
20–24	1.17	1.07	1.20	1.13	1.19	1.09
25–29	ref.	ref.	ref.	ref.	ref.	ref.
30–34	0.84	0.78	0.83	0.92	0.86	0.88
35–39	0.63	0.64	0.78	0.97	0.75	0.91
40–44	0.31	0.30	0.70	1.12	0.71	0.96
45–49	0.07	0.06	0.61	1.09	0.94	1.53
Observations	78,868	47,942	3,458	12,556	64,566	39,575
Births	14,496	8,110	827	3,168	15,101	9,945
Time at risk	57,285	43,327	1,950	9,652	36,869	24,007

SOURCE Alter, "From Data Scarcity to Data Abundance," distributed by the Inter-university Consortium for Political and Social Research, University of Michigan, Ann Arbor (March 29, 2019), available at https://doi.org/10.3886/E109127V1.

claim that he is measuring spacing falls apart. As the next section demonstrates, the Cox model cannot distinguish between stopping and spacing, and a change in stopping usually violates the proportional-hazards assumption in the model. Unfortunately, other scholars—such as Cinnirella, Klemp, and Weisdorf—have adopted this erroneous interpretation of the Cox model, basing their analyses on only closed birth intervals.[45]

Avoiding informative censoring does impose substantial costs on family reconstitution, because Henry's rules typically remove between 40 and 60 percent of family histories from analysis. Fialova, Tesarkova, and Kuprova compared the strictest version of Henry's selection rules to an "extended" sample in which the family history ends with a birth only when no other event is available. Their data set drops from 5,588 to 1,290 observations when the stricter rules are applied. They see the same patterns in both samples, but the magnitudes of coefficients estimated by the Cox model are sometimes vastly different. Alternatives to dropping incomplete

---

[45] Francesco Cinnirella, Marc Klemp, and Jacob Weisdorf, "Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England," *Demography*, LIV (2017), 413–436.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

family histories involve much more complicated computations. Jonker and van der Vaart developed a maximum likelihood approach, and Alter, Devos, and Kvetko show that migration dates can be imputed in some situations.[46]

These comments are not meant to reject the idea that intentional postponement of births may have been important in historical populations (see, for example, Reher and Sanz-Gimeno, "Rethinking Historical Reproductive Change"). The point is methodological not substantive. Van Bavel has done innovative work on the fertility transition and other topics, and his hypotheses about associations between birth spacing and other variables may turn out to have merit. But we can evaluate those hypotheses only when our methods fit both the data and the question. Publications based on these flawed methods betray a loss of knowledge previously shared by historical demographers, as well as a misunderstanding of the new tools that event history analysis has provided. The old tools, like the life table, can guide us in using new tools more effectively.[47]

MOVING FORWARD WITH OLD AND NEW TOOLS The issue of stopping and spacing provides an example of the benefits of combining old and new models of historical demography. Figure 1, which illustrates Hajnal's singulate mean age at marriage also shows how to distinguish stopping and spacing in fertility histories. Compare Figure 1 to Figure 4, both of which are survival curves. In Figure 1, the height of the curve is the proportion of women who survived from birth to each age without marrying. The curve in Figure 4 is the proportion of women who have not yet given birth to their next child. The methods used to obtain the curves are different. Figure 2

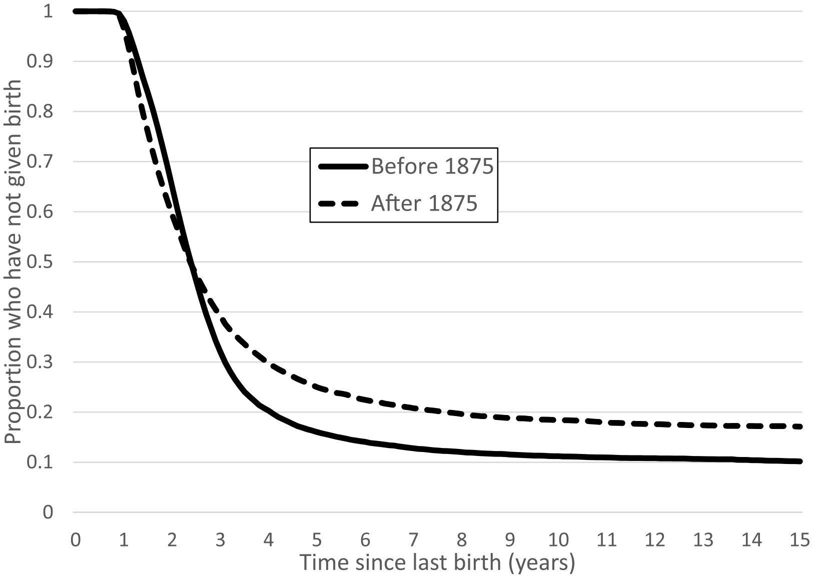
Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[46] Ludmila Fialova, Klara H. Tesarkova, and Barbora Kuprova, "Determinants of the Length of Birth Intervals in the Past and Possibilities for Their Study: A Case Study of Jablonec Nad Nisou (Czech Lands) from Seventeenth to Nineteenth Century," *Journal of Family History*, XLIII (2018), 127–156; Marianne A. Jonker and Aad W. van der Vaart, "Estimation of Average Mortality under Censoring and Truncation," *Journal of Population Research*, XXII (2005), 49–62; Jonker and van der Vaart, "Correcting Missing-Data Bias in Historical Demography," *Population Studies-a Journal of Demography*, LXI (2007), 99–113; Alter, Isabelle Devos, and Alison Kvetko, "Completing Life Histories with Imputed Exit Dates: A Method for Historical Data from Passive Registration Systems," *Population*, LXIV (2009), 327–353.

