Eighth Population Census Conference Examines 1980–81 Experiences

by Robert L. Hearn

Officials from 17 census and statistical offices in Asia and the Pacific met at the Eighth Population Census Conference where they shared their experiences with the 1980–81 round of censuses and talked with researchers and representatives of international organizations about developments in census taking and data analysis. Cosponsored by the East-West Population Institute and the National Bureau of Statistics of the Republic of Korea and held at the Korea Development Institute in Seoul, the conference covered a wide variety of topics, including development and population policy, statistical mapping, migration, fertility and mortality estimation, and census evaluation. The agenda also included reports on the problems the participants encountered during their recent censuses.

Country Reports

Although the presenters focused on the problems of the censuses rather than the successes, the country reports were generally delivered with a sense of relief. Much work may remain tabulating, evaluating, and publishing census results, but the awesome task of preparing for and conducting the actual enumeration is over.

The problem area mentioned most often was the recruitment and training of enumerators. For several countries this was the main problem they faced. Since most censuses attempt a full count of the population and take place in a brief period, the number of enumerators required is large. If the labor market is tight, as it was in Hong Kong, Malaysia, Papua New

Robert L. Hearn is managing editor of the Census Forum.
Guinea, Singapore, and Thailand, simply finding workers to fill the positions is formidable. To make matters worse, the job being offered is only temporary and the pay is usually low. Yet, in order to get good data, the enumerator must understand the questionnaire and be able to elicit meaningful responses from respondents. In addition, training must be uniform to ensure that enumerators conduct the census in the same manner.

Hong Kong chose high-form students as enumerators but ran into a problem when the census was scheduled close to the time for their examinations. Many students who had already been trained dropped out of the operation and more enumerators had to be recruited. Singapore’s solution was to conscript recent high school graduates who were waiting to go to permanent jobs, India and Thailand recruited teachers.

There were different opinions on the efficacy of using students as enumerators. Bangladesh found that students made good enumerators, but India and Pakistan complained that students were not dedicated to the work and were not sufficiently disciplined. Korea deliberately recruited university students on the assumption that they would be better able than most to understand the census operation. Unfortunately, the students often were not local residents and had trouble locating households in the areas to which they were assigned. In Japan, students seemed less likely than older enumerators to intimidate the respondents and were thought to get truer responses, but they were also more easily refused the interview.

Another common problem mentioned in connection with recruitment of enumerators was political interference. Local politicians often use their influence to determine who will or will not be hired. The abuse of this form of patronage in one state in India required the intervention of the prime minister before it was halted.

Enumerator training was another problem. P. Padmanabha pointed out the difficulty of maintaining uniformity of instruction in India, where 1.25 million enumerators were employed. Singapore used university facilities to centralize training, and the air-conditioned classrooms helped trainees sustain their attention, which was otherwise difficult in the warm, humid climate. Hong Kong showed enumerators nine one-hour films on the instructional televisions in public schools, thus ensuring that they all received the same instructions. Papua New Guinea and Nepal had the additional task of training enumerators to reach the remote, mountainous regions of their countries.

The representatives of Pakistan and India commiserated on the phoenix nature of their decennial censuses. Their census operations thrived for the three or so years it takes to set up, conduct, and process the census, but then they reduce themselves to glowing embers for the remaining intercensal period. As a new census approaches, the operation must be rebuilt from the ashes, as it were. Akhtar Hasan Khan of Pakistan suggested that a quinquennial census would keep census offices open and pointed out that the cost of the census is but a small part of the development budget while the data collected are essential to effective planning. A note of warning on adopting quinquennial censuses was sounded by the Australian and New Zealand representatives, both of whom reported an overlapping of work from their 1976 censuses with that of their 1981 censuses.

Although reports about computer-readable forms were generally favorable, especially in regard to time saved during data entry, participants from both the United States and New Zealand reported problems getting their forms printed to the critical specifications required. Australia has recently completed an extensive survey of optical mark recognition (OMR) and optical character recognition (OCR) systems, and Frank Parsons of the Australian Bureau of Statistics said this information could be made available to other countries investigating OCR and OMR equipment.

Maps remain a common source of trouble. In Papua New Guinea, the maps recently completed by the Australian Survey Corps proved of little use because there were few people trained in map reading. Census officials circumvented the problem by tapping local knowledge to help locate people in the rural areas. The use of outdated base maps to make Korea’s census maps sometimes led to ambiguous enumeration-district boundaries. In Nepal and Pakistan, cartographic materials were simply inadequate to the task, and in Thailand the major cities were growing so rapidly that it was difficult to define enumeration-district boundaries.

Representatives from India and the U.S. both stressed the need for a better system of cost monitoring for field operations. Early warning of nearing the spending limit is especially important in the U.S. because obtaining additional funds involves the lengthy process of conducting Congressional hearings and justifying additional appropriations.

Political pressure was mentioned in conjunction with several aspects of the census. Besides interference in the recruitment of enumerators, it took the form of efforts to include additional items on the questionnaires. In New Zealand there was pressure to release figures before an election and in India and the U.S. there was pressure to change the census counts to enable local governing bodies to qualify for financial assistance. Over 50 lawsuits were filed against the U.S. Bureau of the Census, disputing census counts and requesting adjustments (see

---

**IN THIS ISSUE**


Alice Harris has graciously consented to relinquish the space for her column so that we may bring you all of the articles listed above. Publications That Count will return next issue.
Eighth Population Census Conference
28 September–2 October 1981

COORDINATORS
Dr. Lee-Jay Cho, President Pro Tem, East-West Center;
Dr. Dai-Young Kim, Director General, National Bureau of Statistics, Republic of Korea;
Mr. Jay-Soo Park, Fellow, East-West Population Institute;
Dr. Sam Suharto, Director of Population and Social Statistics, Central Bureau of Statistics, Indonesia

PARTICIPANTS
Australia: Mr. Frank Parsons, Director, Processing & User Services Section, Population Census Branch, Australian Bureau of Statistics
Bangladesh: Dr. A. K. M. Ghulam Rabbani, Director General, Bureau of Statistics
Hong Kong: Mr. Benjamin Mok, Commissioner for Census and Statistics
India: Mr. P. Padmanabha, Registrar General and Census Commissioner
Indonesia: Mr. Abdulmadjid, Director General, Central Bureau of Statistics
Japan: Dr. Toshio Kuroda, Professor of Demography, Nihon University; Mr. Shoichi Ueda, Statistical Councilor, Statistics Bureau
Republic of Korea: Dr. Hakchung Choo, Research Director, Korea Development Institute; Mr. Pil-Bong Han, EDI Analyst, Data Processing Division, National Bureau of Statistics; Mr. Young-Kwon Kim, Director, Population Statistics Division, National Bureau of Statistics; Mr. Kap Suk Koh, Research Coordinator, Korean Institute for Population and Health; Dr. Sung-Yeal Koo, Senior Fellow, Korea Development Institute; Mr. Kyong-Jong Yoo, Deputy Director General, National Bureau of Statistics
Malaysia: Mr. Teik Huat Khoo, Chief Statistician, Department of Statistics
Nepal: Mr. K. R. Sharma, Deputy Director, Central Bureau of Statistics
New Zealand: Mr. Michael A. Moore, Senior Research Officer, Department of Statistics

Pakistan: Dr. Akhtar Hasan Khan, Census Commissioner
Papua New Guinea: Mr. Reg Gilbert, Census Director; Mr. Nick Suvulo, Deputy Director, Central Bureau of Statistics
Republic of Singapore: Ms. Chian Kim Khoo, Chief Statistician, Department of Statistics
Sri Lanka: Mr. Wettasinghe A. A. S. Periris, Director of Census and Statistics
Taiwan: Mr. Chung-Li Chang, Executive Secretary of Committee on Census; Mr. Tze-Hwa Fan, Senior Statistician, Directorate-General of Budget, Accounting & Statistics
Thailand: Mrs. Wiwit Siripak, Acting Director, National Statistical Office
United States: Mr. Samuel Baum, Director, International Demographic Center, U.S. Bureau of the Census; Dr. John C. Chao, Demographer, Office of Population, Agency for International Development; Mr. Meyer Zitter, Assistant Director for International Programs and Acting Associate Director for Demographic Fields, U.S. Bureau of the Census
United Nations (ESCAP): Mr. Badr Hanna, Senior Project Expert, Population Division
World Bank: Mr. Barry Wayne Luscombe, Cartographer
East-West Population Institute: Dr. Griffith Feeney, Research Associate; Mr. Robert L. Hearn, Writer/Editor; Dr. James A. Palmore, Jr., Research Associate

ADMINISTRATION
This conference was cosponsored by the National Bureau of Statistics, Economic Planning Board, Republic of Korea and the East-West Population Institute. Many people from both institutions worked hard to make the conference a success. From the National Bureau of Statistics, Mr. Young-Kwon Kim coordinated the Seoul arrangements with the able assistance of Mr. Hyo-Bong Chung, Mr. Seung-Kon Lim, and others on his staff. Preconference arrangements were made from Honolulu by Ms. Janet Mason of EWPI; Mr. Robert Hearn of EWPI assisted with administrative arrangements and served as conference rapporteur in Seoul.

The staff of the Korea Development Institute efficiently provided facilities and assistance during the conference sessions at KDI, and several other Korean organizations graciously provided services and facilities during the course of the conference.

