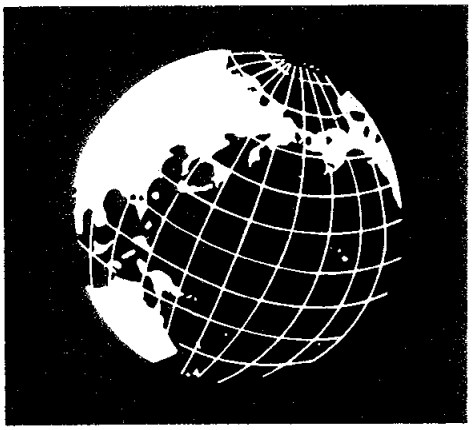
East-West Population Institute
August 1983



East-West Center Honolulu, Hawaii
Volume 10 Number 1
ISSN 0732-0515

Asian and Pacific CENSUS FORUM

A Microprocessor Revolution in Data Collection?

by Griffith Feeney

Economic history teaches that innovation begets bottlenecks, which in turn beget more innovation. Innovating and becoming more productive in one area, we begin to feel handicapped by relatively low productivity in other areas, creating pressure for still more innovation. This counterpoint of innovation, challenge, and response has played an important role in economic development. The manufacture of cotton textiles in England during the Industrial Revolution illustrates the dynamic at work, as does the recent history of electronic computers.

The bottleneck in census and survey data production today is the process of transforming the information on the questionnaires that come from the field into "clean," computer-medium records for further processing. This includes coding, as of occupation or industry; data entry, especially when keystroke devices rather than optical-mark readers are used; and manual editing, including manual editing carried out after data entry with the assistance of computer-generated error reports.

Field operations are completed in a matter of weeks or months. Computer tabulation and offset printing can be accomplished in a comparable period. Yet it is not uncommon to have to wait one year for the results of a survey and two or more years for the majority of census reports. Much of this time is accounted for by coding, editing, and data entry operations.

Is it possible that microprocessors, the devices at the heart of computers, hand calculators, and many other information-processing devices, might relieve this bottleneck? Imagine that the year is 1998. We are about to conduct a household survey. There are no questionnaires, however. Each enumerator carries an "enumerator device" with a keyboard for entering information, a display, and a memory.

To begin an interview, the enumerator presses a button marked "Begin." The device responds by displaying the first question. The enumerator reads the question and enters the respondent's answer at the keyboard. Where coding is required, as in industry or occupation questions, the codes are determined interactively, on the spot, by reference to standards stored in the memory of the device. The enumerator then presses a button labeled "Next," and the next question is displayed.

As the answer to each question is entered, it is compared with the answers already supplied, using edit checks stored in memory. When anomalies or inconsistencies are detected, a query is displayed and the enumerator probes the respondent to confirm or correct the response. When the response to the last question has been entered, the device resets itself and is ready to begin the next interview. The information that has been entered is complete, insofar as the respondent is capable of supplying it, and it is fully coded and edited. The programming of the device makes it impossible to record inconsistent data.

At the end of a day of interviewing, the enumerator goes to the local statistical office, which receives the information collected and transmits it to the central statistical office. When the fieldwork is completed, all the information collected is fully edited and in computer medium, ready for further computer processing. The entire lengthy process of coding, editing, and data entry has been eliminated—or, rather, it has been subsumed into the data collection operation. Final survey results are available within a week of the end of field operations.

The enumerator device envisioned in this scenario is easily within reach of current technology. Indeed, something similar

(continued on page 13)

Griffith Feeney is a Research Associate at the East-West Population Institute.

choice of cutoff points of cohort and birth order for the indices and the best ways of combining the measures to exhibit significant overall quantum trends. Since sample errors for the individual values are considerable, effectively combined indices should be more reliable and consistent. □

REFERENCES

Juarez, F. 1983. Family formation in Mexico: a study based on maternity histories from a retrospective fertility survey. Unpublished Ph.D. dissertation, University of London, London, United Kingdom.

Page, H. J., and Lesthaeghe, R. 1981. *Child Spacing in Tropical Africa: Traditions and Change.* London: Academic Press.

