CENTRAL POPULATION AND HOUSING CENSUS
STEERING COMMITTEE

THE 2009 VIETNAM POPULATION AND HOUSING CENSUS:

MAJOR FINDINGS

Hanoi, 6 - 2010

FOREWORD

The 2009 Population and Housing Census was implemented at 0 am on 01 April, 2009 according to Prime Ministerial Decision No. 94/2008/QD-TTg dated 10 July 7, 2008. This is the fourth population census and the third housing census implemented in Vietnam since its reunification in 1975. The purpose of the Census is to collect basic data on population and housing for the entire territory of the Socialist Republic of Vietnam, to serve national development planning.

Data from the 15% Census sample survey were processed immediately after the data collection was completed and were disseminated on 31 December, 2009. To provide timely users with major findings of the census, following the publication: "The 2009 Vietnam Population and Housing Census of 00.00 hours 01 April 2009: Implementation and Preliminary Result" and pocketbook "The 2009 Vietnam Population and Housing Census: Some key indicators", the Central Population and Housing Census Steering Committee has undertaken the task of compiling and publishing the third report entitled *"The 2009 Vietnam Population and Housing Census: Major Findings"*.

This report contains 4 parts:

Part I explains the design and implementation of the Census, and consists of 2 chapters. Chapter 1 introduces the Census implementation process. Chapter 2 explains the design, estimation and evaluation of census sample survey results.

Part II contains subsequent 7 chapters that present the main results extrapolated from the 15% Census sample survey. Chapter 3 describes features of population size and structure. Chapter 4 presents estimates on fertility and Chapter 5 on mortality. Chapter 6 analyses migration and urbanization. Chapter 7 evaluates population quality through educational attainment and training qualifications. Chapter 8 analyses indicators related to labour and employment. Chapter 9 analyses housing and amenities of households.

Part III presents data tables extrapolated from the 15% Census sample survey.

Part IV contains the appendixes related to the Census such as the concepts and definitions used in the Census, census forms, sample allocation, standard errors for major indicators, ...

This publication was completed with technical and financial support of the United Nations Population Fund (UNFPA). We would like to express our sincere thanks to PhD. Peter Xenos and Ms. Sara Bales, UNFPA funded consultant and staffs of the UNFPA Office in Vietnam for their great contribution to the edition and finalisation of the report.

I highly appreciate and thank the staffs of the General Statistics Office, who have worked with their enthusiasm and whole-heartedness for release of this publication.

We are pleased to present to readers in Vietnam and abroad, a publication with a voluminous amount of information. Although our great efforts have been made in compiling this material, it is difficult to avoid all errors, omissions and limitations, especially demand for the in-depth information use. We hope to receive feedback and contributions from our readers to learn from mistakes and improve subsequent of the Census publications.

DEPUTY CHAIRMAN OF
THE CENTRAL CENSUS STEERING COMMITTEE

Nguyen Duc Hoa

VICE MINISTER OF THE MINISTRY OF PLANNING AND INVESTMENT
CUM DIRECTOR GENERAL OF THE GENERAL STATISTICS OFFICE

TABLE OF CONTENTS

TABLE IN TEXT

FIGURES

MAP OF VIETNAM

R1. Northern Midlands and Mountains
02. Ha Giang
04. Cao Bang
06. Bac Kan
08. Tuyen Quang
10. Lao Cai
11. Dien Bien
12. Lai Chau
14. Son La
15. Yen Bai
17. Hoa Binh
19. Thai Nguyen
20. Lang Son
24. Bac Giang
25. Phu Tho

R2. Red River Delta
01. Ha Noi
22. Quang Ninh
26. Vinh Phuc
27. Bac Ninh
30. Hai Duong
31. Hai Phong
33. Hung Yen
34. Thai Binh
35. Ha Nam
36. Nam Dinh
37. Ninh Binh

R3. North and South Central Coast
38. Thanh Hoa
40. Nghe An
42. Ha Tinh
44. Quang Binh
45. Quang Tri
46. Thua Thien Hue
48. Da Nang
49. Quang Nam
51. Quang Ngai
52. Binh Dinh
54. Phu Yen
56. Khanh Hoa
58. Ninh Thuan
60. Binh Thuan

R4. Central Highlands
62. Kon Tum
64. Gia Lai
66. Dak Lak
67. Dak Nong
68. Lam Dong

R5. Southeast
70. Binh Phuoc
72. Tay Ninh
74. Binh Duong
75. Dong Nai
77. Ba Ria -Vung Tau
79. Ho Chi Minh City

Map legend:
1 Bắc Ninh
2 Hưng Yên
3 Vĩnh Long
4 Hậu Giang

R6. Mekong River Delta			
80. Long An	84. Tra Vinh	91. Kien Giang	95. Bac Lieu
82. Tien Giang	86. Vinh Long	92. Can Tho	96. Ca Mau
83. Ben Tre	87. Dong Thap	93. Hau Giang	
	89. An Giang	94. Soc Trang	

* According to regulations in Government Decree No. 92/ND-CP dated 7 September, 2006, the Northeast and Northwest regions are merged to form the Northern Midlands and Mountains; the North Central Coast and South Central Coast have been combined to form the new region with the name North and South Central Coast. Quang Ninh has been moved from the Northeast region to the Red River Delta. Ninh Thuan and Binh Thuan provinces have been moved from the Southeast to the North and South Central Coast region.