The conference. James A. Palmore presented a theoretical paper that attempted to define some of the issues, Kap Suk Koh described how Korea has used census data to guide its national family planning program, and Hakchung Choo reflected on the social scientist's need for accurate and reliable data. The census officials responded to these presentations with their own views on the subject.

Palmore listed four emerging concerns. The first is about the linkage between data producers and policy-makers and between the data producers themselves—there is a need for administrative and managerial arrangements to coordinate the various data producers and users (the census office, economic planning board, family planning agency, etc.).
A second concern is that policy goals are shifting, thus requiring new types of data. Palmore cited the discussion of optimum population in a Malaysian National Family Planning Board report as an example of the new interest in levels of population as opposed to growth rates. Other new interests are differentials in fertility and mortality, population redistribution and social mobility, the effects of the size and structure of the population on demand for goods and services, and social indicators.

A third concern is the need for small-area statistics to monitor the distribution of the benefits of development. Data and projections are needed to identify areas in need of remedial action, to plan for population movement and labor force availability, and to make sure local development is consistent with national plans. Palmore said that the only good rationale for doing a census instead of a sample survey was to allow tabulations for small areas.

The fourth concern is for the linkage of data sources to permit the combination of data from different offices for analysis. This linkage can cut costs and more efficiently use the data now available. It requires both a routine for accessing the data from another office and that the data be compatible. Palmore cited a Korean survey that saved space on its questionnaire by not asking questions that had already been asked on the census; but for this to work, there had to be some means of matching the survey to the census.

Kap Suk Koh focused on the use of census data in the implementation of Korea’s national family planning program. The success of the program was based on the reduction of the annual rate of population growth instead of less easily measured indicators such as health conditions of the population. The early goals were ambitious: a reduction from 2.9 to 2.5 percent between 1962 and 1966, and a reduction from 2.5 to 2.0 percent between 1967 and 1971. But census statistics warned that the fourth 5-year plan would have to contend with the entry of the baby-boom generation into its reproductive years. Consequently, for the period 1977–81, the government set a realistic goal of reducing the growth rate from 1.7 to 1.6 percent.

Population statistics also helped determine target segments of the population. Differentials in urban and rural fertility rates prompted officials to focus on rural areas in the early years of the program. The 1970 census, however, showed rapid urbanization, which suggested that high-fertility norms of rural areas were being transplanted to urban centers. The government responded with new policies promoting hospital, industry, and military family planning programs and special projects for the low-income urban population.

The census infrastructure, including the enumeration districts, maps, and population composition data, has been used for Korea’s knowledge, attitudes, and practice (KAP) surveys which evaluate the family planning program.

Hachung Choo said the social scientist needs reliable and abundant statistics for analysis and planning. As an example he pointed out the need to consider the population’s age structure in formulating Korea’s latest 5-year plan. The goal of 8 percent annual economic growth was set to avoid high unemployment and underutilization of the economically active population which is expected to grow at a rate of 3 percent per year until the 1990s. In addition, data on population mobility were used to determine housing needs and to help make decisions on the location of industry. Choo said the social scientist needs qualitative data, that is, data on the traits and expectations of the population, which change as development progresses and often determine the success of development.

When it came time for discussion, it was noted that “Development and Population Policy” could encompass “nearly everything under the sun,” but the conference participants soon focused on the limits of the census. Sam Suharto asked whether the census operations, which already seem overburdened, could realistically be expected to meet the new needs for information described by Palmore and Choo. Frank Parsons described Australia’s experience of overloading the 1976 census questionnaire: the public felt that the long questionnaire was too much to bear, that the census was an invasion of privacy, and hence cooperation was low.

P. Padmanabha said each country must decide what it needs from its census. In India’s development plan calls for meeting minimum needs, so the census collected data on household facilities—water, electricity, etc.—and will tabulate the results on a village-by-village basis to indicate where action is needed. Teik Huat Khoo told of Malaysia’s emphasis on monitoring the distributive effects of development policy, that is, who the beneficiaries of development are. He also stated that although census takers are aware of the vast need for information, the census is not the sole vehicle for collecting statistics—it is only part of a total data collection system including surveys and vital registration.

A. K. M. Ghulam Rabbani of Bangladesh said priorities must be set before the census is planned. He suggested that a national committee determine the use of the census before it is designed. Several other participants also expressed the need for meetings between policy-makers, academicians, and census officials before the census is planned. These meetings would determine the type of information the census should produce, possibly even the questions that should be asked.

Samuel Baum of the U.S. made the observation that future (continued on page 17)
Selectivity and the Analysis of Birth Intervals from Survey Data

by Ronald R. Rindfuss, James A. Palmore, and Larry L. Bumpass

An increasing number of demographers are turning their attention to the study of birth intervals because the fertility process is itself a sequential and time-dependent process. Birth interval analysis allows more precision in investigating many fundamental questions; it allows the assessment of the effects of intermediate variables, like contraceptive use or lactation, and the explication of the effects of various socioeconomic variables in terms of intermediate variables.

One factor promoting the increased analysis of birth intervals has been rapid growth in the availability of the necessary data. Birth interval data can, of course, be obtained from a variety of sources: population registers (in Taiwan, for example), prospective studies (like those carried out in Bangladesh as discussed in National Research Council 1981), and even some vital registration systems. Nevertheless, surveys containing retrospective birth histories are the predominant source of interval data, and a quantum leap in the availability of such surveys was made with the advent of the World Fertility Survey. This technical note focuses exclusively on issues raised by data available from retrospective birth histories obtained in cross-sectional surveys.

Three principal methodological issues have been raised about birth interval analysis. The first involves data quality. Respondents are asked to remember the dates of events which may have taken place many years in the past. They may forget events, which would lengthen birth intervals, or they may misdate events, which could either lengthen or shorten intervals. A number of strategies are available for evaluating the quality of birth interval data (Potter 1975; Cabigon 1976; Rindfuss and Bumpass 1979) depending on the nature of the data.

The second methodological concern has come to be known as the “censoring” problem. Censoring occurs when birth intervals are interrupted by the interview itself. Many of these “open” intervals will eventually be “closed” by the next fertile pregnancy, although the timing of this closure is unknown. Since the open intervals are biased towards long intervals, they cannot be excluded from the analysis. Since the closed intervals are biased towards short intervals, they cannot be excluded from the analysis. This is a well-recognized problem. The usual demographic solution involves the use of a life table (Smith 1980; Rodriguez and Hobcraft 1980; Rindfuss et al. in press) or the incorporation of a life table approach into a multivariate procedure (Menken et al. 1981; Bumpass et al. in press; Prentice and Kalbfleisch 1980; Cox 1972, MacDonald and Rindfuss 1981). Since the censoring problem has been discussed extensively in the literature, that discussion is not repeated here.

The third methodological problem is that the intervals available for analysis may be only a selected subset of the universe of interest. Such selectivity has received far less attention than either censoring or data quality. A major exception is the recent paper by Rodriguez and Hobcraft (1980) that carefully distinguishes between censoring and the selectivity that results from the fact that not all of the women in a sample who ultimately enter a particular birth interval have done so by the time of the interview. A more general selectivity is addressed in the present paper. The principal question is: given the range of birth intervals available in the typical cross-sectional fertility survey, which intervals should be included in an analysis? The answer to this question can determine whether the results are biased, i.e., are an artifact of the sets of birth intervals included. Unlike the life table solution to censoring, there is no single solution to selectivity. The most appropriate strategy will depend on the substantive question being asked. In the remainder of this paper, the nature of selectivity is outlined and solutions are illustrated for several substantive problems.

Intervals Available for Analysis

A hypothetical cross-sectional survey is used to explicate the issues involved. Figure 1 is a Lexis diagram representing the intervals available for analysis in a survey of ever-married women under age 50 conducted on 1 January 1971. The horizontal dimension represents time up to the date of interview and the vertical dimension indicates the ages of women at entry into birth intervals. The extreme right vertical line indicates the ages of women at the time of the interview.

Birth intervals are represented in the Lexis plane by the births that initiate them, or, in the case of the first birth interval, by marriage. Thus points on the diagram represent these initiating events as designated by the date of the event and the age on that date of the woman experiencing the event. Areas in the diagram represent sets of initiating events. Any vertical band represents birth intervals begun in the same period (a parity cohort if limited to intervals of the same order); any horizontal band represents birth intervals begun to women of the same age, and any diagonal band from upper left to lower right represents the experience of a birth cohort of women.

Only those intervals initiated within the upper triangle enclosed by solid lines are available for analysis in a survey with an upper age limit of 50. Initiating events that would be located in the lower triangle enclosed by the broken lines occur to women who would be 50 or older at the time of the interview, and hence this experience goes unreported. Women who were 49 on the first of January, 1971, were 15 in 1936, so the years in which intervals could have begun range from 1936 through 1970.

Ronald R. Rindfuss is a professor in the Department of Sociology at the University of North Carolina at Chapel Hill; James A. Palmore is a research associate at the East-West Population Institute and director of the Population Studies Program and professor of sociology at the University of Hawaii; and Larry L. Bumpass is a professor in the Center for Demography and Ecology at the University of Wisconsin–Madison.
Figure 1 Lexis Representation of Birth Intervals Selected for Analysis by a Survey on 1 January 1971 of Women under Age 50. Birth intervals are represented in the Lexis plane by the births that initiate them, or, in the case of the first birth interval, by marriage. Initiating events are represented by points (imagined, not shown), and sets of these events by areas in the Lexis plane. Observed experience refers to the triangle bound by solid lines and is experience available for observation from the birth histories collected in the interview. Unobserved experience refers to the triangle bound by the diagonal and the broken line; it lies outside the range of experience of the women interviewed and is therefore unavailable for observation.

