



1985

Graphics in Demography

Susan Cotts Watkins

Recommended Citation

Watkins, S. C. (1985). Graphics in Demography. *11* (3), 2-21. Retrieved from <https://repository.upenn.edu/svc/vol11/iss3/2>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/svc/vol11/iss3/2>
For more information, please contact repository@pobox.upenn.edu.

Graphics in Demography

Graphics in Demography

Susan Cotts Watkins

Demography can be fruitfully viewed as the study of entries into and exits from a population. The means of entry or exit may be solely by birth or death, as when the population under consideration is that of the world, or, for smaller populations, they may include migration. The population may also be defined as bounded by other characteristics—women of child-bearing age, retired workers, the readers of a particular journal—and then birth, death, and migration must be appropriately defined (e.g., new subscribers.) The size of the population is determined by the balance between entries and exits; its composition is determined by the relevant characteristics of the initial population and the rate at which new members with these characteristics enter and old ones depart.

Since its inception, demography has been centrally concerned with the search for similarities in patterns of entry and exit; the explication of formal relations among mortality, fertility, and the age composition of a population; and, particularly recently, the determinants and consequences of population growth. Graphical representations have been used in all three quests. The application of graphical methods to expose regular patterns in demographic behavior and to analyze the necessary relations among the formal determinants of population composition is seen as early as a 1693 paper by Edmund Halley, preceded only (and only by a few decades) by John Graunt as a major figure in the history of mathematical demography. It continues through Nicander, Milne, and Quetelet, in the nineteenth century, and Lotka, in the early twentieth century, to those who developed the model schedules of mortality, nuptiality, and fertility of the past few decades. The discovery of empirical regularities has permitted the use of graphical techniques in less sublime, even pedestrian, tasks: when demographic data are either lacking or thought to be untrustworthy, as is often the case for both historical populations and those in contemporary developing countries, the existence of known regularities makes it possible visually to compare actual with expected values, and thus to correct what appear to be faulty data.

Graphics, often imaginative ones, have also been used didactically to emphasize differences rather than commonalities. In the late-nineteenth and early twentieth century, demographers became more

interested in differentials than in similarities, as they attempted to describe and understand the contemporary declines in mortality and fertility in Europe and the United States, and the effect of migration on their societies. Currently, graphics are being used in conjunction with the computer to explore and summarize large quantities of demographic information: the Princeton European Fertility project, a study of the demographic transition in more than six hundred provinces of Europe, provides several innovative examples of computer graphics.

Regularities and Differences: Mortality and Age Composition

In demography, practical and theoretical concerns have often been closely interwoven. That mathematical demography began in the seventeenth century with a concern for formalizing and generalizing estimates of the probability of dying at any given age is plausible in the context of a society in which a surge of interest in science and mathematics, associated with Puritanism, was put to the service of attempts to protect survivors from the consequences of high and unpredictable mortality. The estimation of the probability of dying—and its complement, the probability of surviving—are critical for the calculation of both insurance and annuities. Thus, an early interest in the field of demography was the search for a "law of mortality," a single pattern of mortality that would describe the risk of dying at every age and would apply to all populations. The interest of the early demographers in calculating precisely the probabilities of dying by age was related to a second practical concern, that of estimating the number of persons in a given category (for example, the number of men of military age) at a time when there were few attempts to enumerate populations and little reason for confidence in the results. Graunt and Halley, in the late-seventeenth century, as well as others after them recognized that the shape of the survivorship function, formed by the depletion of a birth cohort by death as it aged, was one determinant of the proportion of the population in an age interval.

John Graunt's "Natural and Political Observations upon the Bills of Mortality" (1662) has been described as "the first substantive demographic work to have been written" (Smith and Keyfitz 1977:1). Even though the Bills of Mortality for London provided numbers of deaths by cause rather than by age, by distinguishing causes of death common to childhood (such as teething and measles) and causes of death common to the aged (chronic rather than acute illnesses) Graunt was able to establish the crude outlines of the relationship of mortality risks to age:

Susan Cotts Watkins is an Assistant Professor in the Department of Sociology at the University of Pennsylvania, and an Associate of the Population Studies Center. She has written about changes in marriage and in marital fertility in Europe during the last century.

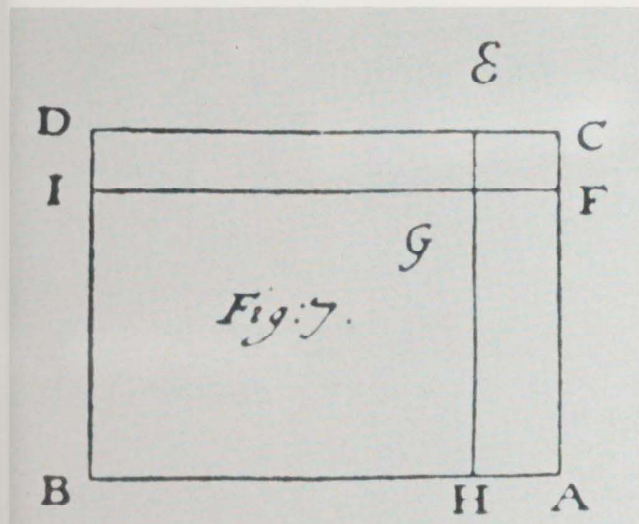


Figure 1 Halley's 1693 Illustration of a Method for Calculating Annuities During Two Lives. From Edmund Halley, *Degrees of Mortality of Mankind* (1942 [1693]).

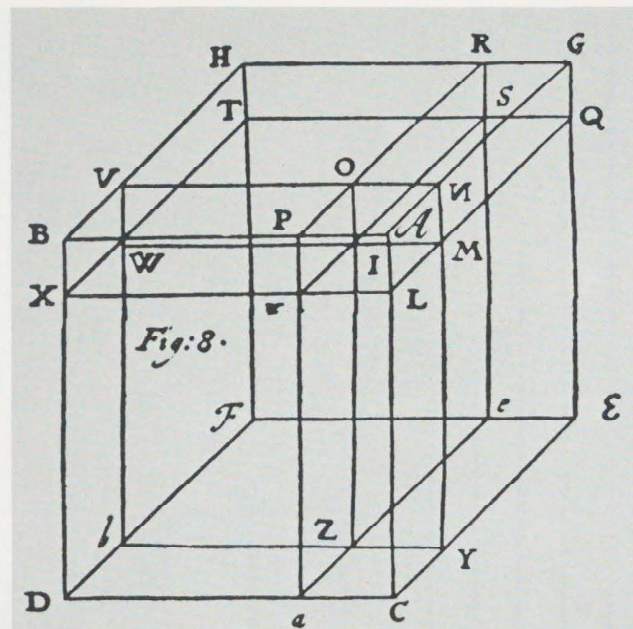


Figure 2 Halley's 1693 Illustration of a Method for Calculating Annuities During Three Lives. From Edmund Halley, *Degrees of Mortality of Mankind* (1942).

Whereas we have found that, of 100 quick Conceptions about 36 percent of them die before they be six years old, and that perhaps but one surviveth 76, we, having seven Decads between six and 76, we sought six mean proportional numbers between 64, the remainder, living at six years, and the one, which survives 76, and finds that the numbers following are practically near to the truth; for men do not die in exact Proportions, nor in Fractions. . . .” [Graunt 1939 (1662):69]

Graunt thus established numerically the relative importance of deaths in infancy and childhood compared to deaths at other ages in a population that suffered from high mortality, a finding equally valid today.

Following Graunt, Edmund Halley used lists from Breslau, Poland, of numbers of births and of deaths by age to construct an age-specific mortality schedule (Halley 1942 [1693]). He used graphs to illustrate his discussion of a method for resolving the practical problem of calculating annuities. Figure 1, the earliest graph I have found in the work of demographers, shows Halley's method for estimating the probability that both or neither of a couple would survive to old age. Halley considers only two age groups, old and young. AB or CD represents the number of persons at the younger age, and DE or BH their survivors after a number of years; AC or BD represents the number

at the older age, and AF or BI the survivors after the same number of years. Thus the rectangle BHGI represents the probability of joint survival, and CEGF the probability that neither of the two lives survives. With Figure 2, Halley's perspective takes on a third dimension as he extends his discussion of the problems involved in calculating annuities for the continuance of three lives.

Like Graunt, Halley described in writing but not graphically a survivorship function not dissimilar in shape to the ones with which we are familiar in contemporary populations suffering from high mortality. Nicander seems to have been the first person to describe graphically the line traced by observed age-specific mortality rates (Nicander 1801). Using information on deaths by age collected by Wargentini, he pictured the progressive diminution by mortality of a birth cohort as it ages. A similar presentation was made by Joshua Milne in his two-volume *A Treatise on the Valuation of Annuities and Assurances on Lives and Survivors* (1815). Figure 3 occurs in the course of a discussion of the construction of mortality tables; the problem is to estimate a continuous line representing survivorship when the only information on deaths by age is for an aggregate age group, so that interpolation between the boundaries of the age group is required to give mortality risks by single years of age.

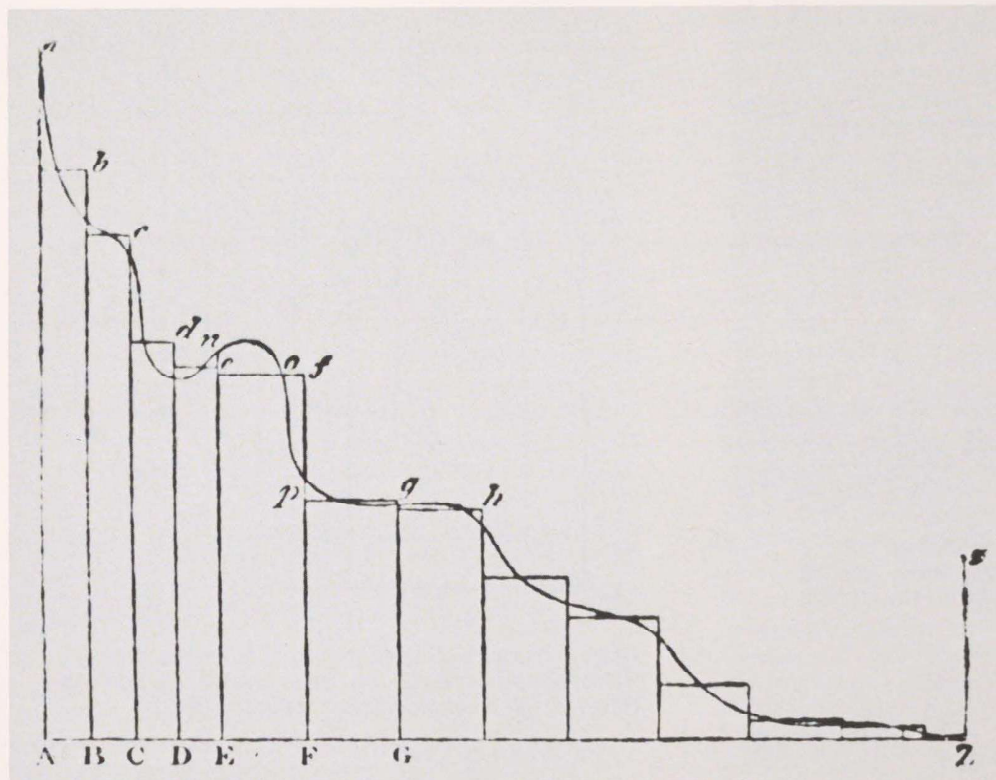


Figure 3 Milne's 1815 Description of the Line Traced by Observed Age-Specific Mortality Rates. From Joshua Milne, *A Treatise on the Valuation of Annuities and Assurances on Lives and Survivorships; and on the Construction of Tables of Mortality; and on the Probabilities and Expectations of Life* (1815), vol. 1, p. 101.