** In the part of figures, the Region 1 throught Region 6 is symboled for short as R1 throught R6, respectively.

Part I
DESIGN AND IMPLEMENTATION OF THE CENSUS

CHAPTER 1
CENSUS IMPLEMENTATION

The 2009 Population and Housing Census was implemented according to Prime Ministerial Decision No. 94/2008/QD-TTg dated 10 July, 2008. This was the fourth population census and the third housing census implemented in Vietnam since the nation was reunified in 1975. The Census aimed to collect basic data on the population and housing for the entire territory of the Socialist Republic of Vietnam, to provide data for research and analysis of population and housing developments nationally and for each locality. It responded to information needs for assessing implementation of socio-economic development plans covering the period 2001 to 2010, for developing the socio-economic development plans for 2011 to 2020 and for monitoring performance on Millennium Development Goals of the United Nations to which the Vietnamese Government is committed.

1. History of the Vietnam population censuses

Vietnam has a long tradition of censuses over several centuries. However, almost all these censuses, in reality, have only served to count the population to know how many people had to pay taxes, or to conscript soldiers to serve in wars. Thus, these counts were not regular and collected few details. During the period of development of Socialism in the North and the struggle to unify the nation (1945-1975), the Democratic Republic of Vietnam implemented two population censuses covering the North, in March 1960 and April 1974. Right after the end of the war in 1975, at the beginning of 1976 the State organized a total population count for all provinces in the South to serve National Assembly elections and reconstruction of the Nation.

The first real population census of the reunified Socialist Republic of Vietnam was implemented at the end of 1979. With the resources and technical capacity available at the time, the 1979 Census provided basic data of good quality to serve as the basis for the country's development.

The second national population census, regarded as the first modern Population and Housing Census in Vietnam, was implemented in April, 1989. This Census used internationally accepted concepts, design and data processing

technologies. Many individuals and organizations that participated in the Census consider that the *de jure* population enumerated in the Census was nearly complete and the results were of high quality. In the following decade, this Census has provided a rich source of data on demography, society and the economy and has been used by many stakeholders.

The third National Population and Housing Census was implemented in April, 1999. Many features of the 1989 Census were included in the design and implementing guidelines for the 1999 Census. In addition, the 1999 Census added questions and expanded the contents to provide more comprehensive and detailed data.

The fourth National Population and Housing Census was implemented in April 2009. Compared to the previous censuses, the 2009 Population and Housing Census had a broader scope and a larger contents. The Communist Party and Government led and directed the Census and the people consented to and supported the Census. At the same time, this Population and Housing Census inherited and drew from the experience of the previous Censuses, especially the 1989 and 1999 Population and Housing Censuses. This was the third Census for which United Nations Population Fund (UNFPA) has provided technical and financial assistance for some important activities of the census.

2. Census subjects

The 2009 Population and Housing Census enumerated all Vietnamese regularly residing in the territory of the Socialist Republic of Vietnam at the reference point of 0:00 on 01 April, 2009; Vietnamese citizens given permission by the authorities to travel overseas and still within the authorized period; deaths (members of the household) that occurred between the first day of the Lunar Year of the Rat (07 February, 2008) to 31 March, 2009; and residential housing of the population.

Population and housing censuses were implemented simultaneously taking the household as the survey unit. The household could include one individual who eats and resides alone or a group of individuals who eat and reside together. For household with 2 persons and over, its members may or may not share a common

budget; or be related by blood or not; or marital or adoptive relationship or not; or in combination of both.

The household head was the main respondent. For information of which the head of household was unaware, the enumerator was required to directly interview the survey subject. For information on labour and employment, the enumerator was required to directly interview all respondents aged 15 and older; for questions on births, the enumerator was required to directly interview women in childbearing ages (from 15 to 49 years of age) to determine the responses. For information on housing, the enumerator was required to directly survey the household head and/or combine this with direct observation to determine the information to record in the forms.

3. Survey contents

The survey contents appeared in two types of survey questionnaire: the complete census form (*short-form*) and the census sample survey form (*long form*).

Contents of the complete census form
a) On population:
- Individual information (full name, sex, month and year of birth/age);
- Relationship to household head;
- Current school attendance;
- Educational attainment;
- Ethnicity and religion;
- Literacy status.
b) On housing:
- Current housing situation;
- Structure of dwelling and primary construction materials;
- Area of dwelling;
- Year the dwelling was put into use.

Contents of the census sample survey form

Besides information collected in the complete census form, the census sample survey form collected additional information as follows:

a) On population:

- Place of usual residence 5 years before the Census;

- Disability status;

- Marital status;

- Highest professional/technical qualifications;

- Economic activity status in the 7 days prior to the Census.

b) On fertility for women aged 15-49:

- Whether or not the woman ever gave birth;

- Number of children born, number who survived and number who died;

- Month and year of birth, number of boys, number of girls for the most recent birth.

c) On deaths:

- Mortality situation in the household;

- Individual information about the death;

- Cause of death and maternal mortality.

d) On housing:

- Use of dwelling;

- Ownership of dwelling;

- Primary fuel source for lighting and cooking;

- Main water source for drinking and eating;

- Type of toilet used;

- Some basic amenities of the household.

4. Comparison with the 1999 Population and Housing Census

It is the fact that experience from the 1999 Census has had a large influence on the design and contents of the 2009 Census. The Population and Housing Census in Vietnam is implemented periodically once every 10 years, facilitating analysis of trends and direct comparison of data between two censuses. In contrast

to previous censuses, the 2009 Population and Housing Census had a larger scope. Besides contents similar to previous censuses on sex, age, ethnicity, educational attainment, marital status and demographic features, etc., this Census has (i) changed the concept for the economic activity indicator, shifting from classification by usual activity in the past 12 months to classification by current economic activity in the past 7 days; (ii) surveyed additional indicators on disabilities, cause of death, maternal mortality and basic amenities in the household.

