

# China's Inter-census Survey in 2005

Deputy Commissioner            Zhang Weimin

Division Director                Li Xiru

Deputy Division Director       Cui Hongyan

National Bureau of Statistics of China

According to government's decision of China, a full population census will be conducted every ten years at the year ended by 0 and an inter-census population survey, also called 1% population sample survey, will be conducted between two full censuses at the year ended by 5. Except for the 1% coverage of population, the inter-census survey is almost the same as the full census in terms of organization, procedures and questionnaire design. The full census and the inter census survey constitute the essential framework for China's population statistics. This framework, providing a system of guaranty the acquisition of accurate and up-to-date demographic information, represents the major source for tracing the change of population quantity and structure.

In the first two decades of the 21<sup>st</sup> century, China will experience dramatic evolution in economic transition and social transformation. During this development period, it is no doubt that China's population, including total size, regional distribution and other characteristics, will have a great change. In this context, the inter-census population survey, which will be conducted in 2005, is of great significance for understanding present situation of China's population and extrapolating the future population change.

Technologically, the inter-census population survey in 2005 (hereinafter referred to as "1% sample survey ") is an improvement and optimization to the 2000 Census of China. In this sense, a careful review of the 2000 Census, especially finding out the imperfections and the causes, may provide substantive information for the design of the 1% sample survey, and for the improvement and optimization of the future population censuses. This paper will talk about the major deficiencies of the 2000 Census and major concerns for the design of the 1% sample survey of China in 2005.

## **I. Major Deficiencies of the 2000 Census**

The 2000 Census is an overall success, which delivers much valued experience in motivating the public participation, launching massive information, education, and communication campaigns, enhancing enumerator training and improving the pre-census discovery of real situation. However, there are still some deficiencies in

data quality and the unprecedented challenge in full coverage of population. According to the post-enumeration check, the under-enumeration rate of the 2000 Census is 1.81%, more than 30 times of the 1990 Census, which is 0.6%. Chinese demographers, through careful investigation and research, hold the opinion that 1.81% reflects the real situation of the under-enumeration of the 2000 Census, and the under-enumeration mainly spreads over the age group 0-9.

Though the 1.81% under-enumeration does not produce critical impact on the quality and use of the census data, and is not more deteriorating than the under-enumeration in most countries around the world, it is a sharp rise compared with the previous censuses in China. Given the gigantic population in China, 1.81% represents under-enumeration of more than 22 million people. What is the reason for under-enumeration? Answering this question justifiably may help us bring the increasing trend of under-enumeration to a halt.

Many factors contribute to under-enumeration in population census, including design, organization, implementation and environment of the census.

### **1. Census Design**

To gain the total population size, the first is to define where the people should be interviewed and enumerated. This is to ensure that everyone is enumerated, but only at one place, so as to avoid under-enumeration or over-enumeration. In this context, the scientific, clear-cut and operable definition of target population of census is critical for quality of census data.

According to the design of the 2000 Census of China, everyone should be enumerated at his permanent residence. This is the same with the previous censuses. The only difference is that the timeline for permanent residence has been reduced from one year to a half year. This concept of permanent residence combines situation of household registration with length of residence time, instead of on the sole basis of length of residence time.

Definitions of population of permanent residence:

- a. People who reside at their legal residence with permanent household registration;
- b. People who reside at a place for more than a half year but the places of their permanent household registration are elsewhere;
- c. People who reside at a place within a half year and the places of their permanent household registration are elsewhere and they have left their legal residence for more than a half year; and if they have departed their legal residence within half year, they belong to the permanent population at their legal residence with permanent household registration;
- d. People whose places of household registration have not been yet settled;
- e. People who are now working and studying abroad and they are still on the permanent household registration.

From the above definitions, we may find the weakness of the onsite enumeration:

- a. This definition of permanent population is viable if the household registration is rather complete and the inconsistency between actual and registered residence has essentially been eliminated. However, in a period of fast economic growth and urban expansion, people who are residing out of permanent residence, in migration and not being listed on household registration can be seen everywhere in China. The number of such people has reached the top of history. The 2000 Census showed that 144 million people have been living out of their places with household registration for more than a half year. These people, plus those away from home for less than six months, have addressed a great challenge for the enumerators to judge their permanent residence, creating a hidden risk for the quality of census data.
- b. While judging whether a person is a permanent one, or whether he should be enumerated locally, the enumerator is in passive position and holds inadequate evidences. Before enumeration, the enumerator has to ask the interviewee how long he has been living there. Some people living out of their places with household registration tend to report a longer time of local residence, resulting in double enumeration at both their current residence and their places of household registration. If they report a shorter-than-actual time of local residence, it may possibly create an under-enumeration case.