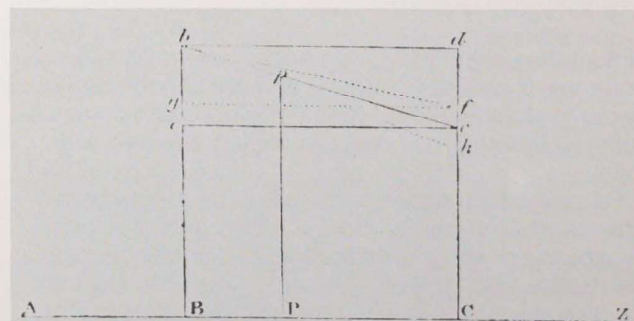


Figure 4 Milne's 1815 Estimate of Person-Years for the Calculation of Mortality Rates. From Joshua Milne, *A Treatise on the Valuation of Annuities and Assurances on Lives and Survivorships; and on the Construction of Tables of Mortality; and on the Probabilities and Expectations of Life* (1815), vol. 2, p. 725.

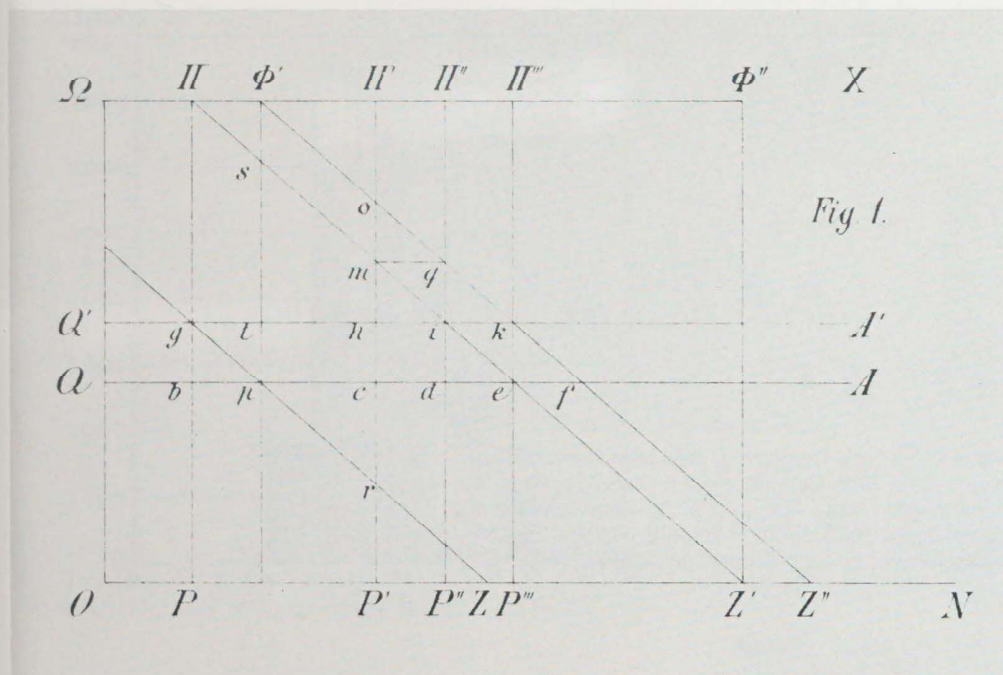


Figure 5 Lexis Diagram, 1875. From Wilhelm Lexis, *Enleitung en die Theorie der Bevölkerungsstatistik* (1875), end page.

Along the abscissa, AZ, the age intervals AB, BC, CD, and so on, are marked off; the width of each is proportional to the ratio of the number living in each interval to the number living in the first interval. The parallelograms are similarly constructed. This done, Milne describes the survivorship function:

Next, let a line (as little curved as the other conditions will admit of) be described through these parallelograms, so that the point describing it, in its motion from the first ordinate, Aa, may continually approach towards the last Zz, and may never change its direction abruptly, so as to form an angle in its path. [Milne 1815:101–102]

The areas of the parallelograms form the basis for the calculations of the number living and the number dying at any single year of age. At about the same time J. B. J. Fourier used a similar graph to demonstrate the calculation not only of age distribution of the survivors of a birth cohort but also of other measures now commonly used to summarize some of the demographic characteristics of a population: the average duration of life, the probable duration of life at every age, and the proportion above or below any given age (Fourier 1821.)

The calculation of demographic rates, in which the numerator is the event of interest and the denominator is person-years of exposure to the risk of those

events, requires a method of calculating person-years at risk. Milne used a graphical presentation, shown in Figure 4, to show the proper calculations. Again, the event of interest is death. Milne begins with the assumptions that there is no migration and that births and deaths are equal; thus, the population is stationary. He has previously noted that the force of mortality is not equal at all periods of life. He suggests that a life table provide exact rates for each of the first three months from birth (when risks of death drop swiftly), and then for each of the remaining three quarters of the first year; from age one to age five single-year intervals are adequate, and above age five intervals of five or even ten years are generally satisfactory. The intervals are apparently chosen by Milne such that within each interval deaths can be assumed to be evenly distributed, since a final assumption is that in the interval "the decrements of life are uniform, equal numbers of persons dying in all the equal intervals of age" (Milne 1815:725).

The lines Bb and Cc are drawn to be proportional to the numbers entering and finishing the interval, respectively. Proceeding with a geometrical argument, Milne concludes that

it will follow that on the hypothesis of the decrements of life being equal during any period of life, the number of persons constantly living in that period will be an arith-

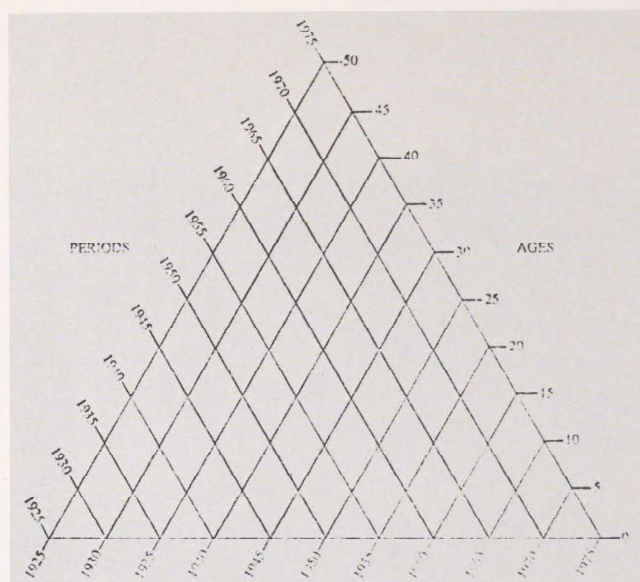


Figure 6 Ryder's Diagram of Available Reproductive Histories for Birth Cohorts, from a Cross-Sectional Survey of Women with Upper Age Limit of 50, Interviewed at the Beginning of 1975. From Norman B. Ryder, "Fertility Measurement through Cross-Sectional Surveys," *Social Forces* (1975), vol. 54, p. 18.

metrical mean proportional between the numbers who enter upon, and who survive the same period during the portion of time expressed by the difference of the ages at the commencement and the end thereof. [ibid.:726]

The area of the trapezium BbCc, then, represents the number of person-years lived during the interval BC.

Milne notes that if there were a change in mortality conditions so that one cause of death—smallpox, for example—were to be eliminated, then the survivorship function must be redrawn (bf rather than bc) and the number of living in the interval would be greater:

But the triangle bcf, which is the excess of the trapezium Bbfc above the trapezium BbCc, is equal to half the rectangle ef: so that the increase in the number of the living in the period of life BC, in consequence of the extermination of that disease, would be equal to half the number of deaths it occasioned in the corresponding period of time when it prevailed. [ibid.:727]

The graphical treatment of deaths and populations in given age groups, used by Nicander and then by Milne and Fourier, subsequently became one of the standard methods for the construction of life tables.

A particular kind of graphical technique has been used in the analysis of the proper measurement of exposure time for a group considered to be exposed to the risk of an event (such as death) when membership in the group is defined not only by its age at the time of observation but also by its location in a specific cohort (e.g., marriages in a given year). The prototype is a Lexis diagram, first used by Wilhelm Lexis

(1875) to discuss aggregate mortality data (Figure 5). Date of birth is measured along the abscissa (0 is the present and the points P, P', etc., up to N, are all dates further in the past); age is measured along the ordinate, with the highest age attained in the population measured by the line O-Ω. A person born at a date along the line ON (e.g., at z) follows a 45-degree trajectory toward O-Ω, although this life line may be interrupted by death (death points are not shown).

Lexis demonstrates that two lines, age and time, define a plane which can be divided in three ways, corresponding to the three ways in which the population can be seen to be structured: (1) the vertical line P-pie divides the deaths occurring before a certain date from those occurring after that date; (2) the horizontal line QA separates those who die before one age from those who die after it; (3) the diagonal (e.g., Z'-π) separates the histories of those born before Z' from those born after that date.

If a survey were taken of individuals at date Z, and the cohorts to which they belong were subsequently followed as they aged, only cohorts born after Z' could be observed for their entire life span; all their combinations of age and time fall below the diagonal Z'-π and will be observed; all those individuals born at a specific point, say Z'', before Z' are observed only for ages above their age at Z', so that the combinations of age and time above the diagonal are the observable parts of that cohort's history. If we then look at age at death for each cohort, and base the average on observable data, it will be biased, for it will take into account only those who survived to a

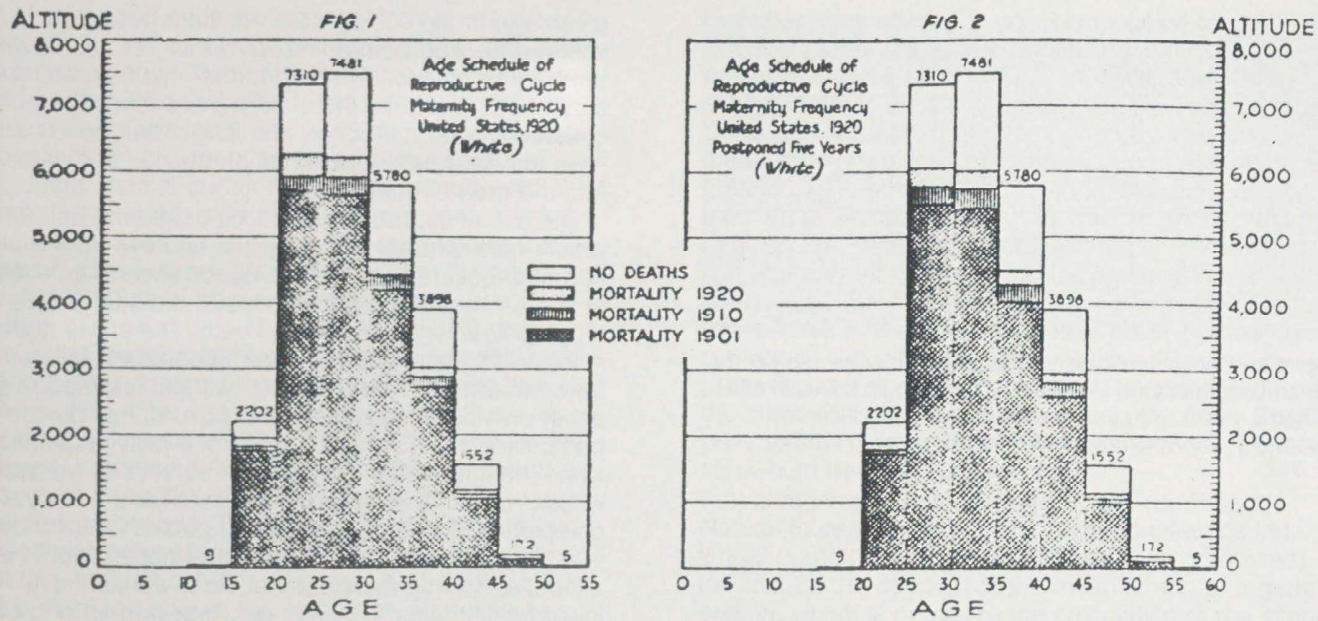


Figure 7 Dublin and Lotka's 1925 Histogram of Total Births of Daughters from a Cohort of 100,000 Females Starting Life Together. From Louis I. Dublin and Alfred J. Lotka, "On the True Rate of Natural Increase," *Journal of the American Statistical Association* (1925), p. 308.

