

creased cardiovascular mortality, violence, and alcoholism.

See also: *Disease, Burden of; Diseases, Chronic and Degenerative; Diseases, Infectious; Health Transition; Mortality Decline; Mortality Reversals; Oldest Old.*

BIBLIOGRAPHY

- Caldwell, John C. 2001. "Population Health Transition." *Bulletin of the World Health Organization* 79: 159–160.
- Caselli, Graziella, France Meslé, and Jacques Vallin. 2001. "Les entorses au schéma de la transition épidémiologique." Paper presented to the general conference of the International Union for the Scientific Study of Population, Salvador, Brazil, August 2001.
- Frenk, Julio, José Luis Bobadilla, Claudio Stern, Tomas Frejka, and Rafael Lozano. 1991. "Elements for a Theory of the Health Transition." *Health Transition Review* 1: 21–38.
- Meslé, France, and Jacques Vallin. 2000. "Transition sanitaire: Tendances et perspectives." *Médecine/science* 16: 1,161–1,171.
- Meslé, France, and Vladimir M. Shkolnikov. 2000. "Russie: Une crise sanitaire sans précédents." *Espace, populations, sociétés* 2: 253–272.
- Myers, Georges, Vicki Lamb, and Emily Agree. 2002. "Patterns of Disability Change Associated with the Epidemiologic Transition." In *Determining Health Expectancies*. Chichester: John Wiley & Sons, Ltd.
- Olshansky, S. Jay, and A. Brian Ault. 1986. "The Fourth Stage of the Epidemiologic Transition: The Age of Delayed Degenerative Diseases." *Milbank Quarterly* 64: 355–391.
- Omran, Abdel R. 1971. "The Epidemiologic Transition: A Theory of the Epidemiology of Population Change." *Milbank Memorial Fund Quarterly* 29: 509–538.
- Robine, Jean-Marie. 2001. "Redefining the Stages of the Epidemiological Transition by a Study of the Dispersion of Life Spans: The Case of France." *Population: An English Selection* 13(1): 173–194.

JEAN-MARIE ROBINE

ESTIMATION METHODS, DEMOGRAPHIC

Demographic estimation methods have been developed to cope with inadequacies frequently found in standard demographic data. In settings where population statistics are of good quality, key descriptive demographic measures are calculated as occurrence/exposure rates, with occurrences recorded by a vital statistics system and exposure time obtained from population estimates, the latter typically census-based. In many developing countries, data from these sources may simply not be available, or may be affected by systematic errors that bias the resulting measures.

An Overview of Estimation

Improvements in demographic data have reduced the need for demographic estimation. For example, the birth histories widely collected in sample surveys in developing countries provide adequate measures of fertility and child mortality from occurrence/exposure data. However, measures for small areas and of other population parameters, such as adult mortality and migration, still often require estimation. Even when population statistics are generally adequate, estimation methods have proved useful for tracing historical trends in demographic parameters, and are also helpful for estimating some parameters of current population dynamics that are particularly hard to measure, such as migration.

Demographic estimation methods can be broadly categorized into three groups: those that estimate rates from *changes in stocks*, those that are based on *consistency checks*, and those that are based on *indirect estimation*. The ideas underlying the three groups are illustrated below with examples.

Changes in Stocks

Stocks, such as the number of people in a population over age 50 or the number of children ever born to a cohort of women, change as a result of demographic events. Changes in stocks can therefore be used to draw inferences about underlying demographic rates. In situations where demographic events are not directly recorded, or are recorded with unacceptable levels of error, changes in population aggregates between two observations can be used as a basis to estimate the number of events between the two observations. Estimation methods

based on changes in stocks are all residual methods, so results are sensitive to even quite small errors in the components.

The estimation of mortality through intercensal survival provides a simple example. Suppose that population censuses have been held in 1990 and 2000 in a population that has experienced negligible migration. The population aged 30–34 in 2000 represents the survivors of the population aged 20–24 in 1990. If the data are accurate, the survivorship ratio approximates a standard life table function:

$$\frac{{}_5L_{30}}{{}_5L_{20}} \approx \frac{{}_5P_{30}^{2000}}{{}_5P_{20}^{1990}}$$

where ${}_5L_x$ is the life table person-years lived between ages x and $x+5$, and ${}_5P_x^y$ is the population aged x to $x+5$ in year y . Thus, in principle, a life table after early childhood can be derived for any population with two census age distributions that has experienced little migration, though in practice, the method is adversely affected by age misreporting errors, particularly at older ages, and possibly by age-differential under- or over-count.

