PN-AAP-925/62
ISN: 34430

WHY COLLECT RETROSPECTIVE DATA?

William P. Butz

December 1981

P-6715

WHY COLLECT RETROSPECTIVE DATA?[1]

William P. Butz[2]

Almost all survey data are recall data: "What foods did your family eat yesterday?" "What were your expenditures last week?" "What was your income in the last four months?" "How many pregnancies did you have in the last three years?" Retrospective data can be defined, only a little facetiously, as recall data with a reference period that is too long. For example, "What foods did your family eat a month ago?" "What were your expenditures six months ago?" "What was your income five years ago?" How many pregnancies have you had since age 15?" Most people, including professional surveyors, would say that respondents cannot give accurate answers to these latter questions, and that the inaccuracies increase the farther back in time the questions try to reach.

Nevertheless, some researchers do gather "retrospective" data, most notably in the fertility area but in others as well. Why do they bother? What would one do with retrospective data? The simplest answer to these questions is the answer to the separate question: What would one do with panel data? The answers are similar, because the form of the data is in both cases nearly identical.

---

[1] William P. McGreevey asked me to speak on this subject at a World Bank Conference on the Measurement of Living Standards, held at Belmont in October 1981. AID, through Grant No. AID/otr-1744, funded the subsequent translation of notes to prose. Previously, the Rockefeller Foundation and AID supported my personal run-ins with retrospective data. I am grateful to all these, and especially to my surveying colleagues from Guatemala, Malaysia, and the U.S., from whom any good ideas herein most likely came. Julie DaVanzo, Dennis DeTray, and especially John Haaga offered useful comments on the initial draft.

[2] Economics Department and Labor and Population Studies Program, the Rand Corporation.

Panel, or longitudinal, data describe characteristics or behavior of persons over a period of time, by observing or questioning them at points throughout that time period. Such data are similar to retrospective data in three important respects:

1. Both types of data document characteristics or behavior of the same individuals or families over time. The result is person- and family-specific histories.

2. Both types can document these characteristics as of arbitrarily specified points, for example, at particular ages of respondents or particular calendar dates. Such data might record data as of respondents' 20th birthday, 25th, 30th, and so on, or as of July 1 in each year. Both types of data can also document <u>changes</u> in characteristics or behaviors, regardless of age or year. In this mode, the data might record employment status at age 18, thereafter noting only the dates of changes and the new statuses; or age of first marriage, and thereafter only changes in marital status and the associated dates.

3. Both panel data and retrospective data contain more information than researchers yet know how to take advantage of, either conceptually or statistically. Though important methodological advances are being made, most behavioral models and the statistical methods developed to estimate them were designed to be used with cross-section data. Panel and retrospective data add to cross-section data a time-series dimension which opens the way to analysis of dynamic processes.

Alongside these similarities of form and complexity, panel and retrospective data are also different in seven important ways:

1. As of the date survey resources are committed, retrospective data document the past while panel data document the future. If the past is of interest, then a panel survey is not an alternative. If the future is of interest, it can be documented either prospectively in a panel survey or retrospectively at the end of the study period.

2. Panel and retrospective data differ in the information available on attrited respondents. Both types of data suffer attrition of respondents in the same way and for the same reasons: death, and frequently migration out of the area. Hence, both types of data can have the same sample at the end of the reference period. However, panel data include partial information on the attrited respondents, namely information collected before the death or migration. Retrospective surveys cannot question these individuals because they are not available at the end of the reference period when the retrospective survey is administered. Hence, the biases from resondent self-selection out of the sample can be directly studied in panel data, but not in retrospective data. Instead, retrospective surveyors must turn to other data sources for estimates of mortality and out-migration rates in the past, or ask the respondents for basic information on family members who have died or outmigrated.