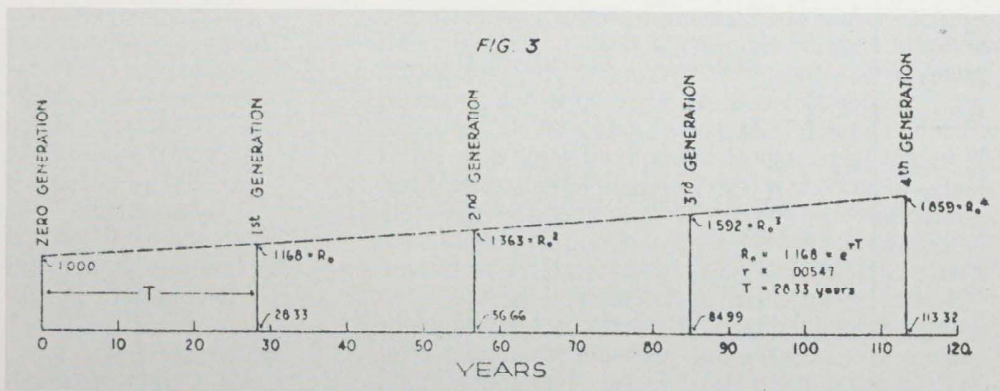


Figure 8 Dublin and Lotka's Diagram Illustrating the Definition of the Mean Length of One Generation. From Louis I. Dublin and Alfred J. Lotka, "On the True Rate of Natural Increase," *Journal of the American Statistical Association* (1925), p. 310.

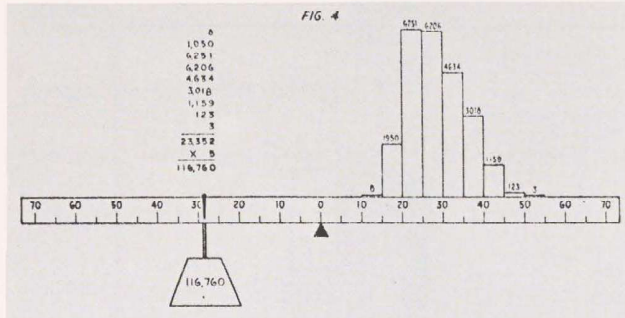


Figure 9 Dublin and Lotka's Mechanical Interpretation of the Mean Length of One Generation. From Louis I. Dublin and Alfred J. Lotka, "On the True Rate of Natural Increase," *Journal of the American Statistical Association* (1925), p. 312.

given age at the survey date. All those who died at young ages and before time Z' will be omitted. It will appear that the cohorts born further into the past are surviving longer than those born more recently because the older the cohort, the longer that cohort will have survived before observation begins, and therefore the greater the bias.

The use of life tables and Lexis diagrams, both of which make precise the entry into and exit from a period of exposure to risk of the occurrence of an event of interest, has been generalized to the study of demographic events other than death. The ability to demarcate segments of exposure, such as the interval between one birth and another, without knowledge of either previous or successive exposure, has proved particularly valuable in the study of fertility, especially when the information comes from surveys of women whose marital and reproductive histories are not yet complete. Bias can be avoided if pooled information is restricted to the same segment of observation. For example, we might expect that older women have lower fertility rates than younger, that women who be-

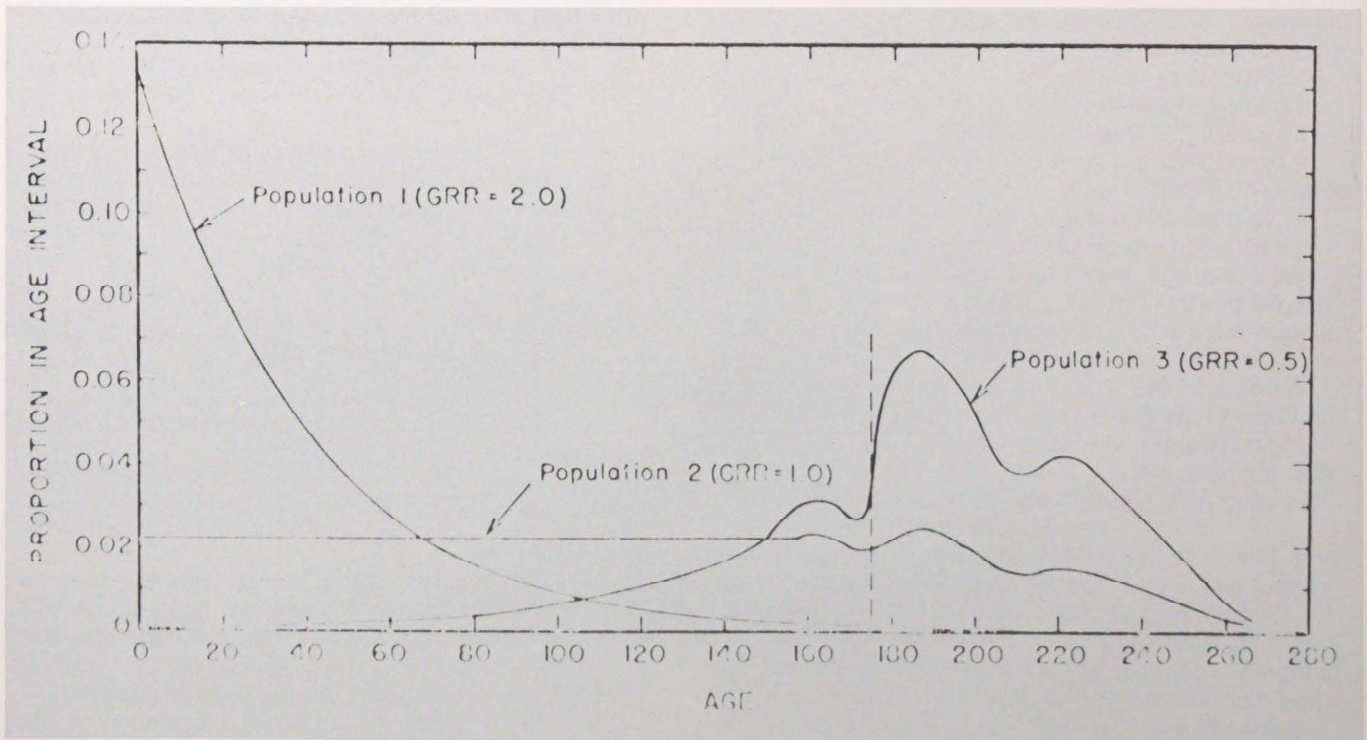


Figure 10 Age Distribution of the Projected Female Population of the United States. Projected from 1960 for 175 Years with No Mortality. From Ansley J. Coale, "Age Composition in the Absence of Mortality and in Other Odd Circumstances," *Demography* (1973), vol. 10, p. 538.

gin childbearing early are different from those who postpone it, and that women who bore children after the introduction of the contraceptive pill bore fewer children than those who had less adequate contraception easily available. Thus, we might want to pool information on childbearing only for women of the same age, marital duration, and time period, and compare them with women of different ages, marital durations, and time periods to test our hypotheses. Figure 6 uses a sideways Lexis diagram to show which segments of observation can be properly pooled: it is taken from N. B. Ryder's (1975) discussion of the measurement of fertility in cross-sectional (period) data obtained from surveys. The negative diagonals follow the age groups in a period; the positive diagonals separate the cohorts. Thus, the positive diagonals represent the boundaries of birth cohort exposure; the negative diagonals represent boundaries of exposure periods. The diamonds so described enclose the proper exposure time for the calculation, so that they include only the experience of women of the same marriage cohort, the same age, at the same time period.

A second major concern of demographers has been to explicate the regularities in the age composition of a population. The early discoveries by Graunt and Halley of similarities in the age pattern of deaths were the basis of subsequent attempts to use these regularities to estimate the proportion of the population in a given age group. In a stationary population, one with equal numbers of annual births and deaths, the age composition is determined only by the schedule of age-specific mortality rates. When entries and exits from a population are not equal, however, the age composition will be a function of the fertility schedule as well as the mortality schedule. This basic notion was set out by L. Euler in a paper in 1760 in which he introduced the concept of a stable age structure, one in which the proportions in all age categories would remain fixed if mortality were constant and births increased exponentially over time (Euler 1970[1760]).

It is Alfred J. Lotka, however, who is given credit for discovering "the link between stable theory and real populations . . . in works that form a single singular achievement in demography" (Smith and Keyfitz 1977:75). In an important article, Lotka showed that if female mortality and fertility schedules are fixed—in other words, if the age-specific risks of dying and of giving birth do not change over time—there is also a single age distribution to which the female population will eventually converge, given any initial arbitrary age distribution (Lotka 1922). By applying the fixed mortality and fertility schedules to the stable age distribution, it is easy to calculate the annual number of births and deaths, and thus the expected rate at which the stable population will increase or shrink.

Even in cases where the age composition has not yet become that of the stable population implied by a particular combination of fixed mortality and fertility schedules, it is possible to calculate the intrinsic rate of natural increase, that rate implied by its current mortality and fertility. The intrinsic rate of natural increase has important uses in population projections that aim at estimating the proportion of the population in given age groups in the future, under the assumption that current mortality and fertility remain the same. For example, although the current rate of natural increase in the United States is positive, the intrinsic rate of natural increase is negative. Thus, were present patterns of birth and death to continue after the current large cohorts born during the Baby Boom pass through the reproductive years, the U.S. population would begin to decline in size.

In another seminal article, "On the True Rate of Natural Increase" (1925), Dublin and Lotka use individual graphics (first separately and then combined) to illustrate the concept of the mean length of a generation, which is crucial to the calculation of the intrinsic rate of natural increase in a stable population and which measures the average number of years between the mean age at which a cohort of mothers gives birth and the mean age at which their daughters in turn give birth. Dublin and Lotka begin, in Figure 7, a histogram of the age schedule of fertility, to illustrate the fact that a cohort of women does not have all its daughters at the same ages. The areas of the complete columns represent potential births of daughters within the age limits indicated, if there were no deaths in the cohort. Areas of truncated columns represent actual births of daughters within the same age limits, as reduced by deaths of potential mothers. Dublin and Lotka then note that, although the births in each generation are spread over a number of years, there is only one (equal) spacing between generations that corresponds to the given secular rate of increase; this is a particular value of T , the mean length of generation. Its definition is illustrated in Figure 8, which shows the successive descendants (second, third, etc., generations) of an initial cohort of women. Each generation is spaced—for the purpose of defining T —an equal distance apart. Figure 9 combines features of both of the preceding graphs to give the mean length of one generation a mechanical interpretation; the histogram of the maternity function of the mothers on the right is matched on the left by the sum of their daughters, balanced approximately T years from the fulcrum.

