

Workshop Two

Use of Discrete-Time Survival Models to Draw Out the Policy Implications of Demographic and Health Surveys

Demographic and health surveys (DHS) provide information on a number of issues of concern to policymakers. Analysis of survey data can reveal trends in fertility, age at marriage and first childbirth, contraceptive use, birth spacing, and breastfeeding, among many other measures. By revealing causal relationships, data analysis can show the effects of factors such as education, urban/rural residence, economic status, and the utilization of health programs on women's reproductive behavior. Careful analysis can also indicate how well family planning and maternal and child health programs are working.

Life-table methodology is appropriate for analyzing DHS data that indicate the time period or interval between a starting event and a terminating event. Several important measures of fertility and health are expressed in this way. Examples include the interval between a woman's birth and her first marriage or the birth of her first child, between the birth of one of her children and the next, between the time when she starts using family planning and when she stops, and between when she gives birth and when she stops breastfeeding.

Discrete-time survival models (also called hazard models) can be viewed as multivariate life tables. They allow the calculation of life tables (and specific measures calculated from life tables) by categories of one characteristic, such as urban/rural residence, while controlling for (i.e., holding constant) other characteristics, such as education.

From a policy perspective, it is important to know which characteristic actually affects important reproductive health variables such as fertility. A simple tabulation of survey results does not answer this important question because characteristics such as urban/rural residence and education, for example, tend to be correlated, meaning that urban persons tend to be more educated than rural persons. A multivariate analysis using discrete-time survival models can show policymakers which characteristic is really having an effect on fertility—urban/rural residence or education—and which characteristic only appears to have an effect because of correlation. Although the methodology is rather sophisticated, such an analysis provides clear information to policymakers in the form of simple tables and charts.

Until recently, discrete-time survival models have been used to analyze cohort data, but not period data. Cohort data pertain to the historical experience of real groups of persons, such as a birth cohort (persons born in a particular calendar year). By contrast, period data pertain to events (for example, third births) and women's exposure to risk of these events in a particular time period (e.g., the year 2000). The application of discrete-time survival models to period data enables multivariate analysis of period measures such as total fertility rate, infant mortality rate, mean and median ages at marriage, and mean and median birth intervals. The methodology also makes possible the multivariate analysis of trends in these period measures.

Another feature of discrete-time survival models is their flexibility and utility for causal analysis. This flexibility, based on separate specifications of calendar time

and life-table time, offers great potential for solving difficult problems of two-way causation. This is achieved by means of lagged predictors, time-varying predictors, and time-varying effects of predictors. For example, a variable representing a woman's work status can be lagged one or two years behind the time in the life table where she might have a third birth, so that causation runs clearly from work to fertility but not from fertility to work.

Discrete-time survival models come in several forms. The most common is the discrete-time logit model. A less common form is the complementary log-log (CLL) model, which is derived from the continuous-time Cox proportional hazards model. The workshop will cover the discrete-time logit model but will focus primarily on the CLL model.

The workshop organizers have developed computer programs for applying discrete-time survival models to DHS data. These programs, which are currently being rewritten in STATA, are for multivariate analysis of levels and trends in total fertility rates, total marital fertility rates, parity progression ratios (including proportion marrying by age 40), mean and median ages at first marriage, and mean and median closed birth intervals. They can also be modified for analyzing levels and trends in other measures, such as infant mortality.

■ Workshop Objective

The basic objective of the workshop is to apply discrete-time survival models to DHS surveys in a number of countries. In most cases, the application will be to marriage and/or fertility data, for which the com-

puter programs need very little modification. It may be feasible for more advanced participants to modify the programs for multivariate analysis of other measures such as infant mortality.

Each participant will either bring to the workshop or will be furnished with a DHS survey, where possible from the participant's own country. Workshop sessions will be divided between presentations on methodology, presentations on the basics of STATA, and hands-on applications of the programs to individual DHS surveys. Each participant will present the results of his or her application of the programs during the last few days of the workshop.

■ Who Should Attend?

The workshop is designed for researchers actively engaged in the analysis of demographic survey data, especially DHS data. Participants should have a good understanding of and experience with multiple regression and logit regression. Some familiarity with the conventional Cox proportional hazards model is also desirable but not required. In addition, participants should be familiar with and have used at least one statistical programming package, but not necessarily STATA.

■ Coordinators

Robert D. Retherford (Ph.D. University of California, Berkeley) is a Senior Fellow and Coordinator for Population and Health Studies at the East-West Center. His research interests include fertility and family planning, infant and child mortality, and population aging in Japan, China, and other countries of Asia. His work in Japan analyzes fertility decline, changes in the institution of marriage, effects of population ag-

ing on the family, and how socioeconomic development and changing values interact with various aspects of demographic change. He is coauthor of *Statistical models for causal analysis* (1993) and *The future of population in Asia* (2002).

Minja Kim Choe (Ph.D. University of Hawaii, Manoa) is a Senior Fellow in Population and Health Studies at the East-West Center. Dr. Choe's research focuses on South Korea, Japan, and other Asian countries, as well as the United States. Her interests include family and gender issues, health behavior of adolescents and young adults, fertility and reproductive health, child survival, and statistical analysis of demographic processes. She is coauthor of *Statistical models for causal analysis* (1993) and coeditor of *The changing family in comparative perspective: Asia and the United States* (1998).

Jiajian Chen (Ph.D. University of Western Ontario) is a Senior Fellow in Population and Health Studies at the East-West Center. His research interests include the HIV/AIDS epidemic in Asia and emerging health and population challenges in China and other countries of East Asia.

■ Resource Person

Hassan Eini-Zinab (M.A. Shiraz University) is a Ph.D. student in sociology at the University of Hawaii at Manoa and a Research Project Assistant in Population and Health Studies at the East-West Center. Mr. Eini-Zinab has worked on reproductive health and fertility in Iran. He wrote the computer programs that will be used during this workshop for the application of discrete-time survival models to DHS data on total fertility rates.