[47] Reher and Sanz-Gimeno, "Rethinking Historical Reproductive Change: Insights from Longitudinal Data for a Spanish Town," *Population and Development Review*, XXXIII (2007), 703–727; Van Bavel, "Diffusion Effects in the European Fertility Transition: Historical Evidence from within a Belgian Town (1846–1910)," *European Journal of Population*, XX (2004), 63–85.

Fig. 4 Kaplan-Meier Survival Curves for Women in Six German Village Genealogies Who Have Not Given Birth, by Time since Last Birth and Time Period



SOURCE Alter, "From Data Scarcity to Data Abundance," distributed by the Inter-university Consortium for Political and Social Research, University of Michigan, Ann Arbor (March 29, 2019), available at https://doi.org/10.3886/E109127V1.

plots percentages by marital status from a census. The survival curves in Figure 4 were estimated with the Kaplan-Meier procedure. But both figures show processes in which a part of the population never experiences the event of interest, that is, marriage or childbirth.[48]

The fact that some people never marry or have another birth distinguishes nuptiality and fertility from the analysis of mortality and from most event-history methods. The curves in both Figures 1 and 4 become flat before reaching zero. Hajnal used the flat part of the curve to measure the percentage of women who will never marry. We can use the height of the curve in Figure 4 to measure stopping, and, like Hajnal, we can adjust the curve to measure spacing by subtracting the proportion who stop. Using the Kaplan–Meier estimate of uncompleted birth intervals after fifteen years as our measure of stopping, we can compute an adjusted survival

---

[48] Edward L Kaplan and Paul Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, LIII (1958), 457–481.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Table 3 Measures of Stopping and Spacing in Six German Village Genealogies

	PERCENTAGE OF BIRTH INTERVALS NOT COMPLETED		AVERAGE LENGTH OF COMPLETED BIRTH INTERVALS	
	BEFORE 1875	AFTER 1875	BEFORE 1875	AFTER 1875
All	10%	17%	2.64	2.65
Number of previous births				
1	3%	8%	2.36	2.44
2	5%	14%	2.52	2.66
3	7%	15%	2.61	2.73
4	8%	18%	2.60	2.76
5	13%	20%	2.61	2.75
6	14%	25%	2.68	2.56
7	22%	25%	2.70	2.49

SOURCE Alter, "From Data Scarcity to Data Abundance," distributed by the Inter-university Consortium for Political and Social Research, University of Michigan, Ann Arbor (March 29, 2019), available at https://doi.org/10.3886/E109127V1.

curve for completed birth intervals. The computation of average birth intervals is slightly different from the way that SMAM is estimated, because the Kaplan–Meier curve corresponds to the number of survivors at an exact age (l_x) in the life table. Results for the data in Figure 4 are shown in Table 3.

Table 3 suggests that the German fertility transition involved an increase in stopping from 10 percent to 17 percent of all birth intervals with no change in the average birth interval. But this implication is somewhat misleading, because stopping also changes the distribution of birth intervals by birth order and age of mother. Women who stop after two children are never at risk of third, fourth, or fifth births. For this reason, it is more informative to analyze birth intervals separately by parity, as in the work of Reher, Sandstrom, Sanz-Gimeno, and van Poppel for historical data and Casterline and Odden for contemporary data. The picture is a little different when we look at stopping and spacing by number of previous births. Stopping increased more among women with two to six children than among those with one or seven children. Average lengths of birth intervals increased slightly at lower parities, but birth intervals became shorter for women with six or seven children.[49]

---

[49] Reher, Glenn Sandstrom, Sanz-Gimeno, and van Poppel,"Agency in Fertility Decisions in Western Europe During the Demographic Transition: A Comparative Perspective,"

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Figure 4 offers a warning that common event-history methods cannot distinguish between stopping and spacing. The Cox model has no way of identifying stopping and spacing separately. Assuming that the baseline hazard rate goes to zero at some point (the survival curve becomes horizontal), a proportional increase in the hazard rate will result in shorter birth intervals and fewer open intervals—in other words, less spacing and less stopping. A decrease in the hazard rate at all durations will increase spacing and increase stopping. The model does not allow stopping and spacing to move independently, as evident in Figure 4 and Table 3. Moreover, the survival curves in Figure 4 cross, which cannot happen if hazard rates change proportionally at all durations. In other words, birth intervals were changing in ways that violated the proportional-hazards assumption in the Cox model.[50]