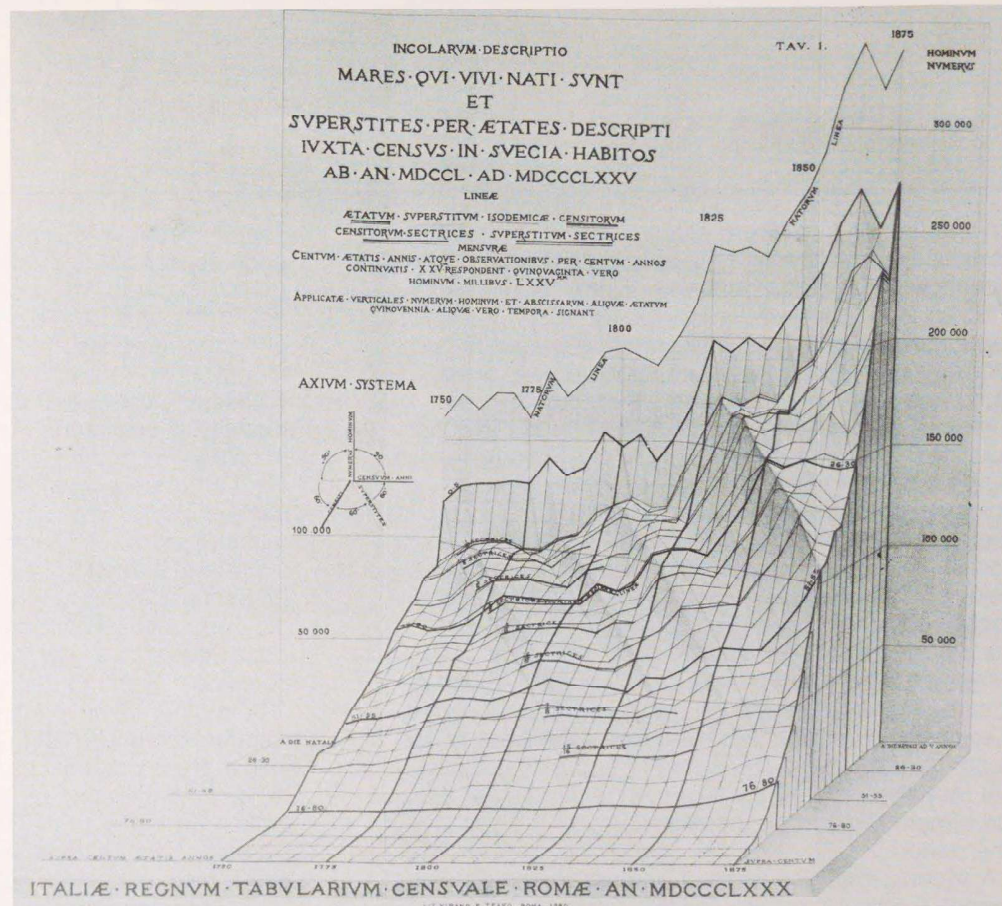
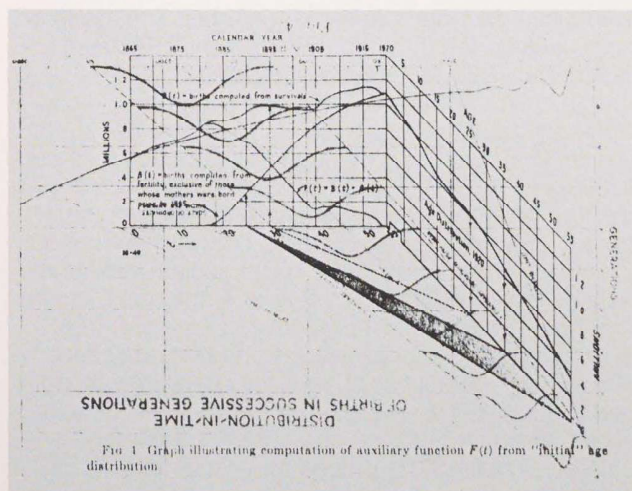


Figure 11 Lotka's 1929 Graph Illustrating Computation of Fertility Function $F(t)$ from Initial Age Distribution. From Alfred J. Lotka, "The Spread of Generations," *Human Biology* (1929), vol. 3.

Figure 12 Perozzo's 1881 Stereogram Depicting the Size and Age Distribution of Successive Cohorts of Swedish Males. From Luigi Perozzo, "Stereogrammi Demografici," *Annali di Statistica*, Ser. 2 (1881), vol. 22, Table 1.



After Lotka established the stability of the normal age distribution, showing that it is the form to which a population will return after displacement or to which it will converge from an arbitrary initial age distribution, a subsequent major problem was revealed: the determination of the effect of changes in mortality or fertility on age composition, a question of theoretical interest but also of practical concern, because many populations had recently experienced changes in death and birth rates. Intuitively, it might seem as if declines in mortality would result in an older population, since more survive to older ages. Normally, however, declines in mortality—such as those experienced in the developing world after World War II—benefit infants and children more than adults and increase their proportion relative to the entire population. Thus, changes in the level of mortality may have the counterintuitive effect of producing a younger rather than an older population. Declines in fertility, on the other hand, always produce an older population, since they reduce the proportion of the population at the younger ages. Figure 10 is a depiction of the shape of the age distribution in three populations, each projected into the future with different fertility and the

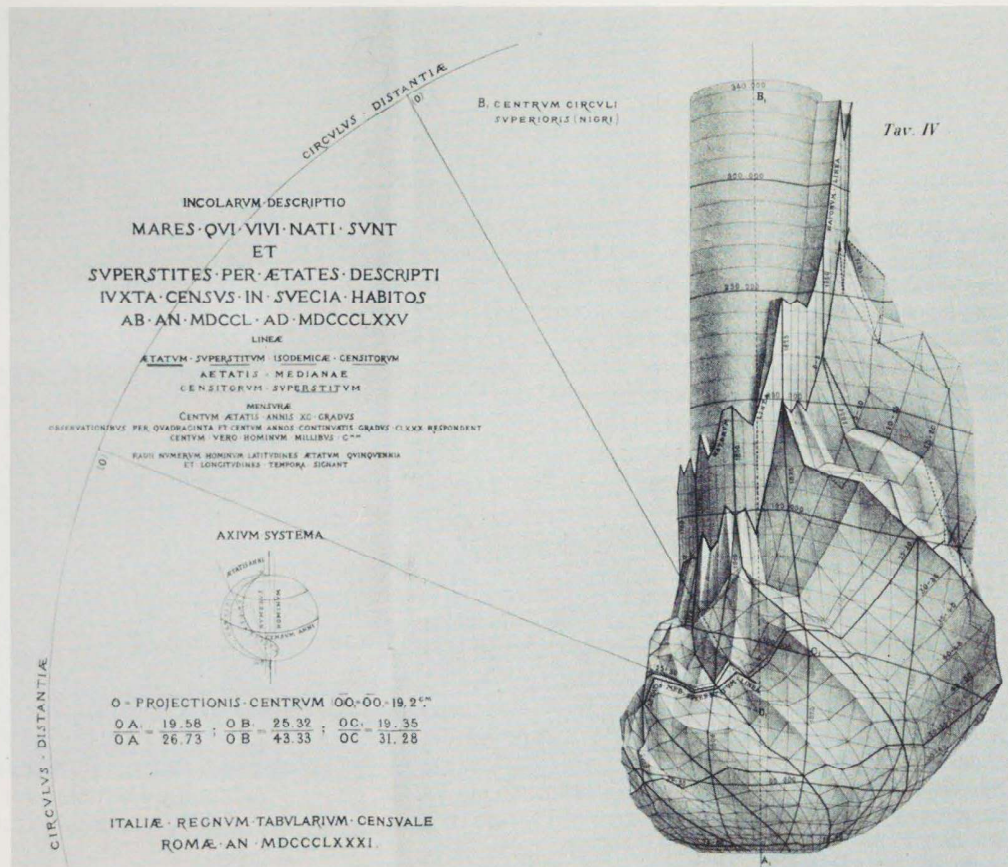


Figure 13 Perozzo's Illustration of the Size and Age Distribution of Successive Cohorts of Swedish Males. From Luigi Perozzo, "Stereogrammi Demografici," *Annali di Statistica*, Ser. 2 (1881), vol. 22, Table 4.

same (no) mortality. The visual comparison helps to make the important point that changes in fertility can have a greater effect on the age distribution than would even the achievement of immortality.

Although it is easy to display graphically the effect of changes in mortality, fertility, or both on the age composition of a population at a single point in time, efforts to picture the experience of successive cohorts have not been notably successful; they have led to the inclusion of more information than can readily be grasped by the eye. For example, Lotka's stereogram, Figure 11, is not easily interpretable. The age distribution of females in 1920 is plotted on a plane reaching forward at right angles to the plane of the paper. The females alive in 1920 are "survived" back to their births by an equation developed in the article, each term of which is plotted separately on the front plane of the drawing. Less immediately apprehensible is Perozzo's (1881) famous stereogram, another attempt to depict the size and age distribution of successive cohorts, in this case Swedish males from 1750 to 1830 (Figure 12). The size is given by the scale on the right, the date is on the horizontal axis, and age composition is on the third dimension. The

somewhat surrealistic presentation in Figure 13 is even more bizarre, although the accompanying cross-sections (Figure 14) provide some clarification by showing population size and age composition at only a single date.

Once regularities have been established, deviations from these expected patterns can be instructive, for calling attention to substantive variation or to suspect data. For example, using a collection of 326 male and 326 female life tables, Coale and Demeny (1966) constructed preliminary model tables to represent the underlying regularities in the force of mortality, and, using graphics such as Figure 15, they visually compared the deviations of individual tables from the age pattern of mortality in the preliminary model tables. In doing so, they found that (1) the pattern of deviation was often similar among life tables for the same population at different times and (2) the pattern of deviation was often similar for populations that were geographically near to one another. The visual comprehension of similarities in these deviations led to the construction of four families of model life tables, distinguished primarily by the relation of infant, childhood, and adult mortality common to a broad

