

**International Union for the Scientific Study of Population
Regional Population Conference, Bangkok
10-13, June 2002**

Perspectives on the Censuses of the New Millennium

**By
K.E.Vaidyanathan**

**Contributed Paper for Session 26 on
Data Collection and Dissemination Systems**

PERSPECTIVES ON CENSUSES OF THE NEW MILLENNIUM

By
K.E. VAIDYANATHAN
vaidyake@hotmail.com

I Why Population Censuses are Essential

In the large majority of the countries there is no alternative to the population census as the means for collecting the most essential socioeconomic data, for establishing the frame for conducting the surveys and for fulfilling the administrative needs such as delimiting areas for elections. Countries like Netherlands where population registration system is efficient, a population system may appear redundant, but in the absence of the population census there is no basis for determining whether the population census system is accurate. It is also doubtful if the range of questions asked in the population census can be included in the population registration system. In countries like Bosnia and Herzegovina and Somalia where large-scale population movements have taken place as a result of war there is an urgent need for fresh census. Since no alternative source of data are available the previous census data are being used with adjustments in spite of these data are becoming obsolete. For instance, in Bosnia and Herzegovina the sample for LSMS survey is based on the pre-war 1991 census results with updating of selected municipalities. In Somalia the author using the pre-civil war census and survey data with various assumptions concerning fertility, mortality and migration during the war and post-war period derived the current estimations of population.

Yet another application of the census is a “master sample” which is a large number of clusters, usually several hundred census enumeration areas which are selected by the statistical agency after the census and for which updated household listings are maintained in order to select from the pool of households the sample of households to be interviewed in each survey. The advantage of the master sample is that there is no need to select a new set of clusters for each survey or to conduct household listing operation each time. Thirdly, cost of updating the master sample will be minimal if interviewers can be selected from the same area. Where censuses are taken every 10 years there is still necessity to conduct sample surveys to obtain the information on current demographic and social statistics. Indeed this is what every developing county is doing. India, for example, has the annual rounds of national sample surveys (NSS) to obtain current data and to compliment the information obtained from the decennial census. Conducting such surveys is far easier where there are information available from a census and where there is staff with experience in conducting a census.

Administrative records are important sources of information of vital events, but any analysis of the data from these records require complimentary information on the “population at risk” which can be provided only by a census. In many developing countries administrative records are incomplete and inaccurate. For a large majority of

developing countries the estimates of vital rates, nuptiality and reproduction during the last decade were based on indirect estimation from census and survey data, and hardly from vital records. Moreover the information provided by the administrative records are not necessarily the same as those obtained from the census. Administrative records can serve different purpose, namely, to monitor the performance of the provider of the administration, and only incidentally they provide statistical information.

II. Benefits to stakeholders from a Census

The potential stakeholders for a census are statistical organizations who carry out the census or use the census data, decision makers in the government, non-governmental organizations as well as international organizations. All these stakeholders benefit from a census in different ways.

1. Statistical organizations in the country will benefit from an expansion of staff and equipment arising from a population census. Significant numbers of staff will be able to upgrade their skills through theoretical and hands-on training. The end result of a census is a statistical system that is capable of designing and implementing quality censuses household surveys with little or no outside assistance.

2. Policy and decision makers at the national, state (or province) and local levels require information provided by the census for various purposes. At the national level census data are required as a key input in national accounts, estimation of requirements for different commodities and services and for formulating social programmes and interventions. National governments have been using census data for preparing the Poverty Reduction Strategy Paper (PRSP), which is increasingly becoming a requirement for assistance from the World Bank. At the State and local levels, in addition to the above, the budget, staff and other requirements are often based on census figures of population.

3. Non-governmental organizations within each country are among the greatest users of population censuses. For example, Chambers of Commerce and Industry, Syndicates and Associations of different trades make use of census data for market research and for informing their clientele about potential business opportunities. National associations in the field of social sciences, population, and family planning use the census data for designing, monitoring and evaluation of their activities besides advocacy purposes. Lately, feminist organizations are using census data to buttress their claims for gender equity and justice. For example, in India feminist organizations are using the sex ratio from the census as the basis for questioning the legality of sex-determination. Even political parties use census information about their constituencies in the selection of candidates. Universities, research institutes and individual researches are among the greatest users of census data.