From the above, we may see that to judge if the people should be enumerated is a more “subjective” matter less restricted by operable criteria. This has added the difficulty for enumerators and brought along adverse impact on accuracy of total population.

## **2. Organization and Implementation**

Effective organization and implementation of an established census program is critical for ensuring the success of a census. This includes budgeting, enumerator recruitment and training, related regulations and cooperation from the public. With the establishment of market economy system, the co-existence of multiple economic elements and diversification of interest groups, the organization of census has been improved to adapt to the changing of census environment. Major deficiencies in the organization of the 2000 Census include:

- (1) Difficulty in enumerator recruitment and finance. Because of massive corporate restructuring, lay-off for productivity enhancement and co-existence of different ownership, enumerator recruitment represents an emerging challenge. In the 2000 Census, some incumbent organizations and companies just appointed enumerators without considering their qualification. What’s more, according to the current finance, the subsidies for enumerator are granted by local government. Enumerators in financially powerful regions may gain appropriate subsidies for their work, while their counterparts in poor regions may be inadequately or even not subsidized for their

involvement in the census. This inevitably influences the commitment and quality of service of the enumerators.

- (2) Difficulty in policy execution. Population census is a policy-based event that involves personal secrecy and confidentiality as well as benefits or punishment to an administrative body. If relevant policies are not formulated or can not be implemented, the quality of census will not be guaranteed.

Before the 2000 Census, the government formulated several census regulations through careful investigation, such as, it is prohibited to use the census findings as the evidence for personal punishment and as a measure for government performance assessment; the census offices should keep the census findings in confidentiality and refrain from furnishing or disclosing them to any others. However, these policies still have to address the enforcement challenge linked to a spectrum of worries--the fertile couples worry about the punishment due to breach against the one-child policy; the migrants worry about additional charge and being expelled; the officials at the grassroots level worry about their performance measured by such indicators as per capita income and local achievement ranking; the companies worry about additional charge due to illegal employment. This, as rumored "individuals worry about punishment, officials worry about performance assessment, and enterprises worry out additional charge", has made it difficult to conduct the on-site enumeration established census program, affecting the quality of the census data.

### **3. Changing Environment of the census**

Much improvement was made to the design of the 2000 Census to adapt to the changing socio-economic development. But the results tell us that we still underestimated the impacts of the changing environment of the census. For examples, the residence instability increases the difficulty for individual access; the broken linkage between the place of household registration and present residence, the removal and resettlement, two residence places pose formidable difficulty for enumerator to judge where they should be enumerated. Many self-employed people working at irregular hours are still inaccessible even at the tenth visits by the enumerators; people are more aware of privacy protection and less cooperative in the enumeration. The enumerators have to complete the questionnaire outside the door during some household interviews.

These factors not only aggregated the under-enumeration rate of the 2000 Census, but also address a formidable challenge for the 1% sample survey in 2005.

## **II. Major Ideas on Design of the 1% Sample Survey in 2005**

In order to ensure the quality of the 1% sample survey, the State Council delivered to the provincial governments the *Notice on Conducting 1% Population Sample Survey* in the end of last October to commence the run-up for the survey.

### **1. Objectives**

Since the 2000 Census, with the accelerating socio-economic development, China's

population, its structure and residence have changed dramatically. A good many of emerging questions are yet to be answered. For examples, how large the population is and how fast it increases? What is the latest change of TFR since the 1990s? How big the impact of the one-child cohorts on fertility level? Whether the floating population increases or decreases? What is the latest change of the structure of floating population? What is the latest relationship between labor supply and demand? How about the old people's living standards and social security coverage with Chinese people getting "elder and elder"? These questions are all directly related to the realization of goals set in China's modernization strategies. The 1% sample survey is just designed to trace the latest change of China's population and structures, to study the developing trend, provide accurate and dependable information for the public and the government for development planning and policy making.

## **2. Organization**

- (1) Setting up leading offices to coordinate all the departments involved in the census is a valued experience derived from the previous censuses. For this inter-census survey, an inclusive leading group comprising of representatives from the departments of statistics, public security, population and family planning, propaganda, development and reform, finance, civil affairs and commerce and industry administration will be set up at various levels, responsible for the overall management and motivation for the survey. Under the leading group, an executive office led by the statistical department and consisting members from the public security and family planning will be responsible for the design, organization and implementation of the survey.
- (2) Household interview and on-site questionnaire fill-in will be adopted. Enumerators will be recruited from the staff of township (streets) and village committees and other institutions, and subsidies will be granted to them. As estimated, some 70,000 enumerators and 20,000 supervisors will be recruited for this event, who will work for a period of about 2 months.
- (3) The reference time of the survey will be zero hour of November 1, 2005. the field enumeration will be completed in 10 days.