Two new major modifications of the Census were proposed and applied. First, the census sample survey size was increased to 15% of the population to allow expansion of the survey contents, to allow estimates to be made at the district level, and to reduce costs. The sizes of the 1989 and 1999 Census samples were only 5% and 3% respectively, which were only representative at the provincial level. Second, Intelligent Character Recognition/scanning technology for data entry replaced traditional keyboard data entry. This technology has many technical advantages, but also proved to be a major challenge for the General Statistics Office, because this technology requires improving the quality and clarity of recording in survey forms and more attention to keeping the forms in good condition.

For the first time in the Vietnam Population Census, a Census hot-line was set up at the Central and Provincial Census Steering offices. The main objective of the hotline was to assist in guiding the technical operations of the Census, and at the same time to receive feedback from the people on the Census activities. These objectives were achieved. In addition, the hotline was a useful tool for publicity, providing information about the Census to the people. The hot-line number was widely disseminated in the mass media. This made the people feel there was greater transparency in the Census, and on this basis they felt more trust and cooperated enthusiastically, providing a large amount of accurate information to the census office and staff.

The 2009 Population and Housing Census is required to satisfy national information needs and international comparison. At the same time, the contents of the census cannot be too burdensome; the contents must be feasible and affordable within the available fund envelope, while securing high quality of the data

collected. To balance the above requirements, the General Statistics Office has actively studied and collaborated with relevant Ministries and sectoral agencies to organize many workshops with data users and to implement pilot surveys to determine the number of indicators that need to be included in the census and to ensure collection of high quality data.

5. Census preparation and planning

5.1 Development of the Census Plan

Preparation of the Census was initiated quite early. It began with organization of a user workshop on 9 October, 2006.

On the basis of information needs, survey indicators were drafted. Based on user needs, two new modifications (expanded sample size and character recognition/scanning technology) were proposed and applied to the Census. These two modifications required much more detailed preparations, especially the trialling of the new data entry technology.

For effective preparation of the Census, on 3 August, 2007, the Prime Minister issued Directive No. 19/2007/CT-TTg on preparation of the Population and Housing Census at 0:00 on 01 April, 2009. Following the spirit of this directive, the Ministry of Planning and Investment (General Statistics Office) took the lead, and collaborated with other Ministries and sectoral agencies to quickly implement preparatory work for the Census.

With the technical and financial assistance of UNFPA, the General Statistics Office developed a 2009 Population and Housing Census Plan and a Census Information Campaign Plan. Already from the beginning of 2008, two booklets of information documents were printed and issued to all relevant Ministries and sectoral agencies, provinces and municipalities. Following this, the questionnaires and survey materials were designed and tested three times before final approval . After each pre-test and the final Census pilot, a workshop was organized to evaluate all activities undertaken and to learn lessons for the subsequent time. The survey questionnaire for the third pre-test and the pilot survey were used to test the scanning technology.

On 10 July, 2008, the Prime Minister issued Decision No. 94/2008/QD-TTg on the organization of the 2009 Population and Housing Census.

On 16 October, 2008, the Party Central Secretariat issued Directive No. 27-CT/TW on leadership to successfully implement the 2009 Population and Housing Census.

Following this, the Census Steering Committees at the central, provincial, district and commune levels were set up to implement the Census protocol, issued under Decision No. 09/QD-BCDTW dated 10 October, 2008.

5.2 Census pre-test and pilot

The census questionnaires and materials were tested through three rounds of pre-testing and one pilot test prior to being officially launched for the Census.

Three pre-tests were implemented to refine the survey protocol, census questionnaires and other technical guidelines, with pre-test locations at the 6 provinces/cities under directly to the Central Government including Ha Noi, Vinh Phuc, Thai Nguyen, Quang Tri, Ho Chi Minh City and Ben Tre.

In August and September, 2008, the final Census pilot was successfully implemented in 48 enumeration areas belonging to 4 provinces (12 enumeration areas per province) representing all regions including: Yen Bai representing the Northern Midlands and Mountains; Nam Dinh representing the Red River Delta; Quang Binh representing the North and South Central Coast and Central Highlands; and Tien Giang representing the Southeast and Mekong Delta.

Provinces and municipalities and relevant Ministries and sectoral agencies contributed staff to participate in the pre-testing and final pilot test. The results of these surveys, along with experience gained, served as the basis for refining the survey protocol, census questionnaires, census procedures and technical guiding documents for collection, processing and synthesizing results of the Census.

5.3 Mapping and listing number of dwellings, households and individuals

The Census Steering Committees at all levels checked and ascertained administrative boundaries between communes, wards and district capitals; urban and rural districts, provincial capitals and provincial level cities, and between provinces and major municipalities in order to detect locations with households and regular residents for which the managing administrative unit is unclear, in order to serve the purpose of drawing maps, making household listings and assigning responsibility for survey work.

Training on drawing maps and listing households was implemented at two levels. Two level training was intended to improve the quality of training. A majority of provinces/municipalities complied with instructions for two-level training. However, because of the large number of trainees, and the short time-frame, some provinces/municipalities took the initiative to train at 3 levels, that is, they added a level of trainer at the provincial level to train the district level, then these district trainers trained lower level staff on drawing maps and making household listings.

Map drawing maps and household listing were implemented right after the training. During this work, some provinces/municipalities relied initially on digitized map systems of the Ministry of Natural Resources and the Environment.

Drawing maps and listing households began in October, 2008 and was basically finished by January, 2009. The map drawing was implemented by cadastral workers at the commune/ward level, while responsibility for listing households was assigned to local people who were well-informed about the local terrain and population distribution in the enumeration area. These people were trained on map drawing and household listing over a 3 day period.