4. More recently international organizations like UNDP, UNFPA and World Bank have emerged as major users of census data and consequently important stakeholders. UNDP's Human Development Report and Human Development Indicators (HDI) draw heavily

upon the data and indicators derived from the population census. UNFPA and has been a major user of census data from monitoring demographic trends and for preparation of the annual State of the Worlds' Population Report. In addition many international organizations have been using census data for their strategy development and project formulation missions and for advocacy purposes. The World Bank has been using the census data for the preparation of documents relating to poverty alleviation besides supporting a number of LSMS Surveys in developing countries, which have relied heavily upon census data for their sample selection.

The extent to which the aspirations and requirements of different stakeholders are met depend upon the resources available, external funding, technological structure (e.g. Internet facilities), etc., and also capacity of the clients to pay for the information. The traditional methods of disseminating the information like hard copies of census publications are fast giving way to supply the information by diskettes CD ROMs, e-mail and even posting of information on the Internet.

III Strategies for Involving Stakeholders

The commitment of the Government to conduct a census is present where there is a constitutional requirement or a strong tradition of conducting a census every ten years (for delimitation of constituencies for elections). This has been a case in India where there has been an unbroken chain of censuses since 1871. Where such constitutional requirement is not present the commitment for carrying out a census is weak unless there is a powerful lobby to press for a census every ten years. This is where different stakeholders can play a major role in ensuring the conduct of a census and in ensuring that censuses meet the needs of different stakeholders. In countries where censuses are taken regularly such as India the different stakeholders are involved in the preparations for the census. For example, in India census data users conferences are held to determine the contents of questionnaires and the procedures to be adopted. In addition non-governmental organizations like as Indian Association for the Study of Population (IASP), Universities and even individual researchers act as pressure groups to influence the selection and development of the topics included in the census questionnaire and the approaches to measurement issues.

In countries where such pressure groups do not exist (such as many African countries) it may be necessary for international organizations and NGOs to step in and act as pressure groups influencing census taking. Many African censuses in the 1960s and 1970s and the census of China in the late 1970s were due to pressures from international organizations and financial support of UNFPA. There is a need to establish scientific societies (or associations) in the field of population in many African countries, and IUSSP should support the formation of such societies through funding in the same way IPPF is supporting several family planning Associations. The support may be towards specific activities such as an Annual Seminar on Population, publication of a journal, or supporting the cost of having a small office. IUSSP should seek assistance from funding agencies like UNFPA, DFID, USAID or private donors to provide necessary support to the national associations so that these associations can act as pressure groups for census

and other population activities in these countries. The support of IUSSP should be made conditional upon the government agreeing to a cost-sharing arrangement and the national associations or societies for population being accorded consultative status with the national and local governments.

IV Adopting New Technologies to Census Operations

The technology of census has come a long way from the stage when manual counting of census slips was done (e.g. in the Indian censuses up to 1961) to the use of the unit record equipment, then the large first generation computers, followed by the use of personal computers and now the laptop computers. The march of technology cannot be stopped and census takers need to keep pace with these changes. These are areas where new technologies could be adopted to census operation in the planning, execution, tabulation and dissemination stages.

(1) Use of Mass Media

Since information pertaining to the census should reach the entire population, media coverage is important. The census administrator should use all available technologies to inform the public about the census and this would include radio, television, internet, e-mails etc. Once the census is completed, census results can be posted on the Internet so that a wider cross section of users can make of the data. Even the data sets can be posted on the Internet besides making them available in CDs. Proper documentation of the census including the questionnaire, method of sample selection, if any, field procedures and quality control techniques, guidelines for using the data, documentation of constructed data sets, description of the data files should be provided. The unit record data should be provided in a user-friendly way in ASCII format or in the format used in various statistical softwares.