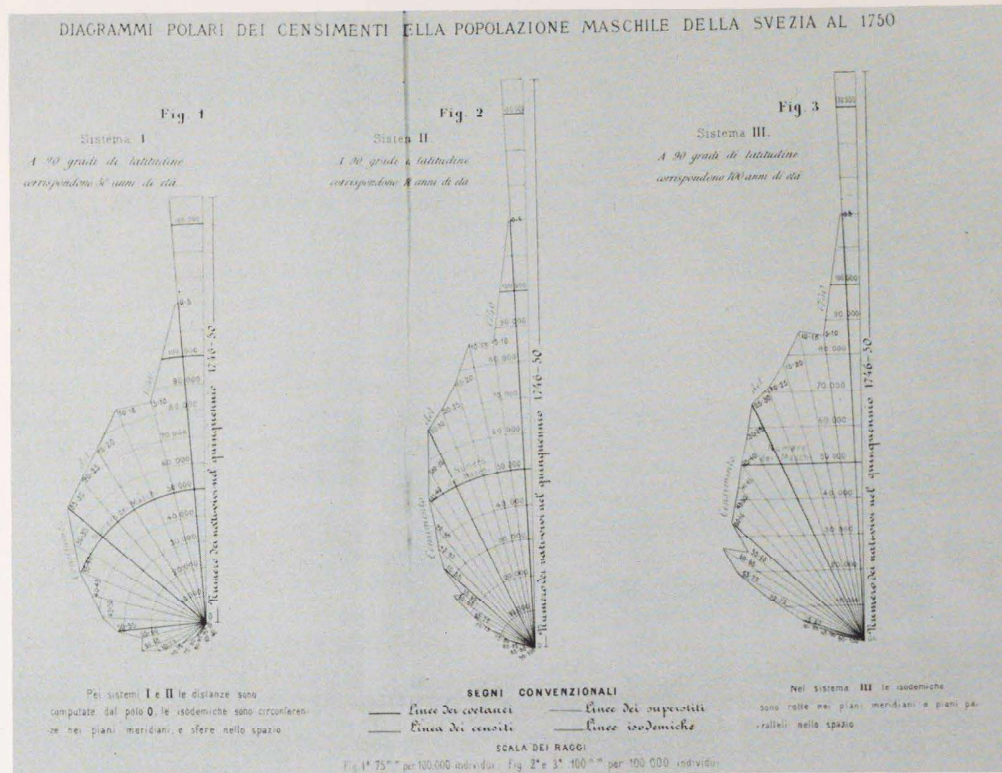
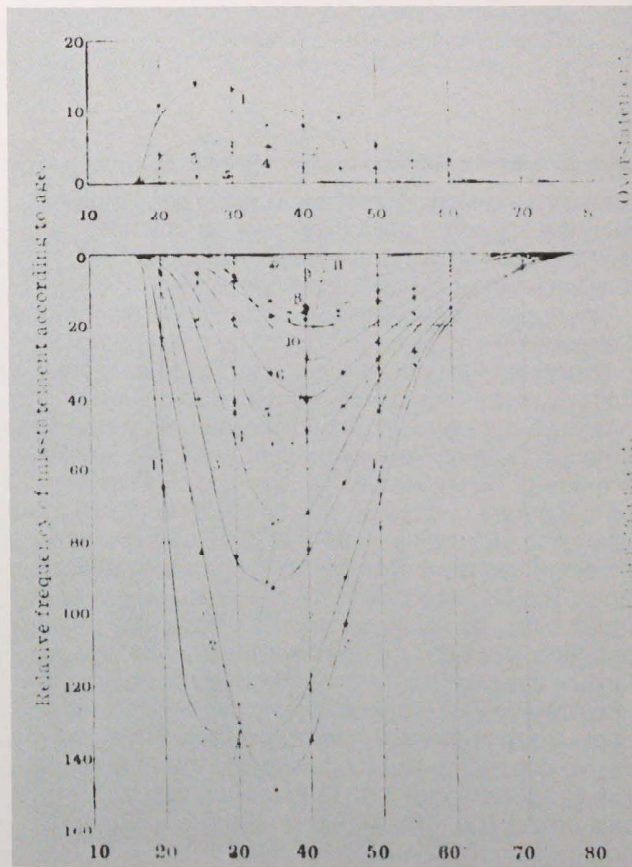
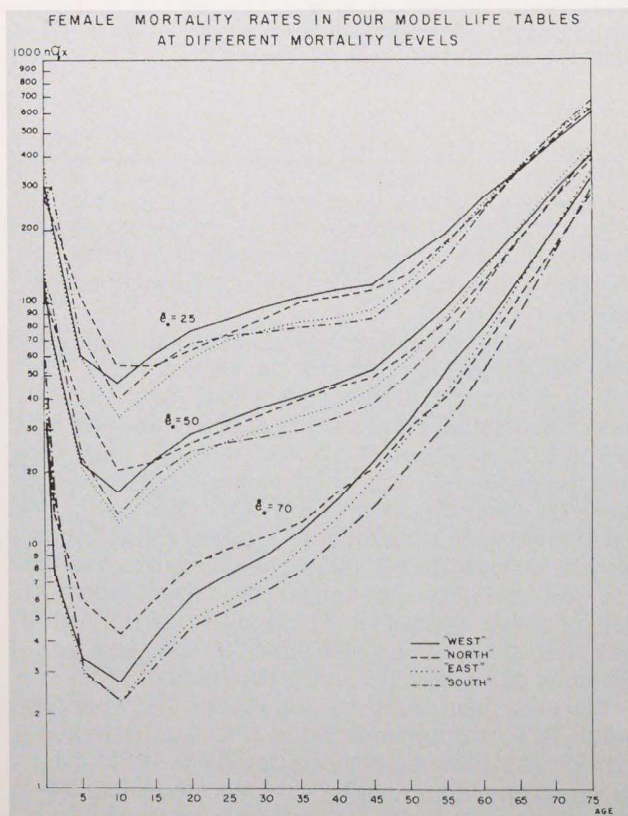


Figure 14 Perozzo's Cross-Sections of the Diagram in Figure 13. From Luigi Perozzo, "Stereogrammi Demografici," *Annali di Statistica*, Ser. 2 (1881), vol. 22, Table 3.

Figure 15 Graphic Comparison of Values of nq_x in the Four Families of Model Life Tables at Different Levels of Expectation of Life, for Females. From Ansley J. Coale and Paul Demeny, *Regional Model Life Tables and Stable Populations* (1966), p. 28.



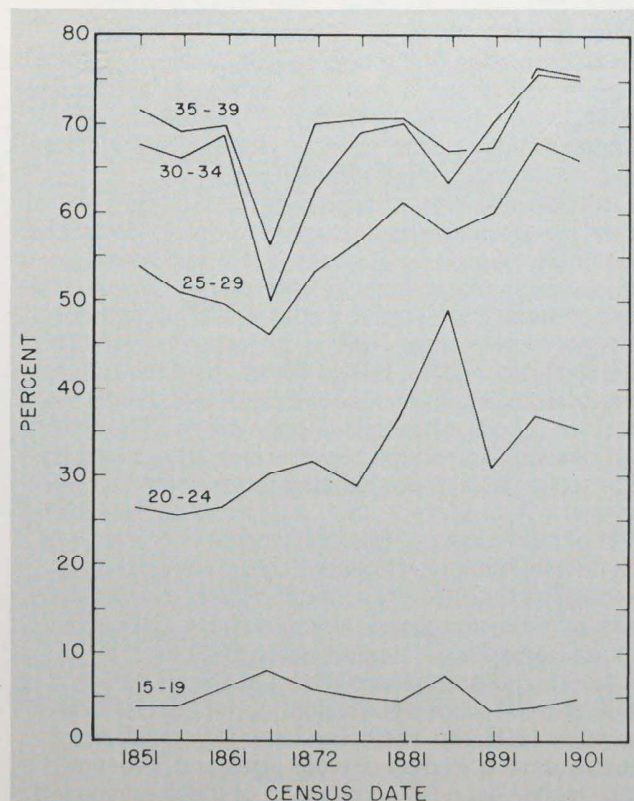
geographical area. Each family consists of mortality schedules with a similar shape but at different levels of overall expectation of life at birth. Thus, if the basic age pattern can be determined, and there is some information to specify the approximate level, the model tables can be used to estimate mortality or survivorship rates for other ages for which information is lacking.

As was shown in the earlier discussion of mortality, the predictability of the age composition of a population, even when fertility and/or mortality have been changing, has made graphical techniques useful in situations where it is necessary to assess the accuracy of demographic observations, and perhaps to reject or adjust these observations. For example, if the stated ages of a population do not conform to its expected age distribution, given its past demographic history, age misstatement may be suspected. Figure 16 is a graph of the age misstatement found by G. H. Knibbs in the 1911 Australian census (Knibbs 1917:115). The vertical axis is the relative frequency of misstatement of age, the horizontal axis is age itself.

The top half of the figure shows overstatement of age, and the bottom half shows understatement; the numbers on the curves denote the amount of misstatement in years. Similarly, to assess the reliability of the French censuses of the nineteenth century, Étienne van de Walle relied extensively on graphical displays of the data for the *départements* of France over more than one hundred years to suggest where the census might be in error. Because the age distribution of a population reflects the idiosyncracies as well as the regularities of its demographic history, unusually large or unusually small cohorts should be visible in successive censuses. Looking at Figure 17, van de Walle concluded that the bulge observed at ages twenty to twenty-four in 1886 is not an accurate description of the size of that cohort in that year (van de Walle 1974). Inaccuracies such as these led him to doubt the quality of many nineteenth-century French censuses and to reconstruct the age composition of the female population of France using stable population theory and information on fertility, mortality, and marriage to correct the census.

Figure 16 Knibbs' 1917 Graph of the Relative Frequency of Age Misstatement According to Age. From G. H. Knibbs, "The Mathematical Theory of Population, of Its Character and Fluctuations, and of the Factors which Influence Them . . .," *Census of Australia 1911* (1917), vol. 1, Appendix A, p. 115.

Figure 17 Proportion Women Married, by Selected Age Groups, Finistère: 1851–1901. From Étienne van de Walle, *The Female Population of France in the Nineteenth Century* (1974), p. 39.



Regularities and Differences: Marital Fertility and Marriage

Regularities in mortality were evident as early as the end of the seventeenth century, but regularities in fertility and marriage took longer to perceive, perhaps because there was less practical reason for interest or because evident differences in level obscured underlying commonalities in patterns.

Recorded differences in the level of fertility among populations are indeed quite striking, ranging from an average of about ten children per couple surviving through the reproductive years for the prolific Hutterites (an Anabaptist group living in the midwestern United States) to under one for the urban population of Vienna in the 1930s. Even in populations where there is no evidence of deliberate attempts to limit childbearing and women continue to give birth until the onset of physiological sterility, the range is from Hutterite levels to about half that much (Coale and Watkins, forthcoming). Despite these differences in the level of fertility, however, the shape of the age-specific fertility curve for these populations is similar: starting with menarche, the ability to give birth to a live child rises for a few years, generally reaching a peak between ages twenty and thirty, and then declines, first slowly and then, after about age forty, rather abruptly. When births are primarily legitimate, the rising portion of the age-specific fertility schedule is a function of age at marriage as well as adolescent sterility, while the falling portion of the curve is determined primarily by the onset of physiological sterility, which is closely related to age.

The establishment of regularities in the age pattern of fertility again makes it possible to detect deviations from the expected pattern, some of which have had substantively important interpretations. A central concern of those interested in population change in the late-nineteenth century and in currently developing countries has been to detect the onset of the beginning of a decline in marital fertility. In eighteenth- and nineteenth-century European populations, it appears that deliberate limitation of births is first practiced by older couples, who presumably reach some desired number of children and then attempt to prevent the birth of more. Figure 18 is of historical importance in the development of historical demography, for it accompanies Louis Henry's introduction of precisely this point into the literature. The figure shows the succession of age-specific marital fertility rates for the women of the bourgeoisie of Geneva born before 1600 and the same information for the cohorts born between 1600 and 1649. The latter falls much more steeply among women of older ages and, presumably, higher parity. A comparison of these curves led Henry (1956) to conclude that during this period older couples adopted some sort of behavior to limit the number of further births. Because populations vary in

the level of fertility for reasons that have nothing to do with attempts to terminate childbearing (for example, because of variations in the extent and duration of breast-feeding), examination of the shape of the age-specific fertility curve is commonly used to identify fertility limitation when direct observation of contraceptive use is not available. Figure 19 shows index values of marital fertility in European historical populations and twentieth-century Asian populations. In each panel, the rates for Sweden 1961–1965—a population in which marital fertility was well under control—are graphed to serve as a comparison. The curve for Sweden falls much more steeply with age than do the curves for the other countries.

The transition to family limitation occurred among the Genevan bourgeoisie and some of the aristocracies of Europe in the seventeenth century, in most of the countries of Western Europe between 1880 and 1930, and in some countries of the currently developing world after 1960, yet the pattern traced out by the schedule of age-specific marital fertility rates in populations in which fertility is deliberately limited is quite similar. In all, the rapid decline of fertility with age (and, again, presumably parity) is quite evident compared to a population in which no attempt is made to terminate childbearing.

