

19<sup>th</sup> Population Census Conference  
26-28 April 2000, Beijing, China

## A Plan for Better Products and Services of the 2000 Population and Housing Census in Korea

by

Hyung-baik Yun  
Director, Population Census Division

and

Hyung-seog Kim  
Demographer, Population Census Division

National Statistical Office  
Republic of Korea

### Introduction

1. The 2000 Population and Housing Census in Korea is scheduled to be conducted as of November 1 this year. Since the population census was commenced in 1925, the forthcoming census in the year 2000 will be the sixteenth one. By being fully utilized in policy formulation, implementation and evaluation, the Korean censuses have undoubtedly contributed to national and regional development. Thus, census-taking has been well supported in Korean society so far. Even though the Asian Crisis limited government finances in 1999, a large budget has been allocated for the 2000 Census. The budget of the 2000 Census in the fiscal year of 2000 amounts to 83.4 billion *won*, twice as much as that for the previous census.

2. However, questions on the validity of census-taking have been frequently raised among users. Increasing costs in census-taking and untimeliness of the results are the major issues. It is far from being feasible to claim that the census should be carried out merely for historical tradition. Without significant improvement in terms of costs and benefits, the extinction of the census would only be a question of time. In this context, the 2000 Census would be a test case for its destiny.

3. On the other hand, what throws census planners into a more embarrassing situation is the deteriorating environment of census-taking. With growing awareness of privacy, more and more people have become reluctant to cooperate with surveys. Also, the increasing number of absent households can be pointed out in this regard. In addition, the weakened statistical function of local governments through the downsizing program of the Korean government threatens the success of census-taking in the year 2000. These changes affect various aspects of census-taking such as census organization, enumeration methodology, enumerator recruitment and training. The rapid shift of society into info-communication is also exerting much influence on the ways of census-taking across all operations from the selection of topics to the dissemination of data.

4. To tackle the problems derived from the changes in the census milieu stated earlier, the Korea National Statistical Office(KNSO) has tried to find some measures to implement the 2000 Census in more economic, accurate and timely manner. This paper aims at sharing the experiences that the KNSO has had in preparation for the 2000 Census.

## **More economic**

5. Since the expense for the enumerators claims almost three-fourths of total costs of the Korean census, census planners have explored ways toward reducing the size of enumerators to be paid. First of all, the possibility of introducing the self-enumeration method has been intensively studied to replace the traditional canvasser method, which requires a large number of paid workers. The self-enumeration method is believed to curtail the number of enumerators to a great extent. This approach was initially proposed by audit authorities.

6. The KNSO has studied the feasibility of the self-enumeration method from the viewpoint of response rate and accuracy through several pilot surveys. The surveys revealed that it was not possible to apply the self-enumeration method on a full-scale since there were a sizable number of households that found it cumbersome to fill out the questionnaires. This phenomenon was pronounced particularly in rural areas where one out of three households is headed by persons 65 years of age and over. The inaccuracy of response in this method

was another problem in adopting this method. Many questionnaires were returned to enumerators without being completed.

7. However, the study left the possibility for a selective application of the self-enumeration method. This method is likely to be applicable to the households living in apartments due to the homogeneity of housing conditions. In general, housing facilities in apartments are homogeneous in terms of period of construction, number of rooms, type of living quarters, piped water, space of building, and toilet, kitchen and bathing facilities. Therefore, the accuracy of housing items in the apartment households can be easily obtained. The convenience of delivery/collection of questionnaires in the apartment households through either security guards or apartment management office was also pointed out. These types of households are around 38% out of total number of households in Korea.

8. Accordingly, the KNSO decided to apply the self-enumeration method only to the apartment households in short-form. An enumerator canvassing these households will be given 3 enumeration areas(EA) instead of 2 EAs in the past censuses. It is expected that around 6 billion *won*, 7 percent of total costs, would be reduced by introducing this method in the 2000 Census in Korea.

9. In relation to the self-enumeration method, the first pilot survey in 1999 unveiled that a mail-out/mail-back method appeared to be inappropriate in Korean context due to not only the low response rate, below 30 percent, but also the inaccuracy of postal address lists. The huge tasks of reconfirming incorrect answers in the questionnaires returned by mail discouraged census planners from using this method. Delivery/collect method will be employed in the 2000 Census.

10. A volunteer enumerator program was also studied in the first pilot survey as one of measures to cut down the cost of manpower in the 2000 Census. It was found that volunteer enumerators had a high level of task incompleteness. At present, the KNSO has a plan to make full use of volunteers registered in 152 regional volunteer centers but on fully paid basis not on voluntary basis.

11. The cost aspect of the census has also governed the type of data processing system. In Korea, OMR data capture had been applied twice to the census since 1990. As far as OMR is concerned, it was quite successful in Korea even though it left some operational problems like transcription errors, high printing costs, etc. In the 1995 Census, domestic questionnaires were printed on high quality domestic paper with the cooperation of the KNSO and private enterprises.

12. However, some critical questions were raised on whether three OMR machines purchased between July 1989 and November 1990 can be utilized for the 2000 Census as they seem to be outdated. In particular, unexpected failures of the OMR machines during data processing are stressed as a major

handicap to carry on using the OMR system in the 2000 Census. The OMR machines were downed 65 