(2) Collecting New Types of Information

Most censuses around the world ask certain standard questions relating to population and housing characteristics. The time has come for us to think about asking some new types of questions which can provide insights into specific conditions such as poverty, human rights, reproductive health, disability, gender differences, energy and environment etc. Some years ago Arriaga had suggested that census could be a means for obtaining information on maternal mortality. This needs to be explored further. Many censuses provide data for computing some basic needs indicators such as education, water, sanitation and type of housing. One could explore the possibility of including some questions pertaining to some “perceived” health conditions, nutrition, immunization, access to basic services (attendance at birth, antenatal care), whether the person receives any pension or social assistance (family allowance, child-care allowance, disability allowance etc). Also questions can be asked whether any one has a chronic illness and the nature of the illness, and whether the person has insurance cover etc. In the 1993 Census of Sudan questions were asked about the type of housing, the fuel used for cooking, source of water etc. The list can be expanded to include other environmental characteristics.

(3) Use of Sampling

With advances in methods of sampling there is great scope for use of sampling in conjunction with the census to reduce cost and to produce quick results. In the 1931 Census of India, tabulations of age and other characteristics were made on a sample of census slips (known as Yeatt's sample, named after the Census Commissioner). The 1993 Census of Sudan included a long questionnaire, which covered one-twentieth of the households, and a short questionnaire, which covered all the households. A comparison of the two results showed that the differences between the two sets of results were not great. While sampling may not be efficient for all characteristics, this is an option that deserves consideration where there are other advantages such as reduction in costs, better control of the quality of fieldwork etc.

(4) Questionnaire Layout

Assuming that the conceptual structure of the census questionnaire (including modules such as fertility, labor, or migration) is well thought out, technology can play a great role in proper formatting of the questionnaire. A good format minimizes the error during the interviewing and subsequent data entry, thus improving data quality and timeliness to which the data can be made available. In my view only one questionnaire should be used for each household. In some countries, such as India, the census questionnaire used to have two parts – a household questionnaire and individual slip. Unless proper identity codes are included in all questionnaires separate questionnaires create the risk of improper matching. A grid may be required in cases where there can be more than one unit of analysis in the household- for example several eligible women for fertility module, several employed persons in labor module etc. The excel software is very helpful in proper formatting of the questionnaires. Other word processing and graphic software packages can also be used for producing the questionnaire layout. Revisions of the drafts can be made more easily with the help of this software. The computerized approach to translations enables the overwriting of the verbal parts of these questionnaires in the local languages leaving skip codes, response codes and the general format unchanged.

(5) Refining Logistics

The logistic of the population census can be mind boggling from the stage of the preparation of questionnaires, printing of questionnaires, distribution of the questionnaires, collection of the questionnaires for the data processing, data entry and processing and storage of the questionnaires and each stage will require thorough planning. Equally important is the selection of the interviewers, their training and deployment and monitoring of their field work. Communication between the census administrator and field units and between the interviewer and supervisor can be improved through the use of cell phones and e-mail communication. Video conferencing can be used for training interviewers in different locations ensuring the uniformity of training. The data collected in the field can be entered into computer through the use of laptop computers and transmitted to the data processing unit through e-mail. In the LSMS

survey in Bosnia-Herzegovina , data collected and entered in the field were transferred immediately to a command center through a modem and software called **computer anywhere**. Use of new technologies involve the additional costs and one should take into consideration the trade-off between cost and time while deciding the use of such technologies.

(6) Data Entry

Commercially produced data entry packages are fast superseding customized data entry programmes adopted in earlier censuses. There are several data entry packages available in the market and their use in the census has also increased considerably. The Sudan census in 1993 produced all the results within two years after the census because of the use of IMPS software developed by the US Bureau of Census. Other similar software like Blaise, ISSA, EPI-INFO can be used but each have limitations. Developments in Software are occurring rapidly and there will be other software packages available soon meeting the needs of the new round of censuses. For example, the new software CS-PRO combines the features of IMPS and ISSA. CS-PRO is projected as "intelligent software" or an "interactive software" implying that the software points out likely errors at the time of data entry and enables the data entry clerk to decide whether to enter the data or to correct it. Concurrent data entry makes it possible to use it in a pilot census or in the first month of the population census to gain insights into the quality of fieldwork enabling the improvement of the training and better control of the fieldwork. However, to my knowledge, the use of this software has not been tested in a census environment. Eventually with advances in the computer technology, enumerators will be able to enter data directly into the portable laptop computer during the interview thus eliminating paper questionnaires completely. This will require totally new approach to questionnaire design, organization of the fieldwork, system of the quality control and management of data.