Fertility can be estimated from changes in the average parity (or average number of children ever borne) of a cohort of women between two observations. The change in average parity measures their cumulative fertility between the two observations, if selection effects through attrition—by death or migration—are negligible. A cumulative fertility distribution for a hypothetical cohort experiencing the fertility rates of the period between the two observations can then be obtained by summing the cohort changes. Period age-specific rates can be estimated from the hypothetical cohort parities in a number of ways, perhaps the simplest of which is fitting of the relational Gompertz fertility model proposed by British demographer William Brass in 1981.

If accurate population estimates—such as from successive census counts—are available for two points in time, net migration can be estimated as a residual from the Demographic Balancing Equation, which expresses the identity that population change during the time interval between the two population counts must be equal to the difference between the additions to the population (births and in-migrants) and the losses (deaths and out-migrants). Thus

$$NM^{t_1-t_2} = P^{t_2} - P^{t_1} - B^{t_1-t_2} + D^{t_1-t_2}$$

where NM is the number of net migrants between t_1 and t_2 , P is the population at time t_1 or t_2 , and B and D are the numbers of births and deaths respectively between t_1 and t_2 . Applying this method to the United States, with a 1990 population count of 248.710 million, a 2000 population count of 281.421 million, and 39.837 million and 22.775 million intercensal births and deaths, net migration between the two censuses is estimated as 15.649 million.

Consistency Checks

Consistency checks seek to compare two or more measures of the same demographic parameter. Consistency between two measures is a necessary, but not sufficient, condition of their validity. If they are found to differ, assumptions about patterns of error can under certain circumstances provide a basis for adjustment, to obtain unbiased measures of the parameter in question even if both data sources are themselves biased. Brass proposed two consistency checks that have been widely used in demographic estimation.

The P/F Ratio Method (Brass 1964) compares current and lifetime measures of fertility. Current fertility estimates for developing countries with weak statistical systems may be in error for a number of reasons: births may not all be registered, and responses to survey questions on both recent births and birth histories may suffer from omission or misreporting of date of occurrence. Brass suggested a simple consistency check for situations in which fertility is not changing rapidly and additional information is available on each woman's lifetime fertility. Age-specific fertility rates can be cumulated from the start of childbearing to obtain measures equivalent to lifetime fertility (for a hypothetical cohort) at exact ages. Measures F comparable to average parities P for five year age groups can then be obtained by interpolating between the point values using standard fertility models (for instance, the relational Gompertz fertility model mentioned above). Consistency is assessed by calculating ratios of average parity P to interpolated, cumulated age-specific fertility F' for each age group.

Brass then goes further to argue that typical errors in the age-specific fertility rates, for example omission of births from registration, may not vary by age, therefore all the F' values will be incorrect by

a constant factor, whereas the reporting of children ever born, the basis of the P 's, may be most accurate for younger women. Thus, when the P/F ratios indicate inconsistency, the ratios for younger women may provide appropriate adjustment factors for the age-specific fertility rates at all ages. A simple assumption about error patterns turns a consistency check into an adjustment method.

In practice the simple assumptions just stated may be incorrect: current and lifetime fertility may be inconsistent because fertility is changing, rather than because of data errors. Changing fertility can be accommodated if information on lifetime fertility is available for two time periods, allowing the calculation of lifetime fertility for a hypothetical cohort.

Brass (1975) proposed a way to use the Demographic Balancing Equation to evaluate information on deaths by age. The equation can be written in terms of rates, and also for age groups. In a population experiencing negligible migration, the open-ended age group x and over ($x+$) experiences exits only through deaths at ages x and over, and entries only through birthdays at age x . Thus

$$b(x+) = r(x+) + d(x+)$$

where $b(x+)$, $r(x+)$, and $d(x+)$ are the entry (birthday), growth, and death rates for the age segment $x+$. $b(x+)$ can be estimated from an age distribution as $N(x)/N(x+)$, where $N(x)$ is an estimate of the population passing through age x in a year and $N(x+)$ is the population aged $x+$. Similarly, $d(x+)$ can be estimated as $D(x+)/N(x+)$, where $D(x+)$ is deaths in a year at ages x and over. If deaths are reported with completeness c , constant at all ages relative to the population numbers, then $d(x+) = (1/c)D^o(x+)/N(x+)$, where $D^o(x+)$ is observed deaths x and over. If the population is then assumed to be demographically stable, the growth rate $r(x+)$ is constant for all x . Thus

$$N(x)/N(x+) = r + (1/c)D^o(x+)/N(x+)$$

If the assumptions are correct, the birthday rates and the observed death rates over a range of ages x should be linearly related; the intercept estimates r , the stable growth rate, and the slope estimates $(1/c)$, the reciprocal of the completeness of death registration. Once again, by making simplifying assump-

tions, the consistency check (of death rates based on recorded deaths against death rates computed from the difference between entry rates and growth rates) provides a basis for adjustment.