6. Census publicity

Experience from previous Censuses indicated that the success of the Census depends, to a large extent, on the attention, guidance and support of State agencies, local authorities at different levels and the cooperation of the people throughout the country. Based on this experience, for the 2009 Census, the Central Steering

Committee gave instructions for a strong Census publicity campaign with a detailed plan, concrete contents and diverse forms so all people could hear about it. An in-depth and widespread information campaign for the 2009 Population and Housing Census aimed to:

(a) Take advantage of the support and assistance of State agencies, mass media agencies and mass organizations through creating adequate awareness of the importance and purpose of the Census;

(b) Make the people understand the benefits of the Census, and become aware of their rights and responsibilities with regard to the Census so they are more willing to cooperate, and to provide complete and accurate information.

Prime Ministerial Decision No. 94/2008/QD-TTg and Central Party Secretariat Directive No. 27-CT/TW on leadership in successfully implementing the 2009 Population and Housing Census was disseminated to all Party cells and to all local block captains and hamlet leaders. In order to implement the above Decision and Directive, all levels of the Party, local authorities, together with the Census Steering Committee at all levels, regularly organized *meetings to verify implementation progress*, and in this way mobilized also the participation of the political system. This is one of the important factors that contributed to the success of the Census.

Publicity materials for the 2009 Census were quite diverse, including: printed materials, audio-visuals, and mobile presentations. The widespread and in-depth publicity campaign began in March, 2009, and was promoted most strongly during the period from 15 March to 20 April, 2009. Publicity activities focused on the period surrounding the Census dates (from 1 April through 20 April, 2009) including: daily Census news items, news reports, Census songs that were broadcast and posted daily on the radio, television, newspaper at both the central and local levels, or even broadcast on the hamlet or block loudspeaker system.

Reports of 63 provinces and municipalities, supervisory teams of the Central Census Steering Committee and teams representing the international donors who monitored the Census were consistent in their assessment that "1 April, 2009 was really the day of the Census festival", with the power to persuade the whole population to actively participate and support census staff.

7. Survey data collection and supervision

7.1 Enumeration area

The country was divided into 172,000 enumeration areas. Clear or relatively clear borders were identified for each enumeration area, so that the number of households per enumeration area was appropriate for one enumerator to complete survey work within the allotted time. The number of households in an enumeration area varied across geographic regions. In general, ***enumeration areas are hamlets (or blocks, villages or other small units)*** with an average size of about 100 households.

7.2 Recruitment and training of enumerators

Training aimed at providing the Census with knowledgeable personnel able to direct and train operational and field staff at all levels, all the way down to the team leader and enumerator level.

In order to be recruited, enumerators had to meet criteria such as a strong sense of responsibility, ability to write numbers and letters legibly, with educational attainment of grade 10/12 or higher (in highland and remote areas, the lowest acceptable educational attainment was grade 7/12) and preferred ages from 25 to 50 years. Enumerators could be farmers, civil servants, students, teachers, military border control staff, etc. Almost all interviewers were recruited from the local area. In a few areas in the high mountains, or extremely remote areas where the majority of inhabitants were ethnic minorities with limitations in educational attainment, the enumerators were brought in from other areas. In these areas, the local people (usually hamlet leaders) helped to guide enumerators around and interpret for them.

The team leader was the workforce that directly managed the work of enumerators. For recruitment of team leaders priority was placed on the following criteria: (1) ability to organize team work, and (2) commitment to allocate 100% of their time to managing the enumerators during the implementation of the survey in their locality.

Besides the team leader, there was a workforce of supervisors at all levels from central, provincial, to district levels with primary responsibility for supervising the activities of the network of team leaders and enumerators, monitoring and providing technical support for the lower level Steering Committees.

Training for census implementation was implemented at 3 levels:

+ Central level trainers instructed provincial level trainers: Training took place over 8 days in which trainees listened to an introduction on the purpose of the Census, clear explanations about the survey forms, guidance on methods for recording information in survey forms, and survey management and supervision.

+ Provincial level trainers instructed district level trainers: Training took place over 6 days. The number of trainees in each province/municipality depended on the number of urban and rural districts in the province/municipality.

+ District level trainers instructed team leaders and census sample survey enumerators: Survey training for 60 000 people who directly participated in the Census sample survey took place over 8 days. By 26 March, 2009, they had run 900 training courses for the Census sample survey throughout the country. Besides the Census sample survey training courses, the district level trainers also ran over 5000 training courses for team leaders and enumerators to understand the complete census form, with each course lasting 4 days. All training courses were comprised of presentations on the enumeration manual, the supervision manual and organization of practical training in the classroom and in the field. The quality of courses was carefully monitored and evaluated.

The careful preparation of training materials helped to improve the quality of training. Besides the technical materials, there were three exceptional training resources made available for the 2009 including: (1) DVD to support training; (2) Large size printed version of the questionnaires; and (3) the Enumerator training manual. These materials were appropriately designed and reproduced in large numbers to serve training at both the central and local levels.

7.3 Survey and supervision requirements and procedures

The 2009 Population and Housing Census applied the method of face-to-face interview. Enumerators came to each household in the enumeration area to which they were assigned to interview and record information on the questionnaire for each person. Interviewing and recording followed instructions that had been explained during training courses. As the interview was about to be completed in a given household, enumerators reread all responses so the household head could listen and confirm that the responses had been recorded accurately.

Enumerators were provided with the resources they needed for daily surveying. Besides the questionnaires and survey materials, the enumerators were also supplied with pens, erasers, enumerator ID cards, a bag to hold papers, a rain poncho (for areas that experience heavy rains). During the Census enumeration days, enumerators were paid a stipend. Although the stipend amount was not high, almost all enumerators exhibited enthusiasm and dedication to the survey work. As they received careful training, they undertook interviewing and recording in questionnaires quite competently. At the same time, the publicity campaign helped the people to understand more completely about the Census, so the actual data collection was implemented quite conveniently and according to plans.