(7) Data Management and Dissemination

Every change in technology, for example improved data entry programmes, portable hardware and computer assisted interviewing will have implication for data management and dissemination. For example, with new data entry programmes the data manager has to insure that the household data files prepared by different data entry operators are included without duplication or omission and convert the large number of households based files into a few thematic files. A master version of the files may be maintained in ASCII, but it may be necessary to convert them into other formats to facilitate their use by data users who use different software packages such as SPSS, SAS, STATA, etc. It should be ensured that the different thematic files could be matched and merged for further analysis. The data manager should take note of the current trends in handling this task.

The traditional method of dissemination data through hard copies of census data will need to be maintained since the large majority of users especially in developing countries still do not have the access to computers, internet, etc. At the same time developing

countries need to adopt these new technologies of dissemination such as internet, e-mail etc side by side with the traditional methods of dissemination. They should also use new softwares for mapping and graphics for better presentation of the information.

IV. Maintaining Census Related Activity in the Intercensal Years

In the past, census taking was a discrete activity, with each new census rising from the ashes for the previous census like the proverbial phoenix. Not any more. There is a continuing need for several post-censal activities. The following activities are particularly important: (i) Conduct and analysis of Post-enumeration Surveys (ii) Establishing a Master Sample of Enumeration Areas, (iii) In-depth analysis of data and Population Projections based on the census results (iv) Undertaking demographic and social surveys during the inter-censal period, (v) Establishing Information Systems, (vi) Census mapping, (vii) Using census results to evaluate other records like population and vital registration systems, (vii) Preparations for the next census. The paper addresses the issues and methodological considerations involved in all the above activities.

(1) Conduct and Analysis of Post enumeration Surveys

In most developing countries the PES, in the manner it is done at present serves as means of self-assurance rather as means for correcting the data. Countries are able to boast that their census are 99.5 is complete and so on, but hardly anything is done to correct the data. This is not the case in developed countries like USA and Canada where the corrections to census data are carried out using PES results. There is a need to change the concept of PES from the means to correct the data to the means of informing the analyst about the pitfalls in the data, so that the analysts can apply their own corrections. The PES should be focused on identifying the coverage and content errors in key variables like rural/urban, employment and unemployment, gender, children ever-born/surviving, occupation, school enrollment, etc. This would help the analyst to look at the data more closely rather than drawing erroneous conclusions. There is a need to thoroughly re-examine the concept and meteorology of PES to make it useful to identify problems that are overlooked in the census and have a consequent impact on the census results.

(2) Establish Master Sample of the Enumeration Areas

A master sample is a large number of clusters of households which are selected by the statistical agency following the census and for which updated household listings are maintained. From these sample of households sub-sample of households are selected for each survey. The master sample should be available for each analytical domain or sub-groups of population – rural, urban, provinces or regions, etc, for which information is required. These master samples should be updated every two years or as when the survey is implemented.

(3) *In-depth Analysis of Data*

Soon after the census only the essential tabulation required by the analysts and policy makers are produced in the form of an abstract. This will typically include the frequency tables of the various variables covered in the census according to the selected socio-economic groups and analytical domains. Further in-depth analysis of the data is necessary to establish relationships between different variables and for studying changes during inter-censal period. Analysis of trends and differentials in vital rates and revision of population projections based on the results of the census are important post census activities.

(4) *Undertaking Demographic and Social Surveys*

Following the census the statistical agency could undertake sample surveys to obtain information on current demographic and social indicators. Examples of such surveys are the World Fertility Survey, Demographic and Health Surveys, Living Standards Measurement Study, the Multi-Indicator Cluster Survey etc. It is a common practice in many countries to undertake special surveys of subgroups of the population such as women, children and elderly for a greater understanding of their problems and needs. Panels of households could be followed to observe changes in employment, fertility etc.

(5) *Establishment of Information System*

Using the census data as a basis an information system can be established providing key socio-economic indicators and this must be updated annually using information available from sample surveys, administrative records, and estimations based on the census. The information system could provide key indicators at the national, regional and provincial levels, and could be accessed by users easily.