Looking first at the age-at-beginning-of-interval groups, (the horizontal bands), it is clear that the range of time when birth intervals could have begun is not invariant over ages at the beginning of an interval. The length of the period diminishes as age at initiation increases. At the extremes, intervals begun at age 15 could have been initiated in any year between 1936 and 1971; and intervals begun at age 49 could have been initiated only in 1970. The bias shifts systematically with age between these extremes; for example, intervals begun at age 30 before 1951 are omitted. In Figure 1, births at age 30 since 1936 are represented by the band AC of which only the BC portion (since 1951) is available for analysis. Thus it is clear that each age at initiation represents a different historical period. This point is likely to be unimportant for adjacent ages but it can become critical as the span of ages compared widens, especially when there has been rapid change in fertility over time.

Next, compare different periods during which the birth intervals were begun (the vertical bands). If we limit the intervals to those begun by women of a specific parity, we compare parity cohorts. Recent periods are relatively unbiased with respect to either age at beginning of interval or birth cohort of mothers because the upper age limit affects ages at which few births occur. Earlier periods, however, are progressively selected toward younger ages at beginning of interval. This introduces a clear bias. In the example indicated by band DE (intervals initiated in 1951) only the births to women under age 30 (the DB segment) are available for analysis. Obviously this selectivity increases as one goes further back in time; at the extreme, intervals begun in the late 1930s are represented only by those begun by women in their teens.

Finally, compare the various birth cohorts of women as represented by their age at interview on 1 January 1971 (as would be represented by diagonal bands). It is obvious that cohorts vary considerably in the possible ages at the beginning of an interval. The youngest cohorts are necessarily biased toward a younger age at initiation, and the range of ages at initiation increases directly with age at interview because of the truncation of cohort experience by the interview. Thus any uncontrolled comparisons between different birth cohorts of women is biased by the differing ranges of age at initiation.

The biases pointed out above could be safely ignored if age at the beginning of an interval, birth cohort, and time period were irrelevant to fertility. There are, however, reasons to believe that all three dimensions are strongly related to the rate at which birth intervals are closed. Because of declining fecundity with age, one would expect the age at the beginning of a birth interval to be related to the pace at which that interval is closed. There are also sociological reasons why the age at the beginning is related to the length of the birth interval; for example, concern about the appropriateness of further childbearing at older ages or after becoming a grandmother favors shorter intervals at younger ages (Rindfuss and Bumpass 1978; Bumpass, Rindfuss, and Janosik 1978). Further, a host of period factors could affect the pace at which birth intervals are closed. Examples include wars, food shortages, the introduction of alternatives to breastfeeding, and increases in education. Perhaps the example most discussed in the literature is the introduction of a family planning program (e.g., see Rindfuss et al. in press). Finally, as Ryder (1959) has emphasized, there are numerous reasons why birth cohort effects may influence fertility. Clearly, one cannot ignore these potential biases.

While it may at first seem natural to analyze all of the intervals reported by retrospective histories, to do so biases the analysis toward younger ages at initiation, toward the experience of older birth cohorts, and toward the experience of most recent time periods. Those who routinely use multivariate analyses might be tempted simply to include two of these dimensions as control variables. However, this approach is unsatisfactory if the effect of any of the other variables in the model on the pace of closing the birth interval is dependent on period, cohort, or age at initiation, and it is difficult to imagine situations where this is not the case. Particularly for rapidly modernizing third world countries, period is likely to interact with the other variables in the model. For example, one would expect the effect of education on the pace of closing the birth interval to change after the successful introduction of a family planning program. Recognizing this problem, it may be tempting to include the various interactions with the other variables in the model. Such an approach would utilize all the information in the survey, but it has several inherent problems. If there are more than a couple of other predictor variables, the number of interaction terms needed would be sufficiently large as to render their interpretation all but impossible, and
An Algebraic Illustration of the Extent of Selectivity Bias

by Chai Bin Park, EWPI Research Associate

Retrospective investigations of time-dependent events have many practical advantages over prospective investigations, especially in the areas of cost, manpower, and time. Unfortunately, biases are more likely to creep into retrospective studies than prospective ones. Epidemiologists, who frequently employ the retrospective method in their attempts to establish an association between a disease and some factor, have long been concerned with the representativeness of the cases and controls in their samples—a problem that is most pronounced when the sample is a cross section of patients alive at a given time. Classical discussions of the problems can be found in the literature as far back as the 1950s, for example in Cornfield (1956) and in Mantel and Haenszel (1959).

In their technical note in this issue of the *Census Forum*, Rindfuss, Palmore, and Bumpass present a typical instance of selectivity in retrospective data stemming from the age cutoff of the interviewees. The authors correctly state that “Recent periods are relatively unbiased,” but “earlier periods are progressively selected toward younger ages.” The interviewees may be a representative sample as of the survey date, but they are not representative for any period much earlier than that.

As an example, suppose $a$ and $b$ represent the beginning and the end of the childbearing age in women. On the survey date we interview women ranging in age from $a$ to $b$. For births (or events related to birth) $t$ years before the survey date, the ages of women at that time range from $a$ to $(b - t)$. Thus, the childbearing ages from $(b - t)$ to $b$ lie outside the range of the study. Many fertility-related variables vary with woman’s age. Therefore, the exclusion of women aged $(b - t)$ to $b$ creates bias in the estimated fertility measures.

The extent of such bias can be algebraically demonstrated. Suppose we are investigating a measure $r$ using retrospective birth history data. It may be birth interval, or it may be infant mortality, or any other such measure that varies by woman’s age. For simplicity, consider the population divided into two subpopulations: 100p percent of the people belong to one subpopulation and 100(1 - $p$) percent to the other. Let the measure for the first group be denoted by $r_1$ and that for the second group by $r_2$. The measure for both groups combined is

$$r = pr_1 + (1 - p)r_2$$

(1)

Suppose the second group is entirely out of our observation at a certain reference period as were the women aged $(b - t)$ to $b$ in the example above. Then we are left with $r_1$, instead of $r$, as the observed measure for the population.

Let us define the relative bias (RB) as

$$RB = \frac{r - r_1}{r}$$

(2)

Assuming that $k$, the relationship between the two subpopulations, is equal to $r_2/r_1$, we can substitute (1) into (2) and simplify to obtain

$$RB = \frac{(1 - p)(k - 1)}{p + (1 - p)k}$$

(3)

Thus the relative bias is a function of the proportion of the population omitted from observation $(1 - p)$ and the relationship between the measurements for the two subpopulations $(k)$.

The accompanying figure and table present the extent of relative bias for selected values of $(1 - p)$ and $k$. Frequently the measurement for one subpopulation is more than twice that of the other (i.e., $k > 2$ or $k < .5$) as is the case with infant mortality of young mothers compared with infant mortality of old mothers. Although the value of $k$ is rarely constant over a study period, it is clear from the figure that for any value of $k$ other than one, the larger the proportion of the population omitted from observation, the greater the relative bias.

REFERENCES


<table>
<thead>
<tr>
<th>$1 - p$</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>-0.33</td>
<td>-0.176</td>
<td>0.00</td>
<td>0.200</td>
<td>0.333</td>
<td>0.500</td>
</tr>
<tr>
<td>0.4</td>
<td>-0.250</td>
<td>-0.136</td>
<td>0.00</td>
<td>0.167</td>
<td>0.286</td>
<td>0.444</td>
</tr>
<tr>
<td>0.3</td>
<td>-0.176</td>
<td>-0.099</td>
<td>0.00</td>
<td>0.130</td>
<td>0.231</td>
<td>0.375</td>
</tr>
<tr>
<td>0.2</td>
<td>-0.111</td>
<td>-0.064</td>
<td>0.00</td>
<td>0.091</td>
<td>0.167</td>
<td>0.266</td>
</tr>
<tr>
<td>0.1</td>
<td>-0.053</td>
<td>-0.031</td>
<td>0.00</td>
<td>0.048</td>
<td>0.091</td>
<td>0.167</td>
</tr>
<tr>
<td>0.05</td>
<td>-0.026</td>
<td>-0.015</td>
<td>0.00</td>
<td>0.024</td>
<td>0.048</td>
<td>0.091</td>
</tr>
</tbody>
</table>

SOURCE: Computed from equation (3) in the text.
with such a large number of interaction terms in the model, very high levels of multicollinearity are probable.

Thus the approach advocated throughout this paper for the avoidance of selectivity bias is to analyze only a subset of all available intervals. Which subset depends on the substantive questions being asked. In subsequent sections of this paper, various strategies are presented for addressing differentials and trends in birth interval analyses. But several additional points need to be made first.

Most of the World Fertility Surveys use an upper age limit of 49 as illustrated here. Some surveys use younger upper age cutoffs; e.g., most of the United States fertility surveys have an upper age cutoff of 44 and most of the Taiwan surveys have an upper age cutoff of 39. As the upper age limit of the sample becomes younger, the number of parity cohorts that are complete, or almost complete, becomes smaller. Furthermore, the younger the upper age cutoff of a sample, the fewer are the number of birth cohorts with completed fertility.