3. These two type of data differ in their possibilities for sample selection criteria. Respondents in a panel survey can be selected only on the basis of their characteristics at the <u>beginning</u> of the reference period. Respondents in a retrospective survey, on the other hand, can be selected on the basis of either initial or terminal characteristics, on the basis, in fact, of characteristics at any intermediate age or date. This is accomplished by documenting the relevant characteristics

in a short screening interview, then picking the sample according to the desired criteria. For example, in a study of rural-to-urban migration, a retrospective survey can include the desired proportions of migrants of various types. But in a panel survey it is not possible to select only persons who will migrate during the period to be studied, or indeed, to be sure than any in the selected sample will migrate. Because of this uncertainty, a larger sample must be selected for the panel than will usually turn out to be necessary. Worse, the sample size will sometimes prove insufficient. This same difference between the two types of surveys applies whenever data are to be used to study characteristics of the transition to a particular condition after the beginning of the study period. Divorce, unemployment, contraceptive use, morbidity, and malnutrition are such conditions whose future frequency in a population cannot be precisely forecast. This advantage of retrospective data becomes more important the less frequent in the population are the events to be studied. On the other hand, this advantage is irrelevant if the sample is to be completely random.

4. These two types of data differ also in quality, as indicated by various measures of validity and reliability. To the extent that quality declines as the recall period lengthens, and it does,[3] panel data are higher quality. However, panel data are subject to "reactive effects," whereby respondents change their behavior as a result of being (continually) interviewed. The changed behavior is then recorded in future interviews. In addition, other factors such as coding

[3] There is little theoretical understanding or systematic empirical evidence concerning the rate of quality deterioration with longer recall periods. See John Haaga, Validity and Reliability of Retrospective Life History Data from the Malaysian Family Life Survey, The Rand Corporation, N-1823-AID, 1982, Appendix.

procedures, and interviewer quality, training, and control also systematically influence data quality. In general, one would expect panel data to be of higher quality than retrospective data, but the advantage could certainly shift in any particular comparison.

5. The interview process is significantly more complex for retrospective data. Collecting them therefore requires higher quality interviewers, more training and standardization, and stricter field control. Data coding is also significantly more difficult.

6. Apart from the interview and coding processes, everything else is much more difficult, costly, and uncertain for panel data. Maintaining a field staff and keeping track of respondents over a period of years is obviously harder than conducting a single survey. More fundamentally, the prospect of being able to finish a panel survey ten years down the road must certainly be judged as uncertain in most underdeveloped countries. Some panel studies have been stopped in midstream, while the investigators in others have restricted their initial investments due to uncertainties about the study's term.

7. Panel and retrospective data differ greatly in the waiting time to use the data. Panel data describing a study period will not be available until the period is over, plus time for final coding, punching, and cleaning. A retrospective survey, on the other hand, requires only about the same time as any cross-sectional survey. It may take somewhat longer because of more difficult surveying and coding procedures, but months at the most, instead of the years a panel survey requires.

In most situations, the principal disadvantage of retrospective data is considered to be their poor quality. The main disadvantages of panel data are their high cost and the long wait before they are available. Similar comparisons can be made between retrospective data and other kinds of information that could be used for the same purposes, for example, past censuses, past cross-sectional surveys, and vital statistics from registration systems. Retrospective data will have the clearest advantages in these comparisons when the other data do not exist or are of poor quality.

These brief comparisons suggest that retrospective surveys are neither better nor worse in all situations than alternative sources of information. Rather, the optimal choice of survey type depends on the other data already available, the survey setting, the available resources, and the uses to which the data will be put.

In practice, though, retrospective surveys have a bad name. In the 1940s and 1950s and even earlier, many retrospective expenditure surveys were conducted in Europe and some less developed countries. Then in the 1950s, a series of retrospective pregnancy surveys occurred in developing countries, primarily in Africa. Subsequent analyses, some quite recent, demonstrated that the expenditure surveys suffered substantial underreporting when the reference period was longer than a week. The pregnancy histories also appeared to miss information in the more distant past--in this case, pregnancies that did not result in a live birth and births of children who later died. More recent conjecture suggests that respondents in retrospective pregnancy surveys may inadvertently report some of their earliest and most recent births as having occurred instead in their middle reproductive years. If so,

fertility rates aggregated from such surveys could show a recent decline in period fertility rates when none, in fact, occurred. In addition, some of these retrospective surveys implied rising mortality over time in places where it was known from other sources to be declining.