## **3. Sampling Methodology**

This survey will cover all the 2,861 counties of China, so as to ensure the representative of the 333 prefectures.

Three-stage cluster sampling method - PPES (Probability Proportionate to Estimated Size) will be used in this survey. The minimal sampling unit is the enumeration district (equivalent to a resident groups of village committee in rural area and neighborhood committee in urban area). On the first stage, about 44,000 township-level sampling units will be selected. On the second-stage, village-level sampling units will be taken from the first-stage clusters. On the third stage, about 60,000 enumeration districts will be taken from the second-stage clusters.

Use of sampling frame. China introduced GIS (resolution down to township level) for its 2000 Census. The 1% sample survey in 2005 will make use of this GIS for the first-stage cluster selection. Addressing the change of township boundaries and population size, we are now updating the GIS maps. For the second- and third-stage sampling, the administrative maps of the affiliates of each selected sampling units are made out, then the number of households and persons in each enumeration district is made clear to form the sampling frame.

#### **4. Target Population**

Target population shall mean which person in the selected enumeration districts will be interviewed and enumerated, or, which person should be filled in the questionnaire. The 1% sample survey in 2005 will change the definition of target population from “permanent” to “current” population, expressed as “all people in the selected enumeration districts”, including the persons who are on the household registration but absent at time of enumeration. The change of target definition is beneficial for the accurate enumeration of the total population.

#### **5. Questionnaire Design**

The questionnaire of this survey is designed on the principles of putting both demands and possibilities into consideration, meeting the demands of both governments and the public, combing visions and insights, justifiability and operability, to substantively reflect the current status and future change of China’s population. Appropriate adjustment will be made on the basis of the long and short form of questionnaires used in the 2000 Census, adding some items that may reflect on the population development and people’s living situations. By far, a pilot to test the feasibility of the questionnaire is in progress. The questionnaire will be finalized by April 2005. Preliminary design is as follows:

(1) 34 individual items:

a. Individual profile, including name, relation with household head, sex, date of birth and nationality.

With these information, we may estimate the population size, age and sex structure, regional distribution, household size and structure, size of aged population and aging speed, composition and distribution of nationalities.

b. Items on migrant status, including current residence, household registration, time and reason of leaving the residence with household registration, usual residence one year ago, and permanent residence five years ago.

With this information, we may judge whether the enumerated persons are permanent population, or coming from outside, or going out, or people on the household registration. These are the major items for the judgment of permanent residence by computer after the adjustment to the target population.

The migration-related items allow us to gain the detailed information about

inconsistency between actual residence and residence with household registration and about population migration, and provide substantive information for study of floating population and transfer of surplus rural labor.

c. Items on the one-child, including question on number of brothers and sisters

The one-child policy implemented for nearly three decades has shaped a unique one-child cohort that has grown up to childbearing ages. These items provide information for the study of qualification, economic activities, marriage, birth and living situations of the one-child cohort and for study of maternal history through of indirect analysis.

d. Items on health, including questions about physical health and disability.

These items are designed to assess the physical health of individuals and map the social perspectives of disabled persons, which may reflect the individuals' self-care and provide information for study on care of the old, service and support.

e. Education, including questions on literacy, educational level and academic fulfillment.

In this survey, the education-related questions on the long form of questionnaire adopted in the 2000 Census will be used without change, which may facilitate the longitudinal comparison, reflect the educational level, and allow us to understand the educational discontinuance of school age groups and assess the effects of the 9-year compulsory education program.

f. Employment and unemployment, including questions on last-week employment, work time, industries, occupation, recent monthly income, employment status, access to job vacancies, unemployment status and reasons, source of income of unemployed persons.

These items provide information about labor resources and structures, employment and unemployment rate, employment scale and structure, and composition of economically inactive population, reflecting on the living standards of Chinese people.

g. Marriage and birth, including questions about marriage status, age at first marriage, number of births, number of surviving children, sex and age of child born in the past 12 months.