Some fertility surveys interview all ever-married women without an upper age cutoff. The various Philippines NDS surveys are examples of this. This reduces but does not eliminate some of the problems discussed here. Instead, it makes them less visible. First, even without an upper age cutoff, intervals begun at a young age are going to be overrepresented because the youngest birth cohorts have not yet had a chance to initiate intervals at an old age. Second, the force of mortality effectively imposes an upper age cutoff. If relatively old birth cohorts are included, their experience represents only a selected subset of all the intervals begun by that cohort; for example, in the Coale-Demeny West model life table with $e_0 = 60$, one third of women reaching age 45 do not reach age 70 (Coale and Demeny 1966).

So far the discussion has referred to all birth intervals, with no distinction as to birth order. Rather than pooling all intervals in one analysis, an analyst would typically be interested in intervals of a particular order, say those begun with a first birth. Clearly, biases in age at initiation change in importance from one order to the next because ages at initiation increase as birth order increases, moving the distribution into the range where it will be truncated.

Intervals begun close to the time of the survey present an additional problem. The point is clearest with respect to intervals begun during the survey year. Suppose that the interviewing was done during the middle of the calendar year. Intervals begun during that year will have, on average, three months of exposure, and none of the intervals will be closed. Including such intervals in the analysis amounts to "padding" the number of intervals being analyzed because these intervals contribute nothing. This problem persists for intervals begun as far back as nine months before the survey, after which it diminishes.

**Differentials**

There is considerable interest in the effects on birth intervals of intermediate variables (such as lactation or contraception) or of socioeconomic variables. As discussed earlier, it would be a mistake to include all available intervals in such analyses be-
These points aside, it would seem that the comparison of education groups within relative age categories would be unbiased. Unfortunately, that is not true and the reason is closest for the youngest age at initiation group. The problem is not a result of the kinds of selectivity being discussed here but rather is a consequence of any significant trend over time in the predictor variable. As can be seen in Figure 2, the youngest age initiation category represents experience between 1936 and 1971. If the diffusion of education has been relatively recent, intervals initiated by more educated women will represent primarily recent experience whereas those of women with no education will be heavily weighted by the early periods. If there has also been a significant trend in fertility this would lead to an artifactual relationship between education and fertility within this age-at-initiation category.

A second approach used by Rodriguez and Hobcraft is to make comparisons within categories of the period in which the intervals were begun. This is illustrated by the areas enclosed in the vertical lines of Figure 2. The major problem with this approach is that the different periods have different maximum upper ages at which the interval could have begun. For the most recent period, the intervals could have begun at any age, up through 49. For birth intervals begun in the year 1951, the latest age at which the interval could have begun is 30. For certain intervals, like the second birth interval, this is a trivial problem. For the later birth intervals beginning at later ages, however, it can become a more substantial problem; for example, comparisons, say by education, can be made without any age bias, but they may be limited to an age range within which little of the experience of interest typically occurs. Note further that this approach becomes more problematic if the upper age cutoff of the survey is less than age 49.

The approach we would recommend is illustrated in Figure 3. This strategy follows that of Rodriguez and Hobcraft noted above but restricts the analysis to intervals begun in some recent period. In a parity-specific analysis this would of course represent a recent parity cohort. If the period is appropriately chosen, then such a parity cohort is relatively unbiased with respect to either age at beginning of interval or birth cohort of women. Figure 3 shows the example of intervals initiated in the decade ending a year before the survey. The intervals begun during the year preceding the survey are excluded for reasons discussed earlier, the exact length of this omitted period being somewhat arbitrary.

Even for intervals initiated recently, the area at the bottom of the diagram indicates that some experience is omitted by the upper age limit of the sample. The intervals contained in this area will always be those begun at a relatively late age (39 for the earliest year in the illustrated 10-year parity cohort). Note that the choice of a decade of experience here is arbitrary, and the appropriate choice of period will depend on the particular substantive problem being addressed. As the period being analyzed is expanded, the number of intervals available for analysis increases, but so also does the proportion of intervals that are omitted, thus increasing the age-at-beginning-of-interval and birth-cohort biases. For example, if the period were extended to 1955, as indicated by the dotted line in Figure 3, the earliest year would omit experience after age 34. An increase in the proportion of intervals omitted is tolerable for the early-order birth intervals but not for the higher-order intervals.

In analyzing the intervals contained in this selected cohort of initiating births, it is possible to add whatever controls are desired. In particular, it is possible to control on age at beginning of interval (which could be represented in the figure by horizontal divisions). Comparisons among such age groups will be unbiased with respect to period except for the oldest age group which involves the bottom triangle.

This approach discards information which was often collected at great expense. Doing so, however, provides a set of intervals to be analyzed which are unbiased by age at beginning of interval, period when the interval began, or birth cohort. Furthermore, the experience being discarded is that which occurred most distant from the survey and therefore is less likely to be reliable data.

Finally, it is important to have the flexibility to decide which parity cohorts to analyze. Thus, data collection procedures should not make this decision in advance by collecting information only on intervals begun in a specified recent period. No matter what period is chosen, for some unforeseen substantive problem, data from an earlier time point are likely to be needed. Further, putting such a restriction on the questionnaire unnecessarily complicates it, increasing the possibility of data collection errors.

**Trend Analysis**

The other major use of birth interval data is to examine change over time. Perhaps the most common strategy is to divide the intervals on the basis of birth cohorts of women and compare the early cohorts with later ones. The problem with this approach has already been noted: as one moves from the oldest to the youngest cohorts, there is successively more bias toward
young ages at the beginning of an interval and toward intervals beginning in more recent periods. At the extremes, those born 15 years before the survey only have one possible age at initiation (15) whereas those born 49 years before the survey have 35 possible ages at initiation (15-49).

Controlling for age at initiation partially solves this bias, but also introduces problems of its own. This approach is illustrated in Figure 4. For convenience, the illustration uses 5-year age-at-beginning categories. The following discussion, however, would apply to any set of age categories. Within any age-at-initiation category, the parallelograms of experience can be compared with one another without any concern about biased results. The assumption, of course, is that there is no significant trend within any parallelogram. If there is, then the boundaries need to be made narrower. Unfortunately, the exact period represented by each parallelogram is almost impossible to describe. For example, the intervals begun in the upper left parallelogram were begun primarily in the years 1940-42. But some were begun as early as 1936 and some as late as 1945. Then the next one to the right begins in 1941. Furthermore, for the oldest age-at-initiation category of each cohort, the age-at-beginning-of-interval bias still exists, and this is visually apparent in the triangles at the right of the figure. Finally, because this approach requires dividing the data into many cells, like any other cross-classificatory procedure, it runs a high risk of having too few cases in some of the cells.

Another possible approach is to control on period, as was illustrated in Figure 2. The age-at-initiation bias associated with this approach was discussed earlier. To try to control for this age-at-initiation bias, Rodriguez and Hobcraft (1980) introduced a control for relative age, illustrated in Figure 2 by the horizontal lines. In addition to the relative age problem discussed earlier, this approach introduces several other problems. First, as in the procedure discussed above, the numbers of birth intervals in some of the categories are likely to get quite small. Second, the relative age control only partly eliminates the age-at-initiation bias. Clearly, comparisons made within any of the three rows of rectangles in Figure 2 will be unbiased. Comparisons across time for the oldest relative age group, however, will still be biased with respect to age at beginning of interval.

A preferable approach for examining trends in birth intervals is illustrated in Figure 5. The approach is simply to focus on parity cohorts which are restricted to the same oldest age at initiation. This cutoff is defined by the oldest possible age at initiation in the earliest parity cohort included. The advantage of this approach is that the comparisons across time are completely unbiased. In the example shown in Figure 5, intervals begun in the time period 1955-70 are being examined. The earliest parity cohort is 1955. For an interval to have begun in 1955 and still be included in the sample, it had to have begun to a woman aged 34 or less at the time. Thus intervals begun by women 35 and above are excluded from the analysis.

There are several features of this approach that are arbitrary and have to be decided on the basis of the substantive questions being examined. First, as before, a decision has to be made about including intervals begun during the year of the interview and during the years immediately preceding the interview year. Second, decisions have to be made about the time span of the parity cohorts. In the illustration here, widths of five years are shown. In other circumstances, wider or narrower widths may be needed. In a recent analysis of Korean childspacing trends, Rindfuss et al. (in press) focused on single-year parity cohorts.

Finally, a decision has to be made regarding how far back in time to go. The further back in time the trend analysis goes,

(continued on page 15)
The Effect of Alternative Matching Procedures on Fertility Estimates Based on the Own-Children Method

by Michael J. Levin and Robert D. Retherford

In the own-children method of fertility estimation, enumerated children in a census or household survey are first matched to mothers within households, usually on the basis of answers to questions on relation to head of household, age, sex, marital status, and number of living children. The matching algorithm checks that relationships and ages of mother and child are compatible, that the mother is not single, and that she is not matched to more children than she states are still living. These matched (i.e., own) children, classified by own age and mother’s age, are reverse-survived to estimate births by age of mother in previous years, and unmatched (non-own) children are distributed by age of mother according to the distribution of own children by age of mother. Reverse survival is similarly used to estimate the number of women by age in previous years. After adjustments are made for incorrect enumeration, age-specific birth rates are calculated by dividing the births by the women. (For methodological details, see Grabill and Cho 1965; Cho et al. 1970; Cho 1973; Ho 1978; Retherford and Cho 1978.)