Rodriguez, G., and Hobcraft, J. N. 1980. *Illustrative Analysis: Life Table Analysis of Birth Intervals in Colombia.* World Fertility Survey Scientific Report No. 16. London: World Fertility Survey, International Statistical Institute.

Ryder, N. 1980. Comments at the Seminar on the Analysis of Maternity Histories, International Union for the Scientific Study of Population, London.

MICRO REVOLUTION *(continued from page 1)*

to the process described occurs already in telephone surveys. Interviewers work at a video display terminal in much the way our hypothetical enumerators work their device in the field. The technological challenge lies in packaging such a system for use by enumerators in the field, and at a cost that is not prohibitive.

Prototypes are already in operation. The Hawaiian Electric Company recently equipped its meter readers with hand-held devices to record the readings from electric meters. In the past, the readings were recorded on optical-mark sensing forms in preparation for computer processing of consumers' electric bills. These meter reader's devices have editing capability. The current meter reading is checked against previous readings, stored in memory, and the meter reader is alerted of any sharp departures from past experience.

What about economic feasibility? Such devices, produced now, would be prohibitively expensive. Nonetheless, it would be imprudent to conclude too quickly that economics will rule out the development of such devices. For the past 20 years, the cost of information processing has been declining at an extraordinary rate. Today's programmable calculators are far more powerful than the large computers of 20 years ago, and they cost one hundredth or one thousandth as much. Another 20 years of comparable development would put enumerator devices within easy reach. Will it happen? We cannot know. We can only say that it may, and that the possibility is worthy of note.

The extent of the market is an important consideration here, for high production volume is necessary to reduce costs. The potential market here is evidently large indeed. Virtually every government in the world is more or less continually engaged in collecting information. Every large business does likewise, though the context and the methods differ. Since many of these data are collected from individuals and households, population size provides a crude index of the potential market.

How much difference would the development of such devices make? Two principal effects can be clearly foreseen, an increase in data quality and an increase in the speed with which data could be produced. Data quality would improve as a result of performing coding operations and edit checks in the field, as the interview was conducted, so that respondents could be probed to confirm anomalous information and resolve inconsistencies. The more striking effect would be the speed with which data could be produced. We could quite conceivably conduct a large survey, or even a census of population, during one month and have the results in hand the next month.

These effects, desirable as they undoubtedly are, must of course be obtained at a reasonable cost. They might in fact reduce costs, making cost another argument in their favor.

The excitement and glamor of high technology sometimes have the lamentable effect of causing a temporary loss of perspective. The principal value of this reflection may lie in calling attention to imbalances in the various stages of the process of data production, and the necessity of paying special attention to the management of the weakest links in the chain. Whether or not there is a microprocessor revolution in data collection, the human resources and the managerial tasks will remain the most critical aspects of the job. Whatever the future may hold, this is a lesson that may be taken to heart in the present. □

NOTES AND ACKNOWLEDGMENTS

On innovation and response in economic history see David S. Landes, *The Unbound Prometheus* (Cambridge University Press, 1969), especially Chapter 2. Thanks to Professor Philip Hauser for valuable comments on an earlier draft, and to Doug Carlson and Alvin Amerino of the Hawaiian Electric Company. Mr. Carlson arranged for me a tour of the company's meter reading section, which Mr. Amerino conducted, answering a great many questions in the process.

ENUMERATIONS *(continued from page 4)*

migration and their impact on the growth of urban Tarawa, and prepared a report for the government on contract labor, equity, and migration in Kiribati.

Connell also visited Papua New Guinea where he worked with the Department of Labor and the National Statistical Office on issues relating to employment and migration, especially in relation to the analysis of data from the last census.

Since June, SPC Demographer Ko Groenewegen has been on a duty tour to several Pacific Island countries. On the first leg of the tour he travelled to Nauru at the request of the government there to assist in the follow-up of the population census conducted in May. Discussions were held on data quality control and manual checking procedures, as well as coding instructions and electronic data processing. In July he discussed the arrangements for the analysis of the results of

(continued on page 16)