In order to improve the quality of information collection, supervision was implemented for all steps of the Census, from supervision of training on map maps, actual drawing of maps and household listing, training for implementation, reviewing and adjusting maps and listings, through recording of information in questionnaires. Among those tasks, the interview and recording of information in questionnaires was identified as the most important, especially when applying the intelligent character recognition and scanning technology for Census data entry.

Supervision of interviews and recording in questionnaires was undertaken by the central, provincial, district and commune level supervisors, supplemented by a large workforce of team leaders who also contributed to success in completing questionnaires. The responsibility of the survey team leaders in the 2009 Census was emphasized more than in previous Censuses.

7.4 Surveying and supervision of enumeration

On 24 March, 2009, under the direction of the Standing Vice Prime Minister of the Government, Chairman of the Central Steering Committee, H. E. Mr. Nguyen Sinh Hung, the Central Steering Committee successfully organized an on-line conference to check on the preparations for the initial steps in data collection, which indicated that all preparatory steps had been completed, and everything was ready.

Early in the morning on 01 April, 2009, the Census steering committee at the district and commune levels convened all census workers and together starting at 7 in the morning, all enumerators and team leaders throughout the country gathered to initiate implementation of the Census, simultaneously commencing interviews and recording information in questionnaires at households.

In parallel with recording information in the household questionnaires, starting at 7 in the morning on 01 April 2009, the district and commune steering committees collaborated with the police force to simultaneously organize surveying of the homeless who live on sidewalks, the streets, markets or on waterways, etc.

The survey in all enumeration areas was implemented over 7 days, and in each census sample survey enumeration area over 15 days. By 15 April, 92% of enumeration areas had completed interviews and questionnaires ; the remaining enumeration areas had to prolong the survey to 20 April for various reasons including rainstorms, complicated geography, too large size or many absent households hindering contact. Therefore, it lasted to 20 April. The rate of progress was in line with the Census protocol.

Starting with the day Census data collection began, many localities organized practice interviews to learn from mistakes, in order to limit errors to a minimum. In the first 3 days of the Census, the steering committee at all levels purposely maintained a slower pace than normal in order to identify all errors of enumerators, and to draw lessons from mistakes in a timely manner.

During the enumeration period, some questions arose related to use of concepts and definitions for cases specific to a given locality. These questions were

answered quickly and in a uniform manner over the telephone, e-mail and on the "2009 Population and Housing Census Webpage for Census worker".

8. Data processing and compiling

The 2009 Population and Housing Census applied Intelligent Character Recognition technology/scanning technology for direct data entry from census forms to the computer to replace the traditional keyboard data entry that is commonly used in Vietnam at present. This is an advanced technology, and the first time it had been applied in a statistical survey in Vietnam, so preparatory work had to be done carefully and meticulously. Through organization of many workshops and 7 pilot applications with technical and financial assistance from the UNFPA, the new technology was mastered, and the Census Steering Committee Standing Committee approved use of this technology to process the entire results of the 2009 Population and Housing Census. Currently the Government has decided to allocate funds through the project on Modernization of the General Statistics Office using World Bank Loan funds to procure the scanning system equipment, software and technical assistance. The successful use of this technology will create a precedent for continued use of scanning technology in other statistical surveys.

After checking and coding at the Provincial/municipal steering committee office, (both the complete census and the census sample survey), forms were checked and accepted then transferred for processing to one of three Statistical Computing Centres in Hanoi, Ho Chi Minh City and Da Nang.

Data processing was implemented in only a few locations, following standard procedures and a fixed timeline. The steering committee at each level and processing centres fully implemented their assigned responsibilities, especially the checking, transmitting and maintenance of survey forms in good condition.

The Central Steering Committee collaborated with the Statistical Computer Centres to set up a plan for processing and compiling results, setting up tabulation plans, interpreting and synthesizing output tables, and developing options for extrapolating from sample to population estimates.

The General Statistics Office completed the work of developing software applications and training using ReadSoft software (the one used in pilot testing), organized training on network management and training on systems and programs for logic checks and data editing, developed a data processing protocol, integrated these systems and completed data flow management programs.

The General Statistics Office collaborated with the contractor, FPT, to develop software applications, train staff, trial the system, complete the programs,... using the new TIS and E-form software.

Compilation of results was implemented in 2 stages. In stage 1 data were compiled from the Census Sample Survey by the end of October, 2009, and in stage 2, data were compiled from the completed census forms, with work finalized in May 2010.

9. Quality assessment

One extremely important requirement of the 2009 Population and Housing Census was that the quality of the survey data be assured. Implementation of the 2009 Population and Housing Census reflected results of human resource development at the General Statistics Office over many years. Much of the design and implementation work during the 2009 Population and Housing Census benefitted from and involved further development of methods applied in the 1989 and 1999 Population and Housing Censuses, from the census publicity campaign through the training program and survey regulations. The results are that the General Statistics Office has gained a high level of self-reliance in its ability to organize the 2009 Population and Housing Census with reduced financial and technical assistance from international and foreign organizations.