Traditional to Chinese census, these items are extremely necessary for gaining the information about childbearing timeline and sequence, especially from the perspective of one-child policy. These items may provide essential data for the study of childbearing age variation, fertility and birth history. Besides, given the under-enumeration of 0-9 age groups in the 2000 Census, which results in less accurate estimate of the actual fertility in the 1990s, the items on birth history may allow us to rectify our past data.

(2) 20 household-related items:

a. Household profile, including questions about household type, number of persons that should be enumerated, births and death in the past 12 months.

These items are designed to understand the essentials of Chinese households, determine the number of persons who should be enumerated, and allow logic check of the individual enumeration.

b. Housing, including questions on number of rooms, floor space, age of housing, construction materials, kitchen facilities, fuel use, source of water, sanitary facilities, housing tenure, and the value or monthly rent of the property.

The 2000 Census introduced for the first time the housing-related items, providing substantive information for the study of Chinese people's housing conditions and living quality and standards. The 1% sample survey in 2005 will keep these items to allow for five-year comparison.

### (3) 6 death-related items

Including name, sex, age, nationality and education of persons died in the past 12 months.

Same as the previous censuses, the death-related items may provide information for the calculation of life expectancy, infant death rate, number and structure of the death, which is of great significance for the study of physical improvement and improvement of sanitary conditions.

## **6. Data Release**

The data of the 1% sample survey will be released in two lots. Preliminary data, including region- and age-specific population and households, birth rate, death rate, natural growth rate will be released 100 days after the enumeration, i.e. early February 2006. The tabulations will be disclosed by the end of 2006.

## **III. Features of Design of the 1% Sample Survey in 2005**

The design of the 1% sample survey in 2005 is well aligned with the current socio-economic development and takes into consideration the variation of population and the census environment, which, compared with the 2000 Census, delivers the following features:

### **1. The concept of target population will be redefined**

Reasonable definition of target population is the first priority for any population census and sample survey. The definition of target population, be it "permanent resident" or "current resident" or "legal resident", is to ensure accuracy of the population. This is the primary and essential principle of any population censuses.

The definition of target population of census or survey should follow the general principles of scientific, accurate and operable, specifically, (1) ensuring avoidance of under-enumeration and over-enumeration; (2) facilitating enumerator's comprehension and operation and interviewee's reporting; (3) ensuring easiness of use

of data; (4) being able to reflect the characteristics of population change and regional distribution; (5) taking into consideration the current population management; (6) taking into consideration the available manpower, material and financial resources. Scientific definition of survey target is crucial for the quality. Despite the fact that the target definition may vary with census or survey, the above general principles must be observed and a scientific and practical program aligned with the socio-economic development should be made.

The 2000 Census introduced a “permanent resident” as the target population that combines household registration with length of residence. The persons enumerated at their legal residence and the persons who are not on the household registration are defined as “permanent population”. The persons not residing at their legal residence are resolved depending on their length of residence (see Part I for detailed definitions). This definition is reasonable and scientific in itself. But the problem is, in actual enumeration, the enumerators must have an acute understanding of the definition, and the interviewees must accurately report their actual place of household registration and the length of their residence. Where a person should be enumerated is not at the discretion of the enumerator, instead, it depends on the answers of interviewees. Given the massive floating population and inconsistency between actual and legal residence and the worries about being expelled and fines, it is hard to ensure the reality and accuracy of the answers of the migrants, leading to under-enumeration or over-enumeration. From the experience of the 2000 Census, the under-estimation of the impacts of inconsistency between actual and registered residence as well as the population migration and the individual cooperation increases the difficulty for field enumeration and affects the data quality.

The 1% sample survey in 2005 will take the “current resident” as the target population, defining “current resident plus absent resident” as the target population, i.e. “all the persons residing at the enumeration districts” and “all the persons who are on the household registration but absent” at zero hour of November 1, 2005 will be counted and enumerated. Whether a person is a permanent resident will be judged through the computer program according to his household registration and the length of residence or absence. For household members absent for more than six months, only the short form of questionnaire is required to fill in. This may minimize the effort needed in filling in of the questionnaire.

#### **Benefits of these improvements:**

First, the enumerator does not need to judge whether a person is a permanent one and whether he should be enumerated. The enumerator just enumerates everyone residing in the enumeration district or on the household registration, without needing to ask how long the interviewee has lived here, does he have a legal residence and where is his legal residence before the enumeration. This may facilitate the enumerators’ understanding of the survey targets and their field work.

Second, all the persons coming from outside, regardless of the length of stay, must be enumerated. This may help avoid under-enumeration because of unreal report of stay

time for fear of fines and being expelled.