The matching algorithm, which is one of several computer programs used in applying the own-children method, sometimes mismatches. For example, an adopted child may be erroneously matched instead of classified as non-own if the census or survey fails to distinguish adopted children from biological children. Or there may be more than one eligible woman in the household (same relation to head, same marital status, etc.), in which case the child may be matched to the wrong woman since the algorithm matches the child to the first eligible woman encountered in the household search. Because of the possibility of mismatch, some censuses and household surveys instruct interviewers to record mother’s person number (i.e., her line number in the household listing) for each child whose biological mother is in the same household. Presumably, the direct match based on mother’s person number (MPN matching) matches fewer children than the indirect match based on relation to head of household (RHH matching), since there is, in the former case, less ambiguity about the mother-child relationship and therefore a reduced probability of erroneously matching the child when the biological mother is actually absent.

Although a direct question on mother’s person number increases the accuracy of matching, it also adds to the cost of the census or household survey. It is therefore of interest to examine how much difference MPN matching, compared with less costly RHH matching, makes in the ultimate fertility estimates. To this end, we examine in this paper two sets of own-

---

Michael J. Levin is a social science analyst in the Population Division of the U.S. Bureau of the Census, and Robert D. Retherford is a research associate at the East-West Population Institute.

---

children fertility estimates, one for American Samoa and the other for East Java Province in Indonesia (the choice of one province instead of all of Indonesia was made in order to economize on computing costs). In American Samoa, adoption is common, households are generally large and complex, and the frequency of in- and out-migration (mainly to the United States) is very high, leading to frequent temporary separations of family members; under these conditions, the likelihood of mismatch is much higher for RHH matching than for MPN matching. In East Java, on the other hand, households are simpler and smaller and the population is less migratory; under these conditions, the likelihood of mismatch is low for both types of matching. Our expectation is, then, that MPN matching should improve the accuracy of the fertility estimates more for American Samoa than for East Java.

Data

Results for American Samoa are based on the 100 percent 1974 census count comprising 29,190 individuals. Results for Indonesia are based on the 1976 Indonesian Intercensal Population Survey, also known as SUPAS II, which included a sample of 35,822 individuals from the province of East Java. Mean household size from these data is 6.9 in American Samoa and 4.6 in East Java. Both of these counts asked mother’s person number as well as relation to head of household and therefore allow computation of fertility estimates based alternatively on RHH and MPN matching.

Relation-to-head-of-household codes used in RHH matching are

for American Samoa:

1. head
2. wife of head
3. son, daughter, nephew, or niece of head
4. grandchild of head
5. brother, sister, brother-in-law, or sister-in-law of head
6. father, mother, father-in-law, or mother-in-law of head
7. other relative
8. no relation
9. unknown

for Indonesia:

1. head
2. wife of head
3. own child of head
4. non-own child of head (adopted child or step-child)
5. grandchild of head
6. parent of head
7. parent of wife of head
8. daughter-in-law or son-in-law of head
9. other family
10. other non-family
11. unknown

Since the Indonesia codes are somewhat more detailed than the American Samoa codes, they allow a more precise match of children to mothers by means of RHH matching; hence the results reported below for the two countries are not strictly comparable. The RHH match algorithm is modified in each application to take advantage of whatever codes are available.

Findings

Table 1 shows the percentage that non-own (i.e., not-matched) children are of all children in each single-year age group according to type of match. As expected, the percentage not matched is higher in American Samoa than in East Java. Also expected is the greater divergence between RHH and MPN matching in American Samoa than in East Java. The percentage not
Table 1 Non-own children as percentage of all children by type of matching

<table>
<thead>
<tr>
<th>Age</th>
<th>American Samoa 1974</th>
<th>East Java 1976</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RHH</td>
<td>MPN</td>
</tr>
<tr>
<td>0</td>
<td>24.8</td>
<td>13.2</td>
</tr>
<tr>
<td>1</td>
<td>23.2</td>
<td>15.5</td>
</tr>
<tr>
<td>2</td>
<td>26.2</td>
<td>18.6</td>
</tr>
<tr>
<td>3</td>
<td>22.4</td>
<td>17.2</td>
</tr>
<tr>
<td>4</td>
<td>24.6</td>
<td>19.2</td>
</tr>
<tr>
<td>5</td>
<td>25.0</td>
<td>21.0</td>
</tr>
<tr>
<td>6</td>
<td>22.2</td>
<td>17.5</td>
</tr>
<tr>
<td>7</td>
<td>24.1</td>
<td>20.3</td>
</tr>
<tr>
<td>8</td>
<td>23.5</td>
<td>20.1</td>
</tr>
<tr>
<td>9</td>
<td>22.8</td>
<td>20.2</td>
</tr>
<tr>
<td>10</td>
<td>24.5</td>
<td>21.9</td>
</tr>
<tr>
<td>11</td>
<td>27.0</td>
<td>24.1</td>
</tr>
<tr>
<td>12</td>
<td>22.4</td>
<td>21.5</td>
</tr>
<tr>
<td>13</td>
<td>29.0</td>
<td>27.9</td>
</tr>
<tr>
<td>14</td>
<td>29.5</td>
<td>27.7</td>
</tr>
</tbody>
</table>

NOTES: Non-own children are those not matched to mothers within the same household. RHH denotes matching based on relation to head of household. MPN denotes matching based on mother’s person number.

matched rises steeply with age of child mainly because older children are more likely than younger children to live in a household other than their mother’s. An unanticipated exception is RHH matching in American Samoa, where the percentage not matched is unusually high among younger children and the rise with age in the percentage not matched is largely eliminated. The difference between RHH and MPN matching in American Samoa is especially great for younger children.

The most startling finding in Table 1 is that in both American Samoa and Indonesia the percentage not matched is higher for RHH matching than for MPN matching, the reverse of what was anticipated. To discover why the MPN match matches more children than the RHH match, we examined in detail each individual match in a household listing from a 25 percent systematic sample of American Samoa’s enumeration districts. (This was not done for East Java, since discrepancies are so small.) The MPN match produced 815 unmatched, or non-own, children. Of these, only 11 were erroneously matched by the RHH match, indicating that overmatching by the RHH algorithm is not a serious problem. Offsetting these 11 children who were overmatched were 107 children who were correctly matched by MPN matching but erroneously not matched by RHH matching, owing to errors in relationship codes. These errors appear to consist mainly of respondent errors, not interviewer errors. Quite commonly, for example, the household head incorrectly reported a child as grandchild, while correctly reporting the child’s mother (identified by MPN matching) as “other relative” instead of daughter or daughter-in-law as would be necessary if the child were truly a grandchild. If the household lacked an eligible daughter or daughter-in-law, RHH matching then incorrectly designated the child as non-own. This kind of error was especially common for very young children, for whom the reported grandparent-grandchild relationship evidently often reflects an affectionate social tie rather than a biological tie, and it is consistent with the large discrepancies between RHH and MPN matching at very young ages in the percentage that non-own children are of all children in the left half of Table 1. In all, because of errors in relationship codes, RHH matching produced $815 + 107 - 11 = 911$ non-own children in the sample, 12 percent more than MPN matching.

The results in Table 1 suggest that fertility estimates based alternatively on RHH and MPN matching should be similar for East Java but somewhat divergent for American Samoa, particularly in the years just previous to enumeration, for which fertility estimates are based on reverse survival of very young children. Table 2 and Figures 1 and 2 confirm this expectation. In the case of East Java, fertility estimates based alternatively on RHH and MPN matching differ by less than 2 percent, and usually less than 1 percent, over single calendar years between 1962 and 1976. In American Samoa, on the other hand, the discrepancy, though small in years close to 1960, increases for years closer to the census date, reaching almost 4 percent in 1974. In American Samoa, estimates of the total fertility rate

![Figure 1](own-children-estimates-of-total-fertility-rates-for-american-samoa-source-table-2.jpg)

![Figure 2](own-children-estimates-of-total-fertility-rates-for-east-java-source-table-2.jpg)
based on RHH matching are consistently higher than estimates based on MPN matching. In East Java, the much smaller discrepancies are in both directions and show no clear pattern. The consistent direction of the discrepancy in American Samoa stems from a systematic distortion in the age pattern of fertility, to which we now turn.

Table 3 examines the effects of alternative matching procedures on the age pattern of fertility. Calendar years are grouped into two 5-year periods in order to minimize effects of heaping of children’s ages on preferred digits on the fertility estimates. The effects of such heaping are evident in the jagged fertility trends observed earlier in Figures 1 and 2.