Right after completing the data collection, the Central Steering Committee Standing Committee implemented many methods to assess and verify outcomes of the Census and the results are as follows:

• Verification and supplementary enumeration through the hotline:

The Central Steering Committee Standing Committee collaborated effectively with Vietnam Television and Voice of Vietnam radio to announce to the people throughout the country that they should help to provide information to identify all survey subjects who may have been missed so supplementary enumeration could be performed. By 20 April, 2009, the Central Population and Housing Census Steering Committee Office had received feedback from 975 households throughout the country who called the hotline. The results were checked, and 744 households were interviewed to add to the census results, while the remaining 231 households were found to have already been enumerated through other household members or through indirect methods. Although the number of households that were supplemented was not large compared to the nation's 22 million households, through this activity, the households and Census Steering Committee at all levels showed their consent and trust in the Census results.

• Post-enumeration survey:

Right after the survey work was completed, a sample post-enumeration survey was undertaken in order to evaluate the accuracy of the Census data. A systematic sample involving 60 enumeration areas was randomly selected for reinterview organized directly by the Central Steering Committee Office.

With the objective of evaluating the completeness of the Census, all usual residents of each household were asked four questions : (1) full name, (2) relationship to household head, (3) sex, (4) month and year of birth or age. The responses to these questions were compared to the survey forms (after they were transferred to the higher levels) to find any cases that were missing from the Census or in the post-enumeration survey. Results of the post-enumeration survey indicated a net error rate of -0.3% (equal to the difference between the undercount rate and the duplication rate), or equivalent to about 258 000 people (see detailed table below). If we compare this rate with the error rate of 1.5% to 4.5% of the total population normally found in population and housing censuses in other countries, and at the same time comparing to the previous Census in Vietnam, completeness of this Census is assessed as high.

Socio-economic region	Duplication rate	Undercount rate	Net error (Undercount rate – duplicate rate)
Entire country	**1.8**	**1.5**	**-0.3**
Northern Midlands and Mountains	2.7	2.0	-0.7
Red River Delta	2.8	2.1	-0.7
North and South Central Coast	1.1	0.9	-0.2
Central Highlands	1.5	2.0	0.5
Southeast	0.7	0.8	0.1
Mekong River Delta	1.8	1.6	-0.2

• *Evaluation of the quality of recording in forms*:

Immediately after the enumeration was completed, the Central Steering Committee Office collaborated with the Statistical Computing Centre to organize teams to check the quality of recording in the census forms. The results of this exercise indicate that the quality of recording in forms met requirements, and ensured the ability to use scanning technology for processing.

• *Checking focal sites for subjects with a high likelihood of being omitted:*

In order to have a stronger basis for releasing preliminary results of the Census, in the first half of July, 2009, the Central Steering Committee Standing Committee decided to undertake sentinel checking for 3 types of subjects with a high probability of being missed. These subjects included: (1) students living in dormitories, (2) students living in boarding houses in the enumeration area, and (3) workers working in industrial zones who rent housing from the local population or who live in tents or other temporary housing of the project or production enterprise far from residential centres. Particularly for Hanoi, Ho Chi Minh City and Binh Duong, places with a large number of in-migrants, the Central Steering Committee directly implemented the checking.

Results of checking indicate that students in dormitories were rarely omitted, because each dormitory has a management board that knows clearly the situation of their housing and information on the individual students, and members of the management board are usually the team leaders or enumerators involved in the survey work. For the remaining two groups, the checking revealed that there is little duplication or omission of these groups, i.e. the undercount and duplication

rates are about the same as those found in the post-enumeration survey for the general population.

Besides the methods for checking and post-enumeration mentioned above, during the process of undertaking the Census, the Central Steering Committee organized checking to review the results of drawing maps and listing dwellings 2 times, and to review the preliminary compilation of results 3 times. Thus, through many different methods of checking and post-enumeration surveys, we come to the assessment that the 2009 Population and Housing Census has collected reliable and stable results, appropriate with the effort and work contributed by the political system, the Steering Committees at all levels, the various sectors and the active participation of the people throughout the country.

CHAPTER 2
DESIGN, ESTIMATION AND EVALUATION
OF THE CENSUS SAMPLE RESULTS

1. Sample size

In the 2009 Population and Housing Census, besides a full enumeration, some indicators were collected in a sample survey. The census sample survey was designed to: (1) expand survey contents; (2) improve survey quality, especially for sensitive and complicated questions; and (3) save on survey costs. To improve the efficiency and reliability of the census sample data, the sample size was 15% of the total population of the country. The sample of the census is a single-stage cluster sample design with stratification and systematic sample selection. Sample selection is implemented in two steps: *Step 1*, select the strata to determine the sample size for each district. *Step 2*, independently and systematically select from the sample frame of enumeration areas in each district to determine the specific enumeration areas in the sample.

The sample size of the two census sample surveys in 1989 and 1999 was 5% and 3% respectively, only representative at the provincial level; sample survey indicators covered fertility history of women aged 15–49 years and deaths in the household in the previous 12 months. In the 2009 Census, besides the above two indicators, many other indicators were also included in the census sample survey. The census sample survey provides data representative at the district level.

When determining sample size and allocation, the frequency of events was taken into account for various indicators including birth and deaths in the 12 months prior to the survey, and the number of people unemployed in urban areas, etc.; efforts were also made to ensure the ability to compare results between districts within the same province/municipality and between provinces/ municipalities.

2. Stratification and sample allocation across strata

To ensure representativeness of the sample for each district throughout the country and because the population size is not uniform across districts or

provinces, the Central Steering Committee decided to allocate the sample directly to 682 out of 684 districts (excluding 2 island districts) throughout the country in 2 steps:

Step 1: Determine the sampling rate $f^{(r)}$ for 3 regions including:

- Region 1: including 132 urban districts;
- Region 2: including 294 delta and coastal rural districts;
- Region 3: including 256 mountainous and island districts.

Step 2: Allocate the sample across districts in each region based on the sampling rates for each region as determined in Step 1 using the ***inverse sampling allocation*** method. Through applying to this allocation method, the number of sampling units in each small district is increased adequately to ensure representativeness.