Fortunately, there are models with explanatory variables that affect stopping and spacing separately. The cure model, also known as the split population model, is a type of mover–stayer model with an event-history component. A new medical treatment can "cure" members of a population with the result that they will not die of the disease. Even those who die, however, may live longer because of the treatment. The cure model has two branches for estimating the effects of explanatory variables on the "cure fraction" and on the timing of transitions separately. The "cure fraction" is analogous to the proportion who never marry or never have another birth. A few applications of the cure model have appeared in demography, including Janssens' work in historical demography.[51]

Demography, LIV (2017), 3–22; John B. Casterline and Colin Odden, "Trends in Inter-Birth Intervals in Developing Countries 1965–2014," *Population and Development Review*, XLII (2016), 173–194.

[50] Reher, Sandstrom, Sanz-Gimeno, and van Poppel, "Agency in Fertility Decisions," 10.

[51] Edith Gray, Ann Evans, Jon Anderson, and Rebecca Kippen, "Using Split-Population Models to Examine Predictors of the Probability and Timing of Parity Progression," *European Journal of Population*, XXVI (2010), 275–295; Lei Li and Minja K. Choe, "A Mixture Model for Duration Data: Analysis of Second Births in China," *Demography*, XXXIV (1997), 189–197; Kazuo Yamaguchi, "Mover-Stayer Models for Analyzing Event Nonoccurrence and Event Timing with Time-Dependent Covariates: An Application to an Analysis of Remarriage," *Sociological Methodology*, XXVIII (1998), 327–361; *idem* and Linda R. Ferguson, "The Stopping and Spacing of Childbirths and Their Birth-History Predictors—Rational-Choice Theory and Event-History Analysis," *American Sociological Review*, LX (1995), 272–298; Janssens, *Labouring Lives: Women, Work and the Demographic Transition in the Netherlands, 1880–1960* (Bern, 2014).

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

Another promising development in historical demography is the introduction of sequence analysis. Sequence analysis describes a life history as a trajectory composed of states marked by important events (starting school, leaving home, marriage, childbirth, etc.). Differences between individuals can be ranked in terms of the similarity of the types, order, and timing of events in their lives, permitting the identification of individuals with similar experiences. Unlike event-history analysis, which focuses on a single event, sequence analysis considers all the events in a life history. Bras and Schumacher adopted this approach to study the effect of age gaps between husbands and wives on childbearing trajectories. They find that relationships in which the partners' ages were roughly equal were more conducive to fertility control.[52]

In 1996, Coale and Trussell reviewed the development of demographic models and commented on recent changes in demographic research:

> The decline in the use of demographic models coincides with the increasing availability of the wealth of survey data collected by the World Fertility Survey and the successor Demographic and Health Surveys. These data are frequently used—without any comparison with demographic models to assess their validity—to analyse individual, not aggregates behaviour. The techniques demographers use to analyse these data are increasingly the techniques of statistics, particularly event-history analysis. While the trend toward the use of sophisticated statistical models appropriate for the problem to be analysed and the data that are available is healthy, the trend toward accepting demographic survey data at face value is not.[53]

A similar development was occurring in historical demography at the same time—a trend away from aggregate analysis based

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025

---

[52] Gilbert Ritschard and Michel Oris, "Life Course Data in Demography and Social Sciences: Statistical and Data-Mining Approaches," in René Levy et al. (eds.), *Towards an Interdisciplinary Perspective on the Life Course* (Boston, 2005), 283–314; Andrew Abbot, "Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes," *Historical Methods*, XVI (1983), 129–147; Hilde Bras and Reto Schumacher, "Changing Gender Relations and Declining Fertility: An Analysis of Spousal Age Differences and Women's Childbearing Trajectories in Zeeland, the Netherlands, 1812–1911," *Demographic Research* (forthcoming).

[53] Coale and Trussell, "The Development and Use of Demographic Models," *Population Studies—a Journal of Demography*, L (1996), 469–484.

on demographic models toward statistical analysis of individual-level data. As a participant in, and promoter of, this trend, I am not arguing against it. New questions, data, and techniques are healthy developments in any field of inquiry. However, Coale and Trussell are correct to point out that much is lost if we forget the lessons of earlier generations. Demographic models can provide deeper insight into statistical models. We should be particularly skeptical of research that emphasizes statistical significance for results that make no demographic or biological sense.

The classic generation of historical demography left an important legacy. Their work is worth serious study not only for its important results but also for its sophisticated understanding of demographic processes and its obsession with data quality.

Downloaded from http://direct.mit.edu/jinh/article-pdf/50/3/325/1693611/jinh_a_01445.pdf by guest on 21 March 2025