In East Java, the difference between the age patterns of fertility estimated alternatively using RHH and MPN matching is inconsequential, consistent with the small differences in matching results and TFR estimates shown earlier in Tables 1 and 2 and the figures. In American Samoa, however, RHH matching results in underestimates of fertility at the younger reproductive ages and overestimates at the older reproductive ages, especially for the second 5-year period, 1968–72. This evidently occurs because RHH matching erroneously allocates too many children to older women and not enough to younger women, relative to results based on MPN matching. The underlying reasons for this misallocation are not entirely clear, but one source of error is the method’s assumption that non-own children are distributed by age of mother in the same proportions as own children. This assumption is violated in American Samoa, where non-own children are in reality concentrated among younger mothers. Because the number of women in each 5-year age group decreases with age, the shift of children

(continued on page 17)
On Grading Infant and Childhood Mortality with Mathematical Curves

by S. Krishnamoorthy

Perhaps the earliest attempt to represent the life table by a mathematical function was made by de Moivre in 1725. In 1825 Gompertz and in 1867 Makeham developed formulas that can be used for adult ages but do not fit very well at younger ages (see Hooker and Langley-Cook 1953: 157–164; Smith and Keyfitz 1977: 273–288). Opperman proposed a formula in 1870 for graduating infant and childhood mortality (Steffensen 1930: 273–276). He represented the force of mortality at age \( x \) by

\[
\mu(x) = \frac{a}{\sqrt{x}} + b + c\sqrt{x}
\]

where \( a, b, \) and \( c \) are parameters to be estimated using the observed data. Even this curve, however, does not fit to an acceptable accuracy.

Recent proposals by Keyfitz (1966), Hartman (1980), and Choe (1981) are decidedly better than the earlier ones. However, no comparisons have been made between these three proposals. Using Australian data, this paper attempts to compare the proposed models to check their fit. In addition, since we have data for the first-year life by months, we can test their ability to fit the earliest period of life.

Recent Models

Let \( \mu(x) \) be the force of mortality at age \( x \), then the probability of surviving to age \( x \) is

\[
x(x) = \exp \int_0^x -\mu(t) \, dt
\]

Keyfitz (1966: 295) proposed a hyperbolic function for the survival probabilities, namely,

\[
x(x) = \frac{ax + b}{x + b}
\]

Hartman (1980: 23) proposed the following logarithmic approximation formula:

\[
x(x) = \frac{b(1 - a)}{(ax + b)(x + b)}
\]

and showed that this fits much better than Opperman’s curve. The Weibull hazard function, which has been used in reliability theories for a long time (Gross and Clark 1975), has been recommended by Choe (1981) for application to the study of infant and child mortality. Choe has also shown the Weibull function’s superiority over the Gompertz model. The Weibull survivorship function is

\[
x(x) = e^{-\lambda x^\gamma}
\]

The hazard functions, the survivorship functions, and the transformations to the linear form of the survivorship functions of these three models are presented in Table 1. A comparison of the linear forms of the logarithmic and Weibull functions shows that \( \ln x \) is linearly related to \( \ln \ln \left[1/x(x)\right] \) or \( \logit x(x) \). For values of \( x(x) \) closer to unity, as is always the case when we consider infant and childhood ages, these two quantities are nearly equivalent (apart from the factor of 0.5). Hence we expect that if these two models fit the data, they should perform equally well.

Data

The Australian Bureau of Statistics provides reasonably accurate data on mortality and provides death statistics by age in months for the first year of life. This permits us to compute survivorship by month till the end of first year of life and survivorship by single years of age till the tenth birthday. The relevant statistics are extracted from Australian Bureau of Statistics (1981a, b).

Findings

The linear forms of the logarithmic and Weibull models given in Table 1 are used to develop normal equations, and unweighted least squares is used for estimating the concerned parameters. Similarly, an unweighted-least-squares solution is obtained for the hyperbola. The observed survivorship function for Australian male children for 1979 and the residuals for the three models are given in Table 2. Table 3 presents the corresponding information for female children. These tables indicate that the hyperbola suggested by Keyfitz is a poor fit when compared to the other two, particularly for the first year of life. The logarithmic and Weibull models fit very well,

\[
x(x) = \frac{1}{1 + e^{2a} x^{2b}}
\]

Table 1 Three Models for Fitting Infant and Childhood Survivorship

<table>
<thead>
<tr>
<th>Models</th>
<th>Hyperbola (Keyfitz)</th>
<th>Logarithmic approximation</th>
<th>Weibull (Choe)</th>
</tr>
</thead>
</table>
| Hazard function \( \mu(x) \) | \[
\frac{b(1 - a)}{(ax + b)(x + b)}
\] | \[
\frac{2 b e^{2a} x^{2b - 1}}{1 + e^{2a} x^{2b}}
\] | \[
\lambda y x^{\gamma - 1}
\] |
| Survivorship function \( x(x) \) | \[
\frac{ax + b}{a + b}
\] | \[
\frac{1}{1 + e^{2a} x^{2b}}
\] | \[
e^{-\lambda x^\gamma}
\] |
| Survivorship transformed to linear form | \[
0.5 \ln \left[\frac{1 - x(x)}{x(x)}\right] = a + b \ln x
\] | \[
\ln \ln \left[\frac{1}{x(x)}\right] = \ln \lambda + \gamma \ln x
\] |
Table 2: Observed and Graduated Survivorship of Australian Male Children Using Hyperbolic, Logarithmic, and Weibull Functions

<table>
<thead>
<tr>
<th>Age</th>
<th>Observed survivorship (^a)</th>
<th>Hyperbolic function</th>
<th>Logarithmic function</th>
<th>Weibull function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>99,175</td>
<td>494</td>
<td>-41</td>
<td>-41</td>
</tr>
<tr>
<td>Months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>99,083</td>
<td>333</td>
<td>-45</td>
<td>-45</td>
</tr>
<tr>
<td>3</td>
<td>98,065</td>
<td>249</td>
<td>-22</td>
<td>-22</td>
</tr>
<tr>
<td>4</td>
<td>98,942</td>
<td>103</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>98,883</td>
<td>160</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>98,848</td>
<td>122</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>7</td>
<td>98,822</td>
<td>88</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>8</td>
<td>98,807</td>
<td>54</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>9</td>
<td>98,782</td>
<td>27</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>10</td>
<td>98,778</td>
<td>6</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>11</td>
<td>98,761</td>
<td>-13</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>98,743</td>
<td>-17</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>98,649</td>
<td>-95</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>98,594</td>
<td>-106</td>
<td>-16</td>
<td>-18</td>
</tr>
<tr>
<td>4</td>
<td>98,539</td>
<td>-87</td>
<td>-20</td>
<td>-20</td>
</tr>
<tr>
<td>5</td>
<td>98,488</td>
<td>-59</td>
<td>-14</td>
<td>-14</td>
</tr>
<tr>
<td>6</td>
<td>98,459</td>
<td>-45</td>
<td>-23</td>
<td>-23</td>
</tr>
<tr>
<td>7</td>
<td>98,425</td>
<td>-22</td>
<td>-22</td>
<td>-22</td>
</tr>
<tr>
<td>8</td>
<td>98,389</td>
<td>5</td>
<td>-15</td>
<td>-15</td>
</tr>
<tr>
<td>9</td>
<td>98,362</td>
<td>25</td>
<td>-14</td>
<td>-14</td>
</tr>
<tr>
<td>10</td>
<td>98,337</td>
<td>45</td>
<td>-13</td>
<td>-13</td>
</tr>
</tbody>
</table>

\( a \) Expected – observed.
\( b \) Survivorships are given per 100,000 births for the sake of convenience. They may be divided by 100,000 to get the probabilities.

Table 3: Observed and Graduated Survivorship of Australian Female Children Using Hyperbolic, Logarithmic, and Weibull Functions

<table>
<thead>
<tr>
<th>Age</th>
<th>Observed survivorship (^a)</th>
<th>Hyperbolic function</th>
<th>Logarithmic function</th>
<th>Weibull function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>99,317</td>
<td>412</td>
<td>-19</td>
<td>-19</td>
</tr>
<tr>
<td>Months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>99,250</td>
<td>274</td>
<td>-29</td>
<td>-30</td>
</tr>
<tr>
<td>3</td>
<td>98,190</td>
<td>203</td>
<td>-13</td>
<td>-13</td>
</tr>
<tr>
<td>4</td>
<td>99,140</td>
<td>157</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>99,104</td>
<td>119</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>99,079</td>
<td>85</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>99,061</td>
<td>55</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>98,047</td>
<td>29</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>98,036</td>
<td>7</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>99,015</td>
<td>0</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>98,005</td>
<td>-15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>98,993</td>
<td>-24</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>98,906</td>
<td>-74</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>98,846</td>
<td>-68</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>98,803</td>
<td>-53</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>98,776</td>
<td>-44</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>6</td>
<td>98,754</td>
<td>-35</td>
<td>-15</td>
<td>-15</td>
</tr>
<tr>
<td>7</td>
<td>98,734</td>
<td>-23</td>
<td>-21</td>
<td>-22</td>
</tr>
<tr>
<td>8</td>
<td>98,702</td>
<td>2</td>
<td>-13</td>
<td>-13</td>
</tr>
<tr>
<td>9</td>
<td>98,683</td>
<td>15</td>
<td>-15</td>
<td>-15</td>
</tr>
<tr>
<td>10</td>
<td>98,655</td>
<td>35</td>
<td>-6</td>
<td>-6</td>
</tr>
</tbody>
</table>

\( a \) Expected – observed.
\( b \) Survivorships are given per 100,000 births for the sake of convenience. They may be divided by 100,000 to get the probabilities.

and based on their fit to these sets of data, it is difficult to say which one is better.

The parameters \( \lambda \) in the Weibull model and \( a \) in the logarithmic model may be interpreted as the level parameters determining the level of childhood mortality, whereas the parameters \( \gamma \) in the Weibull model and \( b \) in the logarithmic model may be viewed as the shape parameters because they determine the rapidity with which the decrement due to mortality takes place.