If information is available about the population age distribution at two points in time, the assumption of stability can be relaxed, and the last expression can be written as

$$b(x+) - r^o(x+) = k + (1/c)d^o(x+)$$

where $r^o(x+)$ is the observed growth rate of the population $x+$, and k can be interpreted as the error in the growth rate due to change in enumeration completeness.

Indirect Estimation

Indirect estimation seeks to estimate a demographic parameter that is difficult to measure directly from some indicator that can be accurately recorded and is largely, but not exclusively, determined by the parameter of interest. The effects of confounding variables on the indicator are then allowed for, so that the parameter of interest can be estimated.

The most widely used example, due to Brass, estimates infant and child mortality from the proportion dead among children ever borne by women classified by age. Prior to the widespread use of birth histories in countries with deficient demographic statistics, infant and child mortality were especially hard to measure because of omission of early infant deaths from registers or retrospective reports. Brass realized that the proportion dead among children ever born was largely determined by the level of child mortality, but was also affected by the time location of the women's births prior to the survey and by the age pattern of mortality risk in childhood. The older the women, the longer on average their children would have been exposed to mortality risk and hence the higher, other things being equal, would be the proportion dead. However, controlling for women's ages, exposure would also be longer in a population of early childbearers than a population of late childbearers, and hence the former would have a higher proportion dead than the latter. Brass used simple fertility and child mortality models to simulate proportions dead for different fertility patterns to develop conversion factors to adjust an observed proportion dead for the effects of the age pat-

terns of childbearing. His initial method has been extended by several authors, increasing the range of model patterns, extending the technique to data classifying women by duration of marriage, and placing reference dates on the estimates in order to estimate trends.

The Brass method and its successors greatly increased knowledge of levels and trends in childhood mortality in the developing world. Although the widespread use of birth histories in surveys has reduced the need to apply the method for national level estimates, the simplicity of the questions needed, and hence the ability to include them in population censuses, makes the method ideal for small area estimates of levels and trends of child mortality.

Indirect methods have also been developed to estimate demographic parameters from population age distributions assuming stability, to estimate adult mortality from proportions of respondents with surviving mother or surviving father and from proportions of brothers and sisters surviving. A method based on survival of sisters has been developed to measure maternal mortality, which is difficult to measure because the events are relatively rare and cause of death is often misclassified. Wendy Graham and colleagues suggest asking female respondents about the survival of their ever-married sisters, and identifying presumed maternal deaths by whether a dead sister was pregnant, delivering, or within two months of delivery at the time of death. Arguing that maternal deaths would follow approximately the pattern of overall fertility, except for rather higher numbers at young and old ages to reflect higher risks, the authors developed a method for estimating the lifetime risk of maternal death by extrapolating from the partial experience of each age cohort. An estimate of total fertility was then used to convert the measure of lifetime risk into the more widely used indicator, the Maternal Mortality Ratio. The indirect estimate obtained from this method refers to a time point at least 12 years before the survey, and the indirect approach has been largely superseded by the use of direct measurement based on a complete sibling history.

See also: *Actuarial Analysis*; Brass, William; *Data Assessment*; *Fertility Measurement*; *Life Tables*; *Mortality Measurement*; *Population Dynamics*.