The formula used to calculate the sample rate for each district in each region is as follows:

$$f_i = \frac{a_i \times f_1}{1 + (a_i - 1) \times f_1}$$

In which: - i is the order number of each district ($i = 2, 3, \ldots m_r$.);
- 1 is the first district in each region;
- m_r is the code for the district in the region (r=1, 2, 3);
- $a_i = N_1/N_i$; N_i is the population of district i;
- f_1 is the sampling rate for the first district which is calculated as follows:

$$f_1 = \frac{f^{(r)} \left(1 + \sum_{i=2}^{mr} a_i^{-1} \right)}{(m_r - 1)}$$

The determination of the sampling rate for each region $\mathbf{f^{(r)}}$ using the above methodology requires weights, with the weights being the average population in the district (Ni) for each region estimated for 1 April, 2009 as follows: $f^{(1)}$ =13.11%; $f^{(2)} = 13.16\%$ and $f^{(3)} = 22.68\%$. The results of allocating sample to districts and provinces are presented in Annex 1.

3. Sampling unit and method

The sampling unit is the enumeration area that was ascertained in the step to delimit enumeration areas. The sampling frame is the list of all enumeration areas that was made following the order of the list of administrative units at the commune level within each district. In this way, the whole country has 682 sample frames (682 strata).

The provincial steering committee was responsible for selecting sample enumeration areas using systematic random sampling as follows:

Step 1: Take the total of all enumeration areas in the district, divide by the number of enumeration areas needed in the sampleto determine the skip (k), which is calculated with precision up to 1 decimal point.

Step 2: Select the first enumeration area (b, with b ≤ k), corresponding to the first enumeration area to be selected. Each successive enumeration area to be selected will correspond to the order number: $b_i = b + i \times k$; here i = 1, 2, 3.... Stopping when the number of enumeration areas needed has been selected.

4. Estimation and extrapolation methods

General weights can be calculated based on probability/weights as follows:

1) Design weights (basic weight): based on probability;

2) Weight adjustment factor due to changes in number of households or number of enumeration areas due to loss without replacement;

3) Weight adjustment factor based on structure of the reference population (weights).

Symbols:

W_{1hji} - Design weights (basic weights) of enumeration area *j*, in strata *h*;

W_{2hji} - Weight adjustment factor for changes in number of households (or population);

W_{3hji} - Weight adjustment factor for variation in average size of enumeration area in strata h;

W_{4hji} - Weight adjustment factor for changes in the number of enumeration areas;

W_{5hji} - Post-stratification weight related to overall urban/rural and sex structure of the reference population;

W_{hji} - Sample weight for household (or male/female population) of enumeration area j strata h.

Determining the basic weight

Assume that a_h is the number of enumeration areas selected in strata h and N_h is the total number of enumeration areas in strata h. Because the sample is selected independently in each strata using systematic random sampling, the basic selection probability is calculated using the formula: $P_{1hji} = \dfrac{a_h}{N_h}$ and the basic weight (design weight) of enumeration area j in strata h is the inverse of the probability of selection, calculated as follows:

$$W_{1hji} = \frac{1}{P_{1hji}} = \frac{N_h}{a_h} \approx \frac{M_h}{\sum m_{hj}}$$

In which, M_h is the total number of households (population) of strata h and $\sum m_{hj}$ is the total number of households (population) of each enumeration area selected in the survey for strata h.

Determining the weight adjustment factors related to enumeration areas

a) Weight adjustment for changes in the number of households (population) in the enumeration area:

Suppose m_{hj} is the total number of households (population) estimated at the time of listing enumeration areas j in strata h and m_{hj}^* is the total number of households (population) enumerated at the time of the survey in enumeration area j in strata h. The adjustment factor for changes in the number of households (population) is calculated according to the following formula:

$$W_{2hji} = \frac{1}{P_{2hji}} = \frac{m_{hj}}{m_{hj}^*}$$

b) Each enumeration area in the 2009 Population Census constitutes a population of non-uniform size ranging from about 100 to 500 people per enumeration area (plus or minus 20 households), therefore it is necessary to determine the adjustment factor for the number of households (or population) of the 2009 Census enumeration areas in terms of the average number of households (or population) in that strata. Supposing that $\overline{m_{hj}}$ is the average number of households (or population) in that enumeration area in strata h and the adjustment factor for changes in the number of households (or population) is calculated according to the following formula:

$$W_{3hji} = \frac{1}{P_{3hji}} = \frac{\overline{m_{hj}}}{m_{hj}}$$

c) Weight adjustment due to changes in the number of enumeration areas :
The design of the Census sample survey on 1 April, 2009 requires that if an enumeration area is selected, but during the process of adjusting census maps and listings it is found that the area has been cleared or lost, the census staff are allowed to replace the enumeration area with an adjacent enumeration area, leaving constant the number of enumeration areas selected. So:

$$W_{4hji} = \frac{1}{P_{4hji}} = 1$$

Determining the post-stratification weight adjustment (weighting according to the reference population structure)

The number of people in the population on 01 April, 2009 used for weighting, is estimated based on preliminary data of the 2009 Census by urban/rural residence and gender for 63 provinces/municipalities allowing post-stratification according to the population share in urban/rural areas and male/female population. Suppose that m_{hji}^* is the total number of households (male/female population) at the time of the survey in enumeration area j strata h;

$m^{*'}_{hji}$ is the total number of households (male/female population), then adjustment for urban/rural share or male/female share of enumeration area j strata h and calculated using the following formula:

$$m^{*'}_{hji} = m^{*}_{hj} \times \frac{M^{*}_{hi}}{M^{*}_{h}}$$

Of which:

m^{*}_{hj} the number of households (male/female population) collected from the sample survey in the enumeration area j strata h;