Concluding Remarks

Weibull and logarithmic functions fit infant and childhood survivorship better than a parabolic function. Based on their fit to the data sets, it is difficult to judge which is better between Weibull and logarithmic functions. However, the Weibull function has already been used frequently in reliability theory and the analysis of patient mortality, and computer program packages are available for the application of the Weibull function (e.g., Dixon and Brown 1979). In addition, the Weibull function is more elegant for further analytic treatment than the logarithmic function. Hence, the Weibull hazard function is recommended for fitting the age patterns of infant and childhood mortality.

REFERENCES


SELECTIVITY (continued)

the younger must be the upper age at initiation cutoff. This is illustrated by the dotted lines in Figure 5: going back another 5 years reduces the oldest age at initiation from 34 to 29. For the earliest parity transitions, a young upper age cutoff may not be problematic, but it would be clearly undesirable for the
later parity transitions. Sometimes different upper age cutoffs will have to be used for different parities. For the most recent parity cohorts, it is possible to determine the effect of the upper age cutoff—that is, it is possible to see the proportion of intervals being omitted. For the earlier parity cohorts that is not possible unless one has information from other data sources. Careful thought must be given to the trade-off between the length of the period observed and the severity of the upper age constraint. While trend analysis for samples such as that illustrated by the dotted lines in Figure 5 (with an age cutoff of 29) are unbiased with respect to age, there is obviously a point beyond which the substantive question being addressed is quite limited by the upper age constraint. The choice will depend upon the issues at hand as well as upon the order of the interval being analyzed.

Summary and Conclusion

In summary, birth intervals represented in retrospective surveys pose a number of potential selectivity biases that have not been generally recognized. This technical note reviews the nature of these biases and suggests separate strategies for coping with them for analyses of differentials and of trends. Because of the increased interest in birth interval analysis, and in related issues such as breastfeeding, infant mortality, and contraceptive usage, it is important that appropriate solutions to the selectivity problem be incorporated into analysis. Further, it must be emphasized that the selectivity problems discussed here are not unique to birth interval analysis. They apply to any time-dependent process represented in surveys. Most directly analogous are concerns with intervals between marital events, such as studies of divorce and remarriage; but studies of occupational transitions or migration histories will be affected as well.

ACKNOWLEDGMENTS

This technical note was begun under grant No. HD-12071 from the National Institute of Child Health and Human Development to the East-West Population Institute (EWPI) and completed under a Special Services Agreement between EWPI and the International Statistical Institute and the World Fertility Survey/London. The authors are indebted to participants in the Workshop on Trends and Differentials in Childspacing Practices, Twelfth Summer Seminar in Population, East-West Population Institute, for provocative questions that led to the development of this paper. The authors are also grateful to Griffith Feeney who encouraged the preparation of this note, to Robert Rutherford for his comments on an earlier draft, and to colleagues at the Rand Corporation and our respective home institutions for discussions about the issues in this paper.

REFERENCES


from younger women to older women increases birth rates at the older reproductive ages more than it decreases birth rates at the younger reproductive ages. This may explain why the TFR computed from age-specific birth rates based on RHH matching slightly but consistently exceeds the TFR computed from birth rates based on MPN matching.

Proportional errors in age-specific birth rates for American Samoans are as high as 8 percent at ages 15–19 and as high as 44 percent at ages 44–49. Absolute errors in these rates are, of course, much smaller than the percentage errors suggest, because fertility in these extreme age groups is very low. Moreover, the errors are largely offsetting. Hence errors in the TFR are quite small. Note that errors in the TFR are smaller in Table 3 than in Table 2 and the figures; aggregation of calendar years reduces the discrepancy in the TFR between the two types of matching.

Conclusion

The above findings suggest that the extra coding costs necessitated by using the mother's-person-number match in place of the relation-to-head-of-household match are justifiable when migration rates are high, households are large and complex, and relationship codes are not very detailed. In this case, use of MPN matching instead of RHH matching can result in significant improvement of the fertility estimates. But when migration rates are low, household structure is simple or only moderately complex, and relationship codes are fairly detailed, RHH matching works almost as well as MPN matching, and the ultimate fertility estimates are affected very little by the choice of matching procedure. Were RHH matching expensive, MPN matching might still be advantageous. But RHH matching is very inexpensive: the cost of RHH matching at the current University of Hawaii computer charge of about US$2.25 per hour is approximately US$0.54 per thousand households, plus about two weeks of programming time to modify the RHH computer algorithm to fit the country's particular data format. MPN matching may, of course, be justified for reasons other than fertility estimation, as, for example, in analyses of family and household structure. In such cases, findings presented in this paper suggest that edit programs should be developed to check for consistency of responses to the relation-to-head question. Census or survey pretests can be used to ascertain whether coding of mother's person number is cost-effective. ☐

ACKNOWLEDGMENTS

We are grateful to the Indonesian Central Bureau of Statistics and to the Government of American Samoa for permission to use census data for this analysis. This research was supported in part by the Office of Population of the United States Agency for International Development.

REFERENCES


CENSUS CONFERENCE (continued)

uses of census data often cannot be anticipated. He pointed out the need for flexibility in data availability and management to make data collected in one period usable in the future.

Sampling

Answers to some of the questions about how to collect the additional data wanted by policy-makers surfaced in the session on sampling. Ironically, some of the participants shifted their stances and asked why the census should be conducted if the full-count data were not going to be used. One of the papers presented gave an account of the use of sampling in conjunction with India's 1980 census, and the other offered suggestions on the use of sampling for extensive tabulations and complex analysis.

P. Padmanabha described India's use of sampling in conjunction with the recent census. Twenty percent of the enumeration districts were selected for sampling. These districts were asked questions on migration and fertility that were not included on the full-count questionnaire. Use of the enumeration district as the areal unit was chosen in part to avoid training all enumerators in the more complex task of conducting the sample survey as would have been necessary with a sample based on a random selection of households. At the tabulation stage, an additional 20 percent of the enumeration districts was selected to make a combined sample of 40 percent for the presentation of economic data (which had been collected from all households). One quarter of the 20 percent fertility and migration sample was selected for early tabulation and presentation of national-level estimates.

Jay-Soo Park told the conference that savings from using samples for tabulation permit more tabulations and more complex tabulations. He argued that many statistics derived from census data need not be calculated from the total census count. The only statistics that must come from the full count are those that require much detail either at the small-area level or by size categories. Full-count data can be used to check sample statistics by comparing variables available in both data sets that are likely to be correlated with the desired statistics. Samples can be selected from the data collected in the census or additional data can be collected during the census from selected households.

Park advocated using special questionnaires at randomly selected households to collect additional data during the cen-
sus. Although, as P. Padmanabha pointed out, this requires additional enumerator training, it avoids the intraclass correlation associated with sampling clusters of units such as enumeration districts. The intraclass-correlation problem is particularly acute when data are being collected on migration or such characteristics as ethnic group, language, and religion that tend to be concentrated in areal subdivisions.

A. K. M. Ghulam Rabbani said that collecting data that would not be used raised a serious ethical question for census takers, and Teik Huat Khoo mentioned that, in Malaysia, the idea of sampling had prompted the outcry, "If you don't count everybody, you can't be right." Participants seemed to agree that if sampling is to be used extensively, the 100 percent questionnaire should be kept to the basic items that will be tabulated on a full-count basis. These data become the controls for sample tabulations. Frank Parsons suggested that sampling might meet with greater acceptance if users are educated to the fact that the error component as a result of sampling is sometimes lower than the error component as a result of the response and processing errors of the full count.

Rabbani noted that extensive sampling is not feasible where tabulations are done manually. In Bangladesh, retrieving and re-storing records from the tons of census documents is too much of a burden for the census operation.

Mapping

Although many of the countries at the conference told of problems with their mapping operations, two of those present appeared well on their way to solving their problems. Japan's grid system of statistical mapping, which has been the envy of many, has been in use for the last two censuses. Korea, on the other hand, is just now installing its computer cartographic system.

Shoichi Ueda described the grid system used on Japan's statistical maps. It is designed to effectively present small-area information to administrators, and in that sense is a form of the linkage discussed in the development and population policy session. Japan's grid system depends on high-technology computers for large-scale storage of data and for computer mapping. Data from other sources (e.g., land use data and location of public facilities) are combined with the census data to form the statistical maps.

The grid system divides the country into areas approximately 1 kilometer square. The squares are bounded by one hundredth degree of longitude and latitude. Census data are coded for location within these squares, which is fairly simple as long as enumeration districts are much smaller than the grid squares. As enumeration districts become larger, they must be assigned to a grid square, or, for very large districts, the data must be divided among grid squares. For the 1975 census, data were apportioned to grid squares according to the location of the household, but this proved an expensive and time-consuming operation which was dropped from the 1980 census. The system provides accurate statistics at very small areas, though not as small as the grid squares where data are assigned to specific squares according to where they approximately belong.

Barry Wayne Luscombe, a cartographer with the World Bank, described the geographic information and automated mapping system being set up in the Republic of Korea. One of the reasons for installing the system was to make possible the more effective presentation of data. The geographic information system must associate locational data with nonlocational data. Locational data are digital representations of boundaries of administrative units, and nonlocational data are the census variables. Boundaries of administrative units are digitized from a set of provincial maps. A digitizing station extracts the planometric location of a particular point on the map and converts this to a world coordinate system, the Universal Transverse Mercator system in Korea's case. In the mid 1970s, the National Bureau of Statistics began geocoding its data, that is, attaching a unique geographic code to the data elements. The computer uses these codes to place graphic representations of the data on maps.