Although the importance of marriage age for completed fertility was noted as early as the ancient Romans (United Nations 1953), similarities in age-specific marriage patterns took longer to detect. In some populations, females marry at a very young age, on average in their teens, and virtually all of them (more than 98 percent) marry at least once, most of them by the time they are twenty-five. In others, females marry much later, on average in their mid-twenties or even older, and substantial proportions of them remain spinsters throughout their lives. As with mortality and marital fertility, however, underlying this apparent diversity in marriage are profound regularities. In an article that has proved to be immensely provocative of further work, J. Hajnal (1965) called attention to a surprising correspondence between geography and marriage in historical populations. He noted that Asian and African societies were (and to a large extent still are) characterized by early and virtually universal marriage for women, while historical Western Europe has been characterized by late age of marriage and a high proportion of women never-marrying. Hajnal and others subsequently have used the sparse data for periods before the nineteenth century to show that this pattern is long-standing, apparently already in place by the fourteenth century (Smith 1979), though neither the origins of the Western European marriage pattern nor its causes are well understood (Watkins 1981, 1984). Its consequences, however, are significant, since in Western Europe, where relatively few births occurred outside

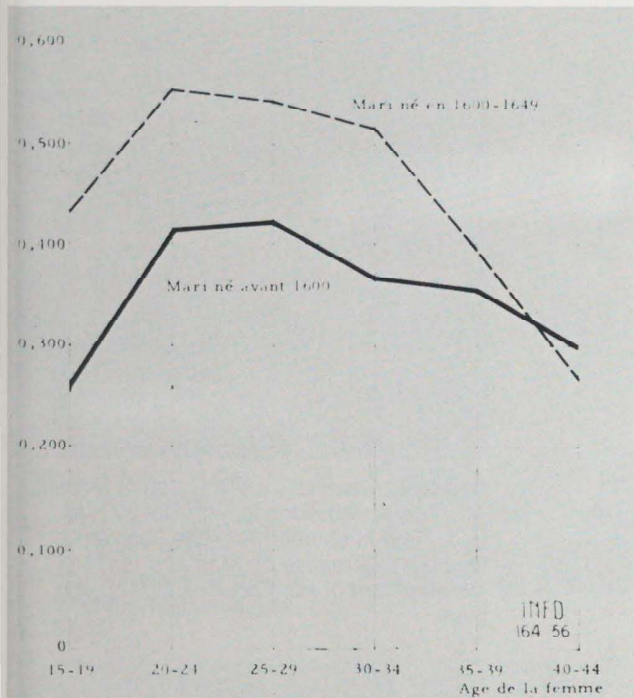
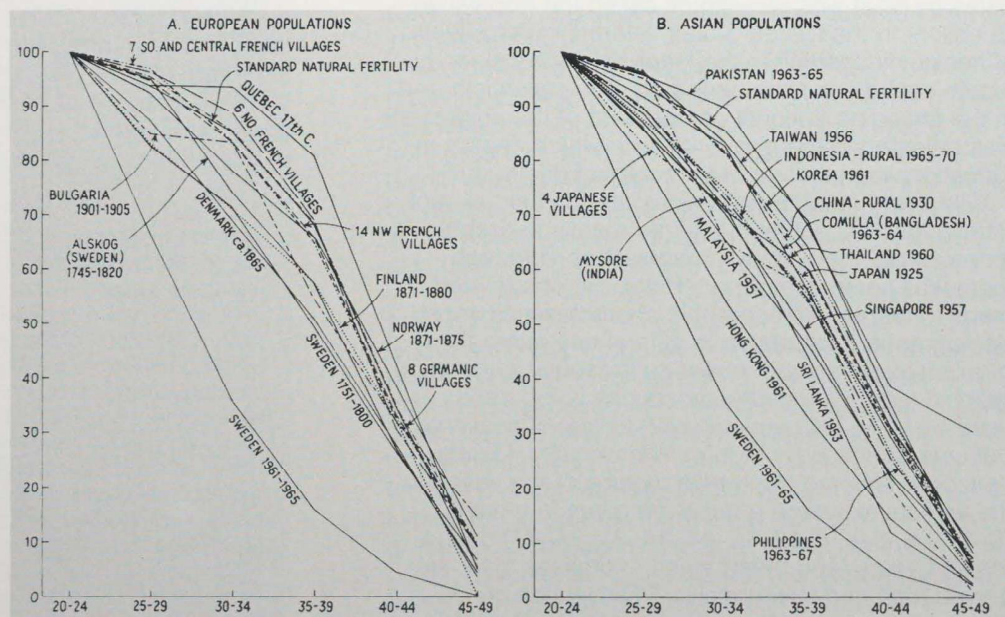


Figure 18 Henry's Fertility Rates by Age, for Women who were Subsequently Fertile, Grouped by Age of Marriage, for Women of the Genevan Bourgeoisie. From Louis Henry, *Anciennes Familles Genevoises* (1956), p. 122.

of marriage, the fact that many potentially fertile years were spent outside of marriage substantially lowered the total fertility rate to about a third of what it would have been had all women married early and born children rapidly throughout their reproductive years.

Nor are geographic uniformities the only ones. Despite the striking differences in the proportions married by age in various geographic areas, it is possible to describe the cumulative frequency of first marriage with a single curve, much as shape of the age-specific mortality schedule is similar in populations with very different levels of mortality (Coale 1973). That populations as widely different in other respects as Sweden in the late-nineteenth and Taiwan in the early twentieth century should have similar age patterns of dying and of living by age is perhaps not so surprising, given the biological bases of both behaviors. That the cumulative frequency of first marriages should be similar is much more unexpected; it suggests a patterned sequence of entry onto and exit from the marriage market that differs from society to society primarily in the age at which first marriages begin to take place, the speed with which the female population exists from the marriage market, and the proportion ultimately exiting.

Figure 19 Graphical Comparison of Index Values of Age-Specific Marital Fertility Rates (20-24 = 100). From John Knodel, "Family Limitation and the Fertility Transition," *Population Studies* (1977), vol. 31, p. 225.



Differentials

In the period before census bureaus were established and charged with the regular collection of demographic information, when such information was collected either by church or state, it was usually limited to marriages, births (or baptisms), deaths (or burials), and sometimes age (for example, the number of males in the age groups liable for military service); demographic data before the mid-nineteenth century relies strongly on church records of burials and baptisms, with little other information describing social or economic characteristics. During the nineteenth century the increasing availability of vast amounts of information, gathered in successively more regular censuses and more accurate systems of vital registration, was exploited to describe and compare the demographic behavior of ever-smaller and more exotic subgroups. Census questionnaires and the categories by which published census information was tabulated suggest particular interest in ethnic and class differentials in mortality and fertility. Differences in risks of dying among various groups—urban v. rural, rich v. poor—came to be of public concern. Once the decline in marital fertility became apparent, it also became evident that these declines were occurring earlier and more rapidly among some segments of the population than among others, leading to the formation of eugenics societies and to vigorous debates on what was sometimes called the Population Question. In the United States, interest in the effects of heavy immigration in the latter half of the nineteenth century and the early part of the twentieth and in the changing national composition of that migration was reflected in the frequency with which census information was tabulated by country of origin.

With some notable exceptions, the presentation of differentials did not lend itself as readily to graphic forms as did the display of regularities. Although there was some use of bar charts, pie charts, and maps to display demographic characteristics, most demographers interested in differences rather than regularities presented their data in tabular rather than graphic form. One of the exceptions is Figure 20, used by L. A. J. Quetelet (1842) to depict seasonal patterns of unusually high or low mortality. On the horizontal axis are the twelve months of the year. On the vertical axis, age groups are given in months (0–1, 1–3, 3–6, 6–12, 12–14, 14–24, 24–37) during the first years, and then by years; crosses mark the annual high and low mortality. Examination of seasonal patterns of deaths in turn has been used to suggest the interaction of season and causes of death.

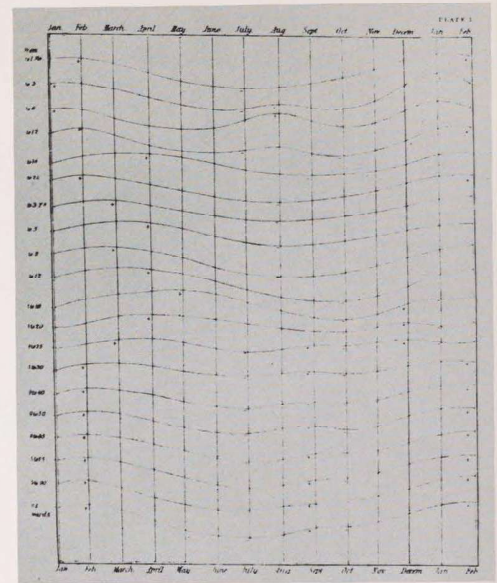


Figure 20 Quetelet's 1842 Graph of Monthly and Yearly Variations in Mortality by Age Group. From Lambert Adolphe Jacques Quetelet, *A Treatise on Man and the Development of His Faculties* (1842), end page.

The most successful forms of graphics used widely in these endeavors were maps that displayed the spatial distribution of population characteristics, such as those seen in E. Levasseur's *La Population Française* (1889) or F. A. Walker's *Statistical Atlas of the United States* (1874). Walker, who is credited with developing the age pyramid used frequently by demographers (see Beniger and Robyn 1978), also provides an example of another successful graphic, shown in Figure 21.

The explanation given in the text is as follows:

The squares are proportional to the population of the states respectively represented (350,000 inhabitants to the square inch). Each square is divided by vertical lines into three rectangles, the left representing the foreign, the middle the native colored, and the right the native white population. Each of the last two rectangles is divided by a horizontal line to exhibit the proportion of each class of the population represented born respectively within and without the State itself. The lower portions of these rectangles thus divided represent the number born in other States and Territories of the Union. Each square has a rectangle of equal height [sic] upon its right, which exhibits, in proportion, the number of persons born in the State, who have become residents of other States. This rectangle is divided by a horizontal line, the upper portion representing colored, the lower white. [Walker 1874, Plate XX]

Figure 21 Walker's Map of the United States. From F. A. Walker, *Statistical Atlas of the United States Based on the Results of the Ninth Census, 1870* (1874), U.S. Census Office, Washington, D.C.

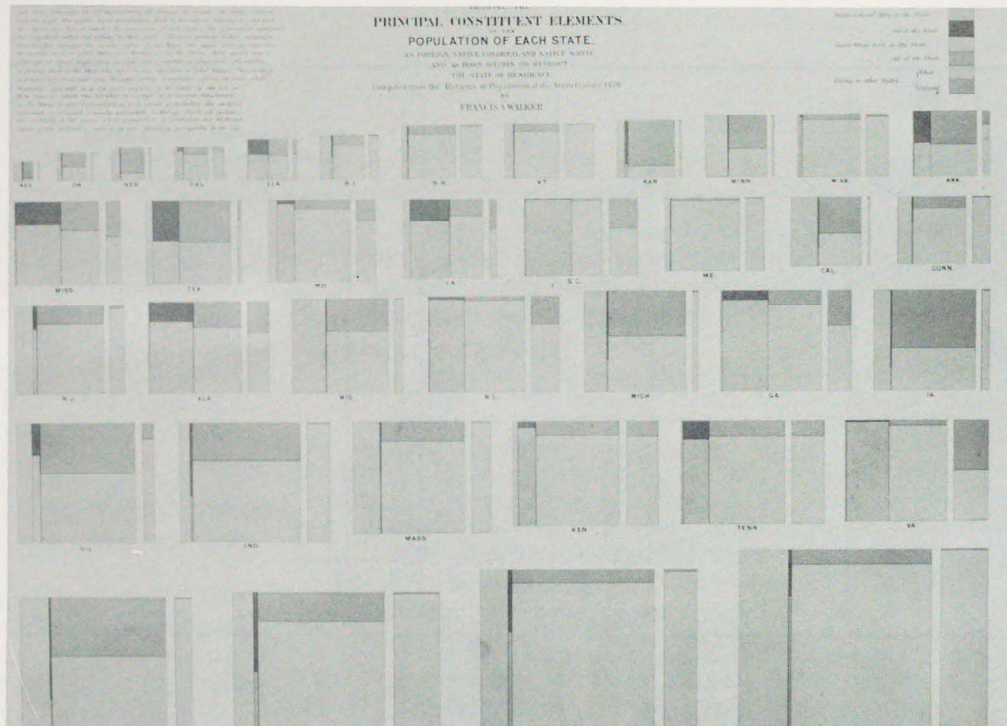
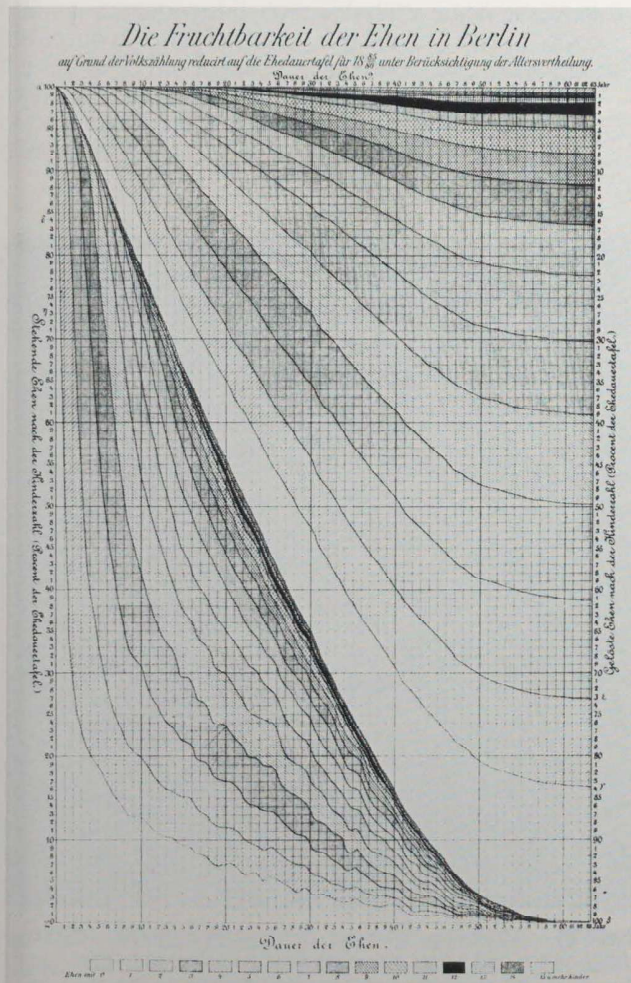