In spite of this evidence, many cultural anthropologists continue to rely on retrospectively reported accounts, and the method is increasingly used by historians. Most other researchers, however, rarely attempt today to collect retrospective data beyond a record of a woman's pregnancies and directly related events. Only a few investigators have been bolder. For example, the 1971 Monterrey Mobility Study documented migration histories. The 1971 Social Accounts Study at Johns Hopkins University included a much broader retrospective life history. And the Adolescent Society Follow-up study at Battelle Memorial Institute in the mid-1970s collected considerable retrospective information. Then in 1975, the Institute for Nutrition in Central America and Panama (INCAP) in Guatemala, in collaboration with the Rand Corporation, attempted a retrospective survey considerably broader than its predecessors. Along with pregnancies and the births and deaths of children, female respondents were asked for a history of lengths of breastfeeding, postpartum amenorrhea, and contraceptive use; types of supplementary infant food used; cohabitation, marriage, separation, and divorce; types of roof, walls, and floors of all houses, as well as ownership status, rent (if applicable), and number of rooms; schooling and training; and hours, wages, and occupations of all jobs. A complementary instrument was administered to the spouses of these women. Two years later in collaboration with Malaysian institutions, Rand included similar retrospective instruments in a survey in Peninsular

Malaysia. The surveyors designed the instruments and interview procedures for both surveys to facilitate cross-referencing across different areas of the respondents' life history. It was hoped that this would jog respondents' memories and improve data reliability.

John Haaga has subjected these retrospective data from the Malaysian Family Life Survey to unusually thorough tests of reliability and validity.[4] His tests include checks against external sources of information, checks of internal consistency, and checks of reinterview reliability. Haaga examined the data on breastfeeding, contraceptive use, fertility, postpartum amenorrhea, fetal and infant mortality, birthweight, respondents' educational attainment, and housing amenities. The external data used for comparison are from censuses or surveys done in 1957, 1967, 1970, and 1974, and from the continuing vital registration system. His conclusions are that the retrospective data on the number and timing of births, the number, timing, and age at death for infant deaths, contraceptive use, birthweight, educational attainment, and housing amenities all appear reasonable. Information on fetal mortality--abortions and miscarriages--is seriously underreported; but this is a general problem in all data, no matter how short the recall period.

Finally, the data on lengths of breastfeeding and amenorrhea are badly "peaked." Women report these lengths at 6, 12, 18, and 24 months far more frequently than these lengths could conceivably occur. Haaga warns that this pattern, which is similar to the phenomenon of "age heaping" in censuses, makes these breastfeeding and amenorrhea data highly suspect for multivariate analysis.

---

[4] John Haaga, op. cit.

Yet, when we investigated the statistical relationship between lengths of breastfeeding and amenorrhea in these data, we found patterns that closely replicate the findings from clinical studies using prospective, observational, data. Our results do not even change importantly when only observations that occurred before 1960 (more than 15 years before the interview) are considered; nor when all the observations with peaked values are dropped from the sample.

The Malaysian setting has peculiarities that might make these data better than retrospective data from most other developing countries. (For example, schooling and income are higher, and an identity card system has recorded most births and deaths since 1953.) Nevertheless, the proof of the data, this example suggests, is in the using. What questions will be put to the data? How, statistically, will the questions be asked? How imprecise can the answers be and still be worthwhile?

Many demographers and survey statisticians preach that any data likely to be lacking high validity and reliability should not be gathered. What damage this has done! How many data sets lack a measure of income, of morbidity, of women's hourly earnings, which, if there, could extend their use to new questions by investigators in different disciplines? And how many policy questions are awaiting "definitive" answers from panel data when retrospective surveys could at least identify the principal possibilities. Two recent analyses of retrospective data in the Malaysian Family Life Survey illustrate this point. In one, James P. Smith has conducted a Dennison-type growth accounting exercise of increases in personal income since the Second World War. Without using retrospective survey data, such an exercise

simply could not be done for Malaysia. The results are sensible and consistent, and they reveal important characteristics of the Malaysian development process. In the other, Sidney and Alice Goldstein have analyzed the relationships between migration and fertility. Being able to locate events temporally in respondents' life cycles has yielded important inferences about causality.