The data in Korea's geographic information system can be output on an inexpensive printer for low-resolution maps for rough analytic use or on a high-quality mapping system that allows data manipulation. The more sophisticated system allows the creation of new variables from existing variables, interaction of maps and their components at the terminal, "windowing in" on areas to look at them in more detail, output on vector plotters or storage-tube devices, or saving maps on a storage medium for later recall. Such a system has a wide variety of graphics available, including various forms of shading, a variety of typefaces, a wide range of symbols, proportional symbols, and direct color output.

Migration

As mentioned in the development and population policy session, one of the areas where better data are needed is migration. Frank Parsons said that Australia's experience indicates that the census offers only a limited opportunity to collect data on migration. To monitor net migration it is important to identify nonmovers, movers into each area by origin, and movers out of each area by destination. This is about all the data the census schedule can provide without overburdening the respondents. Surveys can be used to supplement the census data with information on reasons for moving and length at current usual residence, but comparability must be maintained between censuses and between surveys and censuses. Because of the diverse uses to which migration data are put, Australia will release detailed internal migration tabulations as multidimensional matrices on magnetic tape for user manipulation.

The ESCAP Comparative Study presented by Badr Hanna is an example of how surveys can make up for the limitations of the census. The ESCAP study examines migration patterns in Indonesia, Malaysia, Pakistan, the Philippines, the Republic of Korea, Sri Lanka, and Thailand, focusing on the various forms of population movements, how they have changed over time, and what functions they perform for the individual and for the communities of origin and destination. Census data have been used as the base information for trends of urbanization, patterns of migration, and levels of urban and rural development. But there are problems with the size of geographical areas influencing the volume of migration captured and with return and multiple migration being omitted in census registration. Furthermore, where urban-to-rural appear more impor-
tant than rural-to-urban flows (e.g., in Malaysian and Philippine census data), closer analysis is required to determine if this is due to an enumeration problem and misclassification of information, urban boundaries being too limited so that urban overspill is registered as urban-to-rural movement, or a real reversal of migration.

The ESCAP surveys attempt a clearer specification of the conditions under which different types of population movement occur so that their links to the development process both as cause and as effect can be assessed. They compare different phases of development with different types of policies. With the data that will be generated from these surveys, policymakers will be able to assess and plan for the processes of urbanization and migration in their respective countries.

Fertility and Mortality Estimation and Labor Force Data Collection

Estimation of fertility and mortality has been the meat and potatoes of census data analysis. The presentations on fertility and mortality showed how Griffith Feeney's warning, "Never trust a single set of estimates," can be followed in practice. Indonesia does not have a vital registration system, explained Sam Suharto, so fertility estimates must be based on census and survey data. Unfortunately, these data are not of high quality: age reporting is not very accurate, and reports of children ever born are not reliable. Consequently, Indonesia's Bureau of Statistics plans to use five methods to estimate fertility and to cross-check the results. According to Wiwit Siripak, mortality estimates for Thailand are based on three sources—the 1970 and 1980 population and housing censuses, two surveys of population change, and vital registration statistics—which are used for cross-checking. Feeney himself presented a model of population dynamics based on birth intervals and parity progression—"a more natural way to study fertility," according to Feeney.

Increasingly, census operations are being asked to collect data on the labor force. According to Teik Huat Khoo, Malaysia's approach has been to gather a relatively wide spectrum of labor force statistics, including information on the employed, the unemployed, and the outside labor force. But there is a need for frequent updates of the data bases, so the census results are seldom used without adjustment.

Editing Systems

Benjamin Mok told the conference that for Hong Kong's 1981 census editing was done in two stages: validation and imputation. Validation was restricted to basic checks. Its objectives were to ensure that the structure of the data record was correct, that no record had been duplicated or omitted during data input, that all codes were within the specified ranges, that the household records within the living quarter or person records within a household were consistent in relation to the selected variables, and that the key variables (age and type of household), which would be used as the basis for imputation, were correct. Errors identified by the computer were displayed in validation reports, which were used by approximately 30 clerks who checked the original questionnaires and determined the correction required. For variables other than key ones, amendment could be deferred to the imputation stage by using

a by-pass key which caused the violated checks to be passed over in the ensuing validation cycles.

Imputation was employed to correct errors in consistency between the different characteristics in a record. Both hot-deck and cold-deck methods were used. The program included detailed checks of up to four variables and automatically identified inconsistencies and rectified them. The imputation logic involved a sequence of checks in which items were checked successively with previous accepted items. The first item in the sequence was the key variable, which was established to be correct in the validation stage. Inconsistencies were rectified by amending the lower item in the hierarchy. The maximum imputation rate was about 5 percent.

Griffith Feeney, once the theoretical side of editing he defined the two components of an edit as the edit test or check and the edit action. An edit procedure is a sequence of edits. Feeney noted that rearranging the order in which the edits are executed can lead to different results. Three types of edit actions are: report only without doing anything to the records; check the information against what went into the preceding process to see if the error was introduced by the process and, if so, correct the information, adjust the process, or both; and replace the bad information with the best guess available at this stage and proceed. The latter two edit actions are backchecking and imputation.

Imputation involves changing the record in such a way that, if it were immediately subjected to the edit test again, it would pass. But this does not guarantee that the information would pass the entire edit procedure again. If an earlier edit in the sequence involves the variable whose value was imputed, it is possible that the new record would fail that edit. This is a serious problem because one edit may undo the work of a previous edit and introduce or suppress errors.

Feeney suggested three approaches to avoiding this situation. If priorities are assigned to the variables, the edit procedure can be designed so that once a particular variable has been given an imputed value, no subsequent edit can change this variable. A second approach is to design the system so that whenever there is an imputation, the record is tested against all the edits involving the variable whose value was imputed. The third approach involves logical reduction. If two edit tests involve the same variable, they are linked by a third test that, for example, will allow the record to pass if it passes one, the other, or both tests, but will not allow the record to pass if it fails both tests.

Census Evaluation

A. K. M. Ghulam Rabbani gave the following account of the Bangladesh postenumeration check (PEC). A field check of selected enumeration areas was conducted two and a half weeks after the census—a period that was long enough to get the census records out of the field offices (so they could not be copied) but short enough so that respondents could still remember who was staying in their household on census night. To further ensure the independence of the PEC, the location of the enumeration areas to be re-enumerated by the PEC was kept secret from the census field staff. In turn, the PEC enumerators, who were specially recruited employees of the Bu-
As mentioned earlier, the issue of adjusting census counts after evaluation, was a major problem for the U.S. Census Bureau. Zitter noted that one of the problems in this regard is that the adjustments need to be made at levels of geography too small for direct measurement of enumeration error. He said a method for extending the measurement of undercount to small areas is needed, which means the variables associated with undercount must be identified.

Beyond the Eighth Census Conference

The closing session of the conference was largely devoted to making plans for the Ninth Population Census Conference to be held in 1983. By that time the census offices will have completed most of their tabulations and their census evaluations. Therefore, the conference will focus on experiences with those operations. Questions and problems that were raised at the eighth conference led to the suggestion of the following topics for inclusion on the agenda of the ninth: the census data base in relation to other government data; linking censuses with other surveys; the dissemination of census data (what needs to be published, for whom, and in what form?); and measurement techniques for development indicators, labor utilization, and migration. Also suggested were conference sessions on strategies for mid-decade sample surveys to cover the deficiencies of decennial censuses and developments in data processing to include reports of user experiences with software packages (especially CONCOR) and the hardware being used.

Out of the Eighth Population Census Conference will come a volume on the experiences of the participating countries with their 1980–81 censuses. Like Lee-Jay Cho's earlier volume, Introduction to Censuses of Asia and the Pacific: 1970–74, Censuses of Asia and the Pacific: 1980–81 will present the design, organization, and execution of the most recent censuses in a form that facilitates comparisons between countries and will include samples of census maps and questionnaires.

THE ASIAN AND PACIFIC CENSUS FORUM is a quarterly publication of the East-West Population Institute, supported by a contract between the Institute and the Office of Population, Agency for International Development. It is available without charge to governmental agencies, private institutions, and interested individuals. News items and comments are welcomed and should be addressed to: Census Forum Managing Editor, East-West Population Institute, 1777 East-West Road, Honolulu, Hawaii 96848.

THE EAST-WEST CENTER is a national educational institution established in Hawaii by the U.S. Congress in 1960 to promote better relations and understanding between the United States and the nations of Asia and the Pacific through cooperative study, training, and research. Each year more than 1,500 men and women from many nations and cultures work together in problem-oriented institutes or on "open" grants as they seek solutions to problems of mutual consequence to East and West. For each center participant from the United States, two participants are sought from the Asian and Pacific area. The U.S. Congress provides basic funding for programs and a variety of awards; and the Center is administered by a public, nonprofit corporation with an international Board of Governors.

President, East-West Center: Victor Hao Li
Acting Director, East-West Population Institute: Keith F. Adamson
Forum Technical Editor: Robert D. Retherford
Forum Managing Editor: Robert L. Hearns
Production: Lois M. Bender
Graphics: C. Kamehiro