Figure 22 Böckh's 1890 Graph of Fertility by Duration of Marriage. From R. Böckh, "Die Statistische Messung er ehelichen Fruchbarkeit—Bortrag gehalten in der staatwissenschaftlichen Gesellschaft zu Berlin am 2 Juni 1890," *Bulletin de l'Institut International de Statistique* (1890), vol. 1, p. 187.



The color code is given in the upper right-hand corner: from top to bottom, these are the persons born out of the United States, native colored born in the state, native colored born out of the state, native white born in the state, native white born out of the state, while living in other states, and colored living in other states.

Rare forays during this period into other types of graphics to describe marriage or childbearing have had less pleasing results. Figure 22 is an example; from an article by a leading statistician of the nineteenth century, Richard Böckh (1810), it shows fertility by duration of marriage for several cohorts. The horizontal axis shows the duration of marriage; the vertical, the percent of each duration cohort having 0, 1, 2 . . . 15 or more children. To use E. R. Tufte's terminology (Tufte 1983), in this graph the ratio of data to ink is low; a simpler graph would have been more appropriate. A review of nineteenth-century graphical presentations based on the richer data available from censuses and vital registration systems suggests that, where differentials are of concern, sometimes the amount of information to be conveyed cannot well be expressed in graphic form, as in the Perozzo stereogram discussed earlier; in other cases, as in the Böckh graphic, the particular form chosen is unsatisfactory. In general, however, those demographers whose interest has been primarily in differentials do not appear to rely on graphs in the analysis of their data, and the presentation is usually in tabular form. It is the lumpers rather than the splitters who are most likely to turn to graphics.

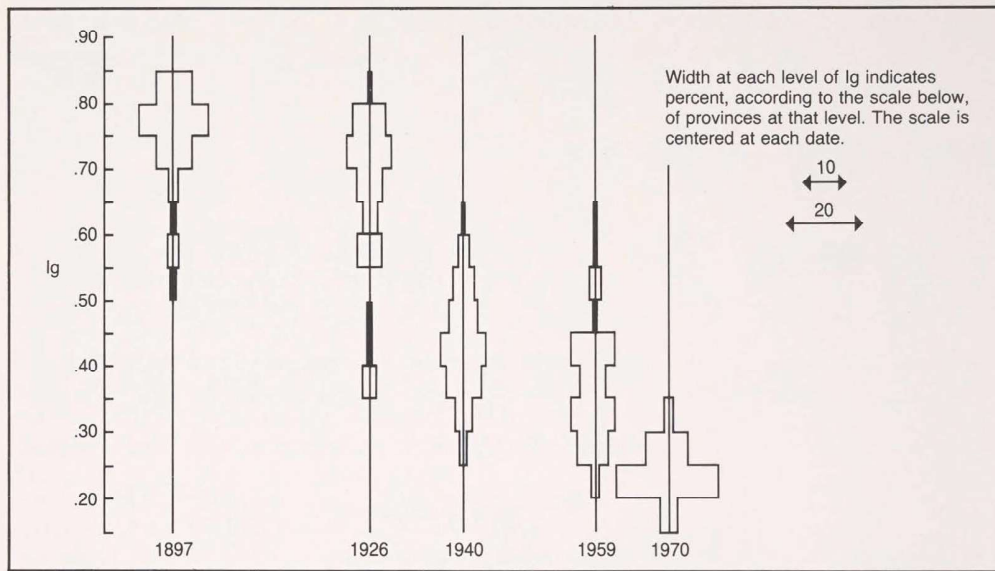


Figure 23 Marital Fertility (Ig) of Total Population, Distribution of Provinces of European Russia by Level: 1897-1970. From Ansley J. Coale, Barbara Anderson, and Erna Härm, *Human Fertility in Russia since the 19th Century* (1979), Figure 2.2.

Computer Graphics

The development of the computer greatly aided in the numerical analysis of large quantities of data, but only recently have the computer's graphic capabilities been exploited by demographers to analyze or display data. Perhaps the most outstanding use of computer graphics to this end so far has been in the Princeton European Fertility Project, a large-scale analysis of the demographic transition among the approximately six hundred provinces of Western Europe between 1870 and 1960. During that century both birth rates and death rates fell by at least 50 percent; the project was undertaken to describe and understand the circumstances under which this transition began (Coale and Watkins, forthcoming).

Marital fertility, illegitimate fertility, and nuptiality together determine the overall fertility of the population. Correspondingly, the decline in overall fertility in the provinces of Western Europe may have been due to changes in any of these components. To measure the relative contribution of each, new measures, which it is necessary to explain briefly here, were developed. Ig is an index of marital fertility that ranges between 0 and 1, constructed by comparing the observed marital fertility rates by age with those of the highest reliability-recorded fertility rates on record, those of the Hutterites. Thus, an Ig of .7, typical of the provinces of Western Europe before the decline of fertility, is interpreted as 70 percent of the level of the Hutterites. Ih is an equivalent measure of illegitimate fertility. Im is a fertility-weighted index of marriage, and can be interpreted as the proportion of women of reproductive ages (15-49) who are married. If, the index of overall fertility, is defined as the product of Ig and Im, plus the contribution to overall fertility of births by unmarried women (which was generally so low in Western Europe that it can be ignored in most discussions of fertility change.)

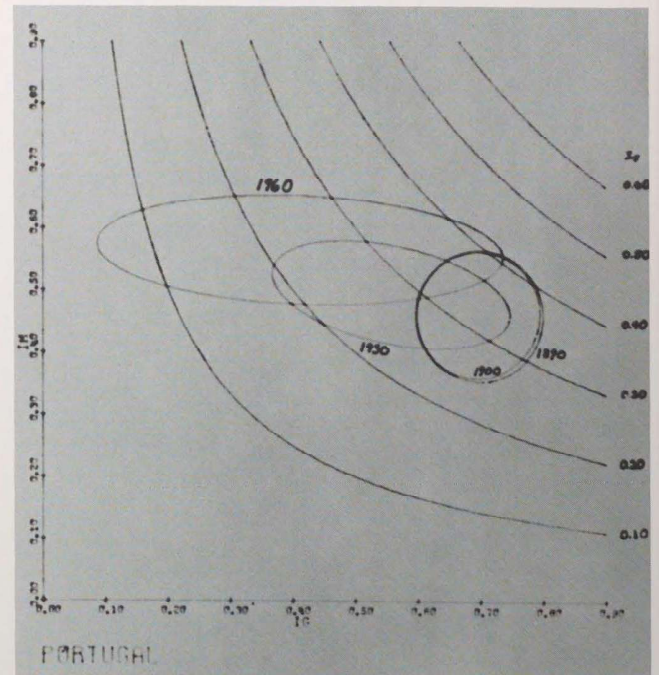


Figure 24 "Marching Ellipses," Districts of Portugal, 1870, 1900, 1930, and 1960. From Ansley J. Coale and Roy Treadway, "A Summary of Changing Fertility in the Provinces of Europe" (1979), in Coale and Watkins (forthcoming), Figure 2.23.

Two new graphics were developed subsequently to explore and summarize province-by-province changes in these measures in Europe over the past century; these graphics are known as "sliding popsicles" and "marching ellipses." In Figure 23, the sliding popsicles illustrate the changes in marital fertility in Russia at five dates: 1897, 1926, 1940, 1959, and 1970. The percent of the provinces in which I_g falls within a specified interval is indicated by the width of a bar. The total area of the bars is the same at each date, and each set of bars is centered on the date for which the I_g 's are calculated (Coale, Anderson, and Härm 1979)

Marching ellipses demonstrate effectively the sequence of change in marital fertility and proportions married, and the contribution of each to the overall fertility of the population. Figure 24 shows the ellipses drawn for Portugal at four dates. The vertical axis shows I_m , the proportion married; the horizontal axis shows I_g , marital fertility. The curved lines, called isoquants, are labeled I_f , overall fertility; each isoquant is the locus of combinations of I_g and I_m that produce the same I_f . The location, orientation, and principal dimensions of each ellipse summarizes the dispersion of points in the plane described by I_m and I_g at a given date.

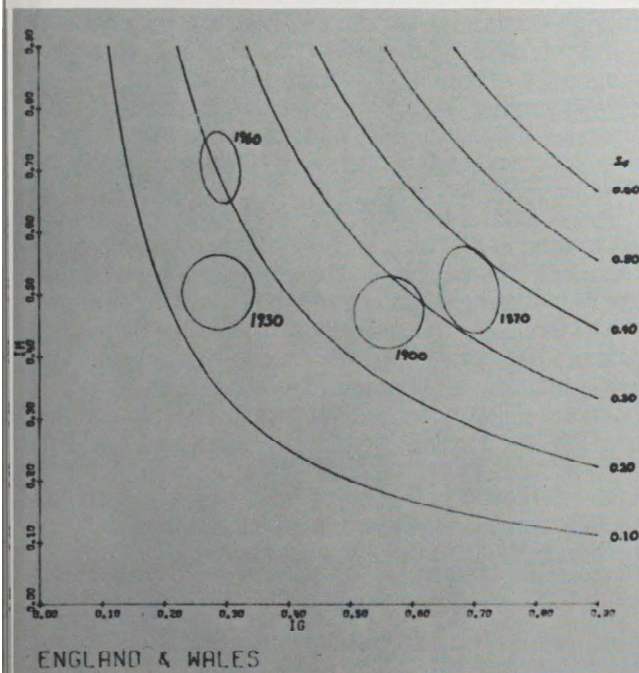


Figure 25 "Marching Ellipses," Counties of England and Wales, 1871, 1901, 1931, and 1961. From Ansley J. Coale and Roy Treadway, "A Summary of Changing Fertility in the Provinces of Europe" (1979), in Coale and Watkins (forthcoming), Figure 2.12.