(6) *Census Mapping*

This is the weakest area in the preparation for the census in every developing country and needs to be addressed. There are two kinds of problems – maps may not be available or missing due to the war, mishandling, etc. Secondly, maps may be available but are incorrect due to changes in boundaries, construction of new dwelling units, factories, etc, changes in street names, street numbers of buildings. This is particularly significant in urban areas, which undergo change due to annexation of rural area or re-drawing of boundaries between the areas within the locality. It is necessary to strengthen the cadastral and mapping authorities to carry out revisions of maps on regular basis at least once every five years. The census authorities should secure these maps before the house listing operation and update them based on the information obtained during the house listing operation.

(7) Using Census records to evaluate other administrative records

There is a great potential for matching the records of the vital registration systems with the records of the census to determine the sources of errors in one or both systems. For example, the children born during the year preceding the census can be matched with the birth records to find out the extent of omissions. The dual record approach to demographic estimation can be attempted if independence of the two systems can be ensured.

(6)) Preparation for the Next Census

Preparations for the next census would normally require two years prior the date of the next census. This would include house listing operation, preparation of census maps, preparation of the questionnaire and pre-testing it, and selection and training of interviewers and getting all the logistics of the census organized. In the past census organizations were similar to the proverbial Phoenix, which arose from its ashes. This was the case in India with the census organization up to the Fifties. However it was found necessary to have a permanent institution of Census Commissioner and Registrar General with a core staff on a permanent basis. At the state level the office of the director of census operations is also maintained with a core staff of census statisticians and computer personnel. However the large body of staff – enumerators, supervisors, data entry operators etc are recruited prior to the census and disbanded soon after the census. The census organization in India is separate from the central statistical organization and national sample survey. When census taking is a responsibility of statistical agency, the staff of the agency serves as census planners and administrators. This is the only practical way continuity and institutional memory of the earlier census can be retained and utilized in later census. In some countries (like Iraq) census records are stored for a long time as a reference material for the administrative purposes (e.g. for purchasing property in Baghdad the persons should be enumerated in the earlier census). There are countries which keep their records for ten years and then discard them. With the advance of computers census records can be stored in the CD ROMs, Diskettes, Bernoulli boxes, etc and they can be retrieved anytime.

V Conclusions

In a large majority of countries in Asia there is no alternative to population censuses as the means for collecting a wide range of social and demographic data required for administrative and planning needs. Even where vital registration and population registration systems are efficient there is need for population censuses to provide the information for small areas and for providing the frames for surveys. Moreover it is doubtful if the range of topics covered by the population census can be included in the population registration systems. Consequently, population censuses have acquired new meanings and new dimensions in the context of the growing needs for information and the changes in the technological environment. This paper addresses the issues and dimensions of the population census of the new millennium.

The technology of census taking has come a long way from manual counting to the use of unit record equipment, then to first generation computers, personal computers and presently laptop computers. There has been advancement in the software available for data processing and analysis. Opportunities for application of new technologies are present in the planning, execution, tabulation and dissemination stages. The paper illustrates the use of new technologies in (i) mass media, (ii) questionnaire layout, (iii) refining logistics, (iv) data entry, (v) data management, (vi) data dissemination. The traditional methods of disseminating information like hard copies of census publications are fast giving way to supply of information through diskettes, CD-ROMs, email and posting of information on the internet.

In the past, census taking was a discrete activity, with each new census rising from the ashes for the previous census like the proverbial phoenix. Not any more. There is a continuing need for several post-censal activities. The following activities are particularly important: (i) Conduct and analysis of Post-enumeration Surveys (ii) Establishing a Master Sample of Enumeration Areas, (iii) In-depth analysis of data and Population Projections based on the census results (iv) Undertaking demographic and social surveys during the inter-censal period, (v) Establishing Information Systems, (vi) Census mapping, (vii) Using census results to evaluate other records like population and vital registration systems, (viii) Preparations for the next census. The paper addresses the issues and methodological considerations involved in all the above activities.

Censuses will no longer remain statistical exercises carried out by the statistical (or census) agencies of the countries with or without involvement of other stakeholders. With so many domestic and international stake-holders interested in census results, the challenge of census taking now depends on balancing the demands of these stake-holders and the resources available, external funding, technological infrastructure etc.