BIBLIOGRAPHY

- Arretx, Carmen. 1973. "Fertility Estimates Derived from Information on Children Ever Born Using Data from Censuses." *International Population Conference, Liège 1973*, Vol. 2: 247-261.
- Brass, William. 1964. "Uses of Census or Survey Data for the Estimation of Vital Rates." Paper presented to the African Seminar on Vital Statistics, Addis Ababa, December 1964.
- . 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data*. Chapel Hill, NC: University of North Carolina.
- . 1981. "The Use of the Gompertz Relational Model to Estimate Fertility." *International Population Conference, Manila 1981*, Vol. 3. Liège: International Union for the Scientific Study of Population.
- Brass, William, and Kenneth Hill. 1973. "Estimating Mortality from Orphanhood." *International Population Conference, Liège 1973*, Vol. 3. Liège: International Union for the Scientific Study of Population.
- Coale, Ansley, and Paul Demeny. 1968. *Methods for Evaluating Basic Demographic Measures from Limited and Defective Data*. New York: United Nations.
- Feeney, Griffith. 1980. "Estimating Infant Mortality Trends from Child Survivorship Data." *Population Studies* 34: 109-128.
- Graham, Wendy, William Brass, and R. W. Snow. 1989. "Estimating Maternal Mortality: The Sisterhood Method." *Studies in Family Planning* 20: 125-135.
- Hill, Kenneth. 1987. "Estimating Census and Death Registration Completeness." *Asian and Pacific Population Forum* 1: 8-24.
- Hill, Kenneth, and James Trussell. 1977. "Further Developments in Indirect Mortality Estimation." *Population Studies* 31: 313-334.
- Sullivan, Jeremiah. 1972. "Models for the Estimation of the Probability of Dying Between Birth and Exact Ages of Early Childhood." *Population Studies* 26: 79-97.
- Timæus, Ian. 1992. "Estimation of Adult Mortality from Paternal Orphanhood: A Reassessment and a New Approach." *Population Bulletin of the United Nations* 33: 47-63.

- Timæus, Ian, Basia Zaba, and Mohamed Ali. 2001. "Estimation of Adult Mortality from Data on Adult Siblings." In *Brass Tacks: Essays in Medical Demography*, eds. Basia Zaba and John Blacker. London: The Athlone Press.
- Trussell, James. 1975. "A Re-estimation of the Multiplying Factors for the Brass Technique for Determining Childhood Survivorship Rates." *Population Studies* 33: 97–107.
- United Nations. 1983. *Manual X: Indirect Techniques for Demographic Estimation*. New York: United Nations.

KENNETH HILL

ETHICS, POPULATION

See *Animal Rights; Asylum, Right of; Data Collection, Ethical Issues in; Environmental Ethics; Euthanasia; Future Generations, Obligations to; Genetic Testing; Reproductive Technologies: Ethical Issues*

ETHNIC AND NATIONAL GROUPS

The term *ethnic group* has no single agreed-on definition in English. Its Greek origins refer to a nation or a people, but by the fifteenth century the word "ethnic" had a connotation of "heathen." A typical modern American definition (Merriam-Webster's Collegiate Dictionary) is "a: of or relating to large groups of people classed according to common racial, national, tribal, religious, linguistic, or cultural origin or background; b: being a member of an ethnic group; c: of, relating to, or characteristic of ethnics." However, if "religious," "racial," and "national" groups are included under the general rubric of "ethnic," what explains the common usages "ethnic and national groups," "ethnic and religious groups," and "racial and ethnic groups"? Such linguistic ambiguities cannot be resolved easily; therefore, the use of the term in this article is broad and inclusive.

The Use of Classifications

Governments around the world routinely collect demographic data classified according to ethnic, na-

tional, tribal, racial, linguistic, cultural, and other categories. The categorizations are typically country-specific—deemed to be suitable to their particular circumstances—thus, cross-country comparisons are difficult and can be misleading. As an example, although some states collect detailed data on the national origins or religions of their residents, others consider these categories improper in official statistics.

Classifications of ethnic and national groups within a country often are linked to sensitive political issues of power, control, and contested territorial boundaries. Therefore, they typically are determined by political rather than scientific or technical decisions.

Overall, the distinguishing quality of data on ethnic and national groups is variability. Even states with otherwise similar recent histories, cultures, political structures, and economic systems (e.g., the liberal democracies of Europe, North America, Asia, and Oceania) demonstrate marked differences in the manner in which they collect, classify, and report data on ethnic and national groups. Moreover, even within a state attitudes toward parallel categories of data are often inconsistent, applied differentially depending on the degree of historical and political sensitivity.

For example, in the U. S. censuses conducted from 1790 through 2000 more or less detailed data have been collected on race, national origin, ethnic origin, and ancestry. However, over the same period of more than two centuries essentially no official U.S. data have been collected on religion.

Changes in the U.S. Census

Early U.S. racial categories were simple—white, Negro, and Indian—but there have been numerous subsequent revisions, mostly in the direction of greater complexity. Until 1960 a respondent's "race" was determined by census interviewers; since that time it has been based on self-identification on a mailed questionnaire. As a result of pressure from a few influential members of Congress, the 2000 U.S. Census "race" classifications had no fewer than sixteen categories, with most of the new categories identifying a substantial number of distinct "races" in the Asian and Pacific regions. Many of these categories (e.g. "Samoan," "Guamanian or Chamorro," "Asian Indian," "Filipino," "Korean," and "Vietnamese") would more commonly be treated as national rather than racial categories.