A mathematical scheme for the calculation of the ellipses was devised by Michael Stoto. The method for determining the location, shape, and size of the ellipse is described below:

(1) The points to be represented are divided into the top third, middle third and bottom third with respect to position along the axis (I_m or I_g) or in which the inter-quartile distance is the greatest.

(2) The median value of I_m and the median value of I_g are found within the bottom set and the top set of the points. The slope of the line connecting the (I_m , I_g) doublets thus located is taken as the slope of the major axis of the ellipse; the minor axis is perpendicular to the major axis.

(3) A preliminary length of each axis is calculated as the distance between the quartiles, when the points are arranged in order along each axis. The center of the ellipse is located at the median position along each axis.

(4) The first three steps determine the location, direction of major axis and shape of the ellipse. Its final size is fixed by multiplying both axes by a scale factor chosen so that the adjusted ellipse encompasses 75 percent of the points. [Coale, Anderson, and Härm 1979]

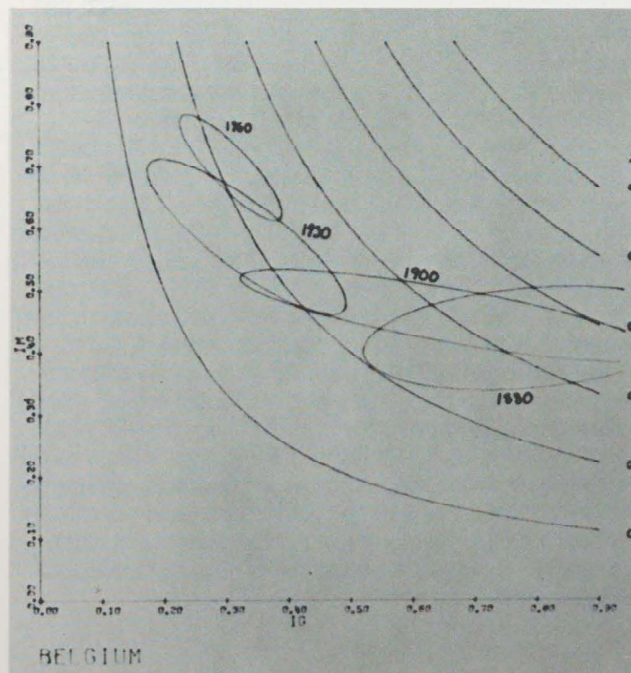


Figure 26 "Marching Ellipses," Arrondissements of Belgium, 1880, 1900, 1930, and 1960. From Ansley J. Coale and Roy Treadway, "A Summary of Changing Fertility in the Provinces of Europe" (1979), in Coale and Watkins (forthcoming), Figure 2.210.

It is evident that the ellipse is relatively unaffected by extreme values.

The ellipses facilitate the examination of both the level and the distribution of two variables as they change over time. In Figure 24, the succession of ellipses from the high marital fertility of the past to the low fertility of the present is represented by the movement from left to right on the horizontal coordinate. The ellipse that is farthest to the right represents the districts of Portugal in 1900, when the circular shape of the ellipse shows that there was about the same amount of variation in nuptiality as in fertility. In 1960, the ellipse is extremely elongated on the horizontal axis, showing that while there had been little change in the dispersion in nuptiality (the vertical axis), some districts had substantially reduced levels of marital fertility while others had levels only slightly lower than in 1900. Matters were quite different in England and Wales, as shown in Figure 25. Change in all the counties was rapid and nearly simultaneous; as a result, the ellipses are small and do not overlap. Nuptiality does not change significantly until after 1931, when falls in the age of marriage and in the proportion of spinsters caused *lm* to rise.

One of the more consistent findings in this study of fertility decline in Europe was the importance of regional location in determining the timing of the onset of fertility limitation. Marital fertility declined first in France, and in areas contiguous to France, such as the canton of Vaud in Switzerland, Catalonia in Spain, and the French-speaking parts of Belgium. Figure 26 shows separate ellipses drawn for the two language groups of Belgium; the red ellipse is the French-speaking Wallonia, the blue ellipse is Flemish-speaking Flanders, and the black ellipse is drawn for all the provinces together. The tilt of the ellipse for each region describes the tilt of each part of Belgium toward its linguistically similar neighbor. Thus, the provinces in Wallonia that are closest to the French border are the most similar to the *départements* of France in low marital fertility and high proportions married; these characteristics become more diluted as the distance from the French border increases. Luxembourg exerts a similar pull on the linguistically similar provinces of Flanders. Although in this case language seems to be the basis of demographic similarity, in other countries the geographic contiguity that is reflected in common demographic patterns seems to be based on other aspects of a common culture (Watkins, in Coale and Watkins, forthcoming).

Conclusions

In demography, as in many other fields, graphs have played an important role both in presentation and in discovery. Most of the uses of graphs that have been discussed here fall into two groups: the pedagogical and the pattern-finders.

Halley, Nicander, Milne, Lotka, Lexis, and Ryder, in the examples considered here, used graphical displays pedagogically, to illustrate a previous mathematical or verbal analysis. Although they may in fact have used displays of the data in the exploration that led to the analytic formulation, the published graphs themselves appear as illustrations that are intended to give the reader a better intuitive notion of points previously made.

The pattern-finders display real data, either to search for underlying commonalities or to summarize visually what cannot be succinctly expressed verbally. Quetelet's graph of seasonal mortality, Perozzo's depiction of the effects of mortality and fertility on successive cohorts of Swedish males, the model mortality, fertility, and nuptiality schedules, and the "marching ellipses" of the demographic transition all present considerable quantities of data in a way that exposes empirical regularities that might otherwise go unnoticed. Once the regularities have been established, visual comparison of new data with old patterns has sometimes led to the rejection of the data as inaccurate, and sometimes to significant substantive findings. The use of graphical techniques to summarize differentials in demographic behavior among a variety of subgroups of the population—for example, by nativity or occupation—has in general been less typical and often less successful, with the notable exception of maps.

The emphasis here has been on the most interesting graphs in the work of a handful of men who made important contributions to demography. The examination of the use of graphics in the history of demography emphasizes the distinction between two very different traditions in the field: those who have looked for regularities, the lumpers, and those who have looked at differences, the splitters. Thus, it would appear that the study of graphics in a specialized field not only adds to the ways in which visual depictions have been used in theoretical and empirical work in another specialized field, but also to understanding of the intellectual history of the field itself.

Acknowledgment

An early version of this paper was presented at the annual meeting of the Council on Social Graphics. I owe much to Jim Beniger for proposing that I write it for his session, and for helpful comments at every stage.

References

- Beniger, J. R., and Robyn, D. L.
1978 Quantitative Graphics in Statistics: A Brief History. *The American Statistician* 32:1–11.
- Böckh, R.
1890 Die Statistische Messung der Ehelichen Fruchbarkeit—Bortrag gehalten in der Staatwissenschaftlichen Gesellschaft zu Berlin am 2 Juni 1890. In *Bulletin de L'Institut International de Statistique* 1:159–187.
- Coale, A. J.
1973 Age Patterns of Marriage. *Population Studies* 25(2):193–214.
forthcoming The Decline of Fertility in Europe Since the 18th Century as a Chapter in Demographic History. In *The Decline of Fertility in Europe*. A. J. Coale and S. C. Watkins, eds. Princeton: Princeton University Press.
- ———, B. Anderson, and E. Härm
1979 Human Fertility in Russia Since the 19th Century. Princeton: Princeton University Press.
- ———, and P. Demeny
1966 Regional Model Life Tables and Stable Populations. Princeton: Princeton University Press.
- ———, and R. Treadway
forthcoming A Summary of Changing Fertility in the Provinces of Europe. In *The Decline of Fertility in Europe*. A. J. Coale and S. C. Watkins, eds. Princeton: Princeton University Press.
- ———, and S. C. Watkins, eds.
forthcoming The Decline of Fertility in Europe. Princeton: Princeton University Press.
- Dublin, L. I., and A. J. Lotka
1925 On the True Rate of Natural Increase. *Journal of the American Statistical Association* 20:305–339.
- Euler, L.
1970 [1760] A General Investigation Into the Mortality and Multiplication of the Human Species. Translated by N. and B. Keyfitz. In *Theoretical Population Biology* 1:307–314.
- Fourier, J. B. J.
1821 Notions Générales sur la Population. In *Recherches Statistiques sur la Ville de Paris et le Département de la Seine*. Paris: C. Ballard.
- Graunt, J.
1939 [1662] Natural and Political Observations Made Upon the Bills of Mortality. Walter F. Willcox, ed. Baltimore: Johns Hopkins University Press.
- Hajnal, J.
1965 European Marriage Patterns in Perspective. In *Population in History*. D. V. Glass and D. E. C. Eversley, eds. London: Edward Arnold.
- Halley, E.
1942 [1693] Degrees of Mortality of Mankind. Edited with an introduction by L. J. Reed. Baltimore: Johns Hopkins University Press.
- Henry, L.
1956 Anciennes Familles Genevoises. INED Cahier No. 26. Paris: Presses Universitaires de France.
- Knibbs, G. H.
1917 The Mathematical Theory of Population, of its Character and Fluctuations, and of the Factors Which Influence Them. Appendix A, Volume I: Census of the Commonwealth of Australia, 1911.
- Knodel, J.
1977 Family Limitation and the Fertility Transition. *Population Studies* 31:219–249.
- Levasseur, E.
1889 La Population Française. 2 vols. Paris: Arthur Rousseau.
- Lexis, W.
1875 Einleitung in die Theorie der Bevölkerungsstatistik. Strassburg: Karl Trubner.
- Lotka, A. J.
1922 The Stability of the Normal Age Distribution. *Proceedings of the National Academy of Sciences*, Volume 8.
1929 The Spread of Generations. *Human Biology*, Volume 3.
- Milne, J.
1815 A Treatise on the Valuation of Annuities and Assurances on Lives and Survivorships; and on the Construction of Tables of Mortality; and on Probabilities and Expectations of Life. Volumes I and II. London: Longman, Hurst, Rees, Orme and Brown.
- Nicander, H.
1801 Svenska (Kongliga) Vetenskaps Academiens nya, Handligar 22:56–72.
- Perozzo, L.
1881 Stereogrammi Demografici. Seconda Memoria dell'Ingegnere. In *Annali de Statistica*. Ser. 2 (22):1–19.
- Quetelet, L. A. J.
1842 A Treatise on Man and the Development of His Faculties. M. A. Quetelet, trans. Edinburgh: W. and R. Chambers. End page.
- Ryder, N. B.
1975 Fertility Measurement Through Cross-Sectional Surveys. *Social Forces* 54:7–35.
- Smith, D., and N. Keyfitz
1977 *Mathematical Demography: Selected Papers*. Berlin, New York: Springer-Verlag.
- Smith, R. M.
1979 Some Reflections on the Evidence for the Origins of the 'European Marriage Pattern' in England. In *The Sociology of the Family*. C. Harris, ed. London: Keele.
- Tuft, E. R.
1983 *The Visual Display of Quantitative Information*. Cheshire, Conn: Graphics Press.
- United Nations
1953 Determinants and Consequences of Populations Trends. New York: The United Nations. Department of Social Affairs, Population Division. *Population Studies* No. 17.
- van de Walle, E.
1974 The Female Population of France in the Nineteenth Century. Princeton: Princeton University Press.
- Walker, F. A.
1874 *Statistical Atlas of the United States Based on the Results of the Ninth Census, 1870*. Washington, D.C.: U. S. Census Office.
- Watkins, S. C.
1981 Regional Patterns of Nuptiality in Europe, 1870–1960. *Population Studies* 35(2):199–215.
1984 Spinsters: An Introduction. *Journal of Family History* 9(4):310–325.
forthcoming Conclusions. In *The Decline of Fertility in Europe*. A. J. Coale and S. C. Watkins, eds. Princeton: Princeton University Press.