Third, enumeration of floating population at both their original and current residence may ensure integrity of the information about floating population, for this allows us to evaluate the quality of data about floating population and analyze the floating population characteristics by the way of original and current residence balance.

Fourth, since the target population are redefined as “current population plus persons on household registration”, the computer program may distinguish the permanent population, population with household registration, emigrants (time sensitive) and immigrants (time sensitive), freeing the enumerators from personal judgment and increasing the accuracy of the data.

Disadvantages:

Enumeration of persons coming from outside and persons going outside may increase the workload of enumerators by 11~12% compared with the 2000 Census. This increase might be higher in regions with large amounts of immigrants and emigrants.

The redefinition of target population is a proactive measure taken against the inconsistency between actual and legal residence, massive migration and weakening individual cooperation, the aim is to avoid under-enumeration and over-enumeration and ensure accuracy of information about distribution of permanent population.

## **2. New items in Questionnaire will be added**

Giving full consideration to the present socio-economic development, new items in the questionnaire designed for this survey, on the basis of the long form and short form of questionnaires introduced in the 2000 Census, will be added about emigration, one-child cohort, social securities, physical health, disability, income, and adjustment will be made on the items of migration, employment/unemployment and maternal history.

The feasibility of new items is now under test through pilot survey.

## **3. Samples will be controlled effectively through strict sampling method**

PPES will be introduced in this survey. Two controls will be implemented to minimize the sampling bias to ensure the effectiveness of samples and increase the data quality.

First, sampling frame at different level will be made according to the principle of administrative region. A majority of township or village in China have their distinct administrative boundaries, despite the geo-overlapping between some township or village. For a sampling unit, if the administrative boundary is not clear, it should be made clear before the enumeration (only for survey purpose, irrelevant to the administrative jurisdiction). Such clarification should lead to non-overlapping of enumeration districts. All the persons residing inside all kinds of buildings (including factory, shop, government complex, school and hotel) at the selected enumeration districts, regardless of their administrative region, will be regarded as the target

population.

Second, intentional change of sample will be strictly prohibited. The minimal sampling unit for this survey is enumeration district. Several enumeration districts will be taken randomly after village-level sampling units are selected. To prevent intentional change to another enumeration district with better traffic, good financial conditions and easier enumeration, it is requested that after the village sampling unit are selected, the enumeration districts will be divided and coded in certain serial number and a detailed village map containing the boundary, code and geo-location of every enumeration district will be made out. The maps should be sent to the provincial census office for supervisory. Then, the enumeration districts will randomly be taken. Once selected, any enumeration district can not be changed as all the enumeration districts are coded consistently, the code of enumeration districts selected in a township or a county are uniformly distributed. Any intentional changes of enumeration district will be avoided.

In addition to the above three features and improvements, we will take some other effective measures to address the enumerator recruitment and financial challenges, work with related departments to formulate substantive regulations, launch public awareness programs to solicit support and cooperation from the public.

#### **IV. Conclusion**

The 1% sample survey in 2005 is a large-scale national survey after the 2000 Census, which will be launched amid the critical and strategic period of socio-economic development in the first 20 years of this century, the 1% sample survey in 2005 is of great significance for understanding China's current population, studying developing trend of population and driving the harmonious development between population and socio-economy. While building upon the success of previous censuses in leadership style, enumerator recruitment and training and household interview, this survey will make some improvements and optimization to the definition of target population and sample control to ensure accuracy and quality of the data. In addition, it makes necessary adjustment on the items in questionnaire for the purpose of better utilization, analysis and research on the survey data.

Data quality represents the core and lifeline for population survey. With the change survey environment, how to ensure the data quality is the focus of the survey design. It is respected that this survey may further improve data quality to provide more accurate information for the government decision-making and for the public use, and provide successful experience for the 2010 population census.

#### **Reference**

1. Qiao Xiaochun, *China Census Studies*, China Population Press, 1995.10
2. Y. C. Yu, *Methods to Collect Demographic Data*, China Statistics Press, 1997.3

3. United Nations, *Principles and Recommendations for Population and Housing Censuses Revision 1*, ST/ESA/STAT/SER.M/67/Rev.1
4. Zhang Weimin and Cui Hongyan, *Evaluation of Data Quality of China's 1990 Census*, International Symposium on China's 1990 Census, 1992.10, Beijing
5. Zhang Weimin, "Reflections on Problems in the 2000 Census", *Population and Economics*, Capital University of Economics and Business, 4<sup>th</sup> issue, 1998, Beijing