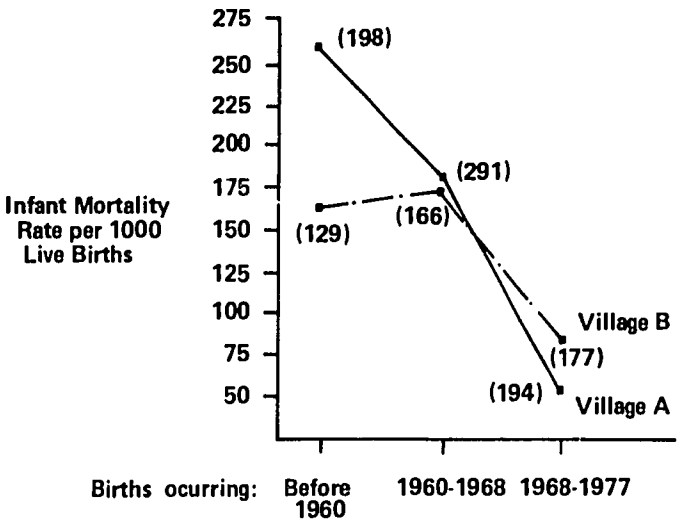
The doctrine of validity and reliability for their own sake should yield to a broader calculus. In it, we value these characteristics only for their contribution to the project's ultimate objectives. We recognize that the opportunity cost of boosting reliability coefficients is assuredly a longer survey or more analysis. We don't let data quality go to pot, but we choose objectively how much of it to pay for.

I suspect that such a calculus will recommend a retrospective survey in several circumstances:

Wherever prospective panel data would be useful, consider whether retrospective data would be useful enough. They are cheaper and faster. Can they answer the questions?

Wherever baseline data are required, consider whether information on the baseline trend would also be useful. It is usual practice in field interventions to ascertain through baseline surveys the initial comparability of persons, families, or villages with respect to the evaluation indicators. For example, the figure shows data on infant mortality rates for two villages chosen to be comparable in 1968, for a nutrition intervention that began in Village A in 1969.[5] Comparing the

[5] The figure and accompanying discussion are from Jean-Pierre Habicht and William P. Butz, "Measurement of Health and Nutrition Effects of Large-scale Nutrition Intervention Projects," in Robert E. Klein, et al. (eds.), Evaluating the Impact of Nutrition and Health Programs, New York, Plenum Publishing Company, 1979, pp. 150, 151. An



Infant mortality rate trends in two villages chosen for a nutrition intervention. Source: Female Retrospective Life History Questionnaire from *INCAP-RAND Survey in Guatemala.*

Note : Numbers of live births in each period are in parentheses.

rates in 1960-1968 suggests that the two villages were quite similar with respect to this indicator. However, looking at the trends between the 1950s and 1968 in these villages sorely tries one's confidence in the comparability of future infant mortality data across the villages. Such confidence is, of course, a prerequisite for believing that the reduced mortality seen in Village A relative to Village B after 1968 is due to an intervention applied to Village A and not to Village B.

These data are primarily from a retrospective survey conducted in 1975. Had such questions been asked in 1968, villages would presumably have been differently chosen or matched.

More generally, wherever information about the past is needed, but otherwise missing or deficient, consider that it is still possible to get it. A retrospective survey is always a substitute for a vital registration system, a history of censuses, or several past cross-sectional surveys. "How good a substitute, considering the

---

accompanying health intervention occurred in both villages, causing infant mortality rates in both to fall faster than previously.

purpose?" is the question. "Far from perfect, but inexpensive and good enough," may be the answer.