M^{*}_{hi} the number of households (male/female population) divided into urban and rural areas estimated from preliminary Census results on 1 April, 2009 for strata h; (i = 1 – urban ; i = 2 – rural)

M^{*}_{h} number of households (population) estimated from preliminary Census results on 1 April, 2009 for strata h

Post-stratification weights for the structure of the population (or number of households) as estimated for 1 April, 2009, is determined as follows:

$$W_{5hji} = \frac{1}{P_{5hji}} = \frac{m^{*'}_{hji}}{m^{*}_{hji}} \times \frac{M^{*}_{h}}{M_{h}} = \frac{m^{*}_{hj}}{m^{*}_{hji}} \times \frac{M^{*}_{hi}}{M^{*}_{h}} \times \frac{M^{*}_{h}}{M_{h}} = \frac{m^{*}_{hj}}{m^{*}_{hji}} \times \frac{M^{*}_{hi}}{M_{h}}$$

Because the sample allocation is not proportional to the population being studied, sample weights will be used for all analysis of data from the 2009 Population and Housing Census sample survey in order to ensure representativeness of the sample. The sample weights for each household (or for population type i) in enumeration area j in strata h is the inverse of the probability of selection:

$$W_{hji} = 1/P_{hji} = W_{1hji} \times W_{2hji} \times W_{3hji} \times W_{4hji} \times W_{5hji}$$

$$W_{hji} = 1/P_{hji} = \frac{M_{h}}{\sum m_{hj}} \times \frac{m_{hj}}{m^{*}_{hj}} \times \frac{\overline{m}_{hj}}{m_{hj}} \times \frac{m^{*}_{hj}}{m^{*}_{hji}} \times \frac{M^{*}_{hi}}{M_{h}} = \frac{\overline{m}_{hj}}{\sum m_{hj}} \times \frac{M^{*}_{hi}}{m^{*}_{hji}}$$

The Central Census Steering Committee Office has collaborated closely with the Central Statistical Computing Centre to program, make concrete

estimates, and check accuracy of the weights for all 30 720 sample enumeration areas in the 2009 Population and Housing Census.

5. Standard error calculation method

Estimates from the Census sample survey were affected by two types of error: (1) non-sampling error, and (2) sampling error. Non-sampling error is the result of errors in implementation of data collection and processing such as visiting the wrong dwelling, interviewing the wrong household, mis-understanding of questions by respondents, other errors on the part of respondents or enumerators and data entry. Although many efforts were made to minimize these errors while implementing the survey, non-sampling errors cannot be completely avoided are difficult to evaluate statistically.

On the other hand, sampling error can be evaluated statistically. The sample of respondents in the Census sample survey is only one of many possible samples that could be selected from the total population using the same sample design method and required sample size. Each of these possible samples could give different results from the sample actually selected. Sampling error results from variation in results from the many different possible samples. Although it is not possible to know this variation precisely, it can be estimated from the sample survey results.

Sampling error is usually measured using *standard errors* for specific statistical indicators (means, proportions,...). The standard error is, in fact, the square root of the variance. Standard errors can be used to calculate confidence intervals which contain the true value for the population. For example, for a given statistical indicator calculated from the sample survey, the true statistical value will fall within the interval plus or minus two times the standard error for that indicator with confidence equal to 95% for all possible samples of the same size and sample design.

If the sampling unit is selected following simple random sampling, then it is possible to use formulas to directly calculate standard errors for the sample. However, the Census sample survey was designed with strata, and therefore a more complicated formula must be used. Computer software that could be used to

calculate standard errors for stratified samples includes the standard error calculation module of ISSA or STATA. These programs use the *Taylor linear expansion* method to estimate variance for estimates of means and proportions for sample surveys.

The Taylor linear expansion method considers the mean or proportion as a ratio estimate, $r = y/x$, for which y is the total sample value of variable y, and x is the total number of events in the group or sub-group being studied. The variance of r is calculated using the following formula, where standard errors are the square root of the variance:

$$SE^2(r) = var(r) = \frac{1-f}{x^2} \sum_{h=1}^{H} \left[\frac{m_h}{m_h - 1} \left(\sum_{i=1}^{m_h} z_{hi}^2 - \frac{z_h^2}{m_h} \right) \right]$$

where:

$$z_{hi} = y_{hi} - rx_{hi}, \text{ and } z_h = y_h - rx_h$$

In which:

  h - indicates the strata ranging from 1 to H,
  m_h - is the total number of enumeration areas selected in strata h ,
  y_{hi} - is the total of weights of variable y for enumeration area i, in strata h,
  x_{hi} - total number of weighted events in enumeration area i in strata h, and
  f - the overall sampling rate, (if this value is very small it can be dropped).

Sampling error in the Census sample survey is calculated for a few selected key indicators. Results are presented in an appendix for estimates at the national, urban and rural, and 6 socio-economic region levels and for 63 provinces/municipalities. For each variable, the statistical estimate (R), standard error (SE), relative standard error (SE/R) and 95% confidence interval (R±2SE) are presented in Annex 4.

The confidence interval (for example, when estimating the indicator *sex ratio at birth*) can be explained as follows: the national estimate for sex ratio at birth is 110,5 boys to 100 girls with a standard error equal to 0.54. Therefore, for a 95% confidence interval, adding and subtracting 2 times the standard error to the estimate gives us, 110.5±2×0.54. With a high probability (95%) the national sex ratio at birth lies in the interval from 109.5 to 111.6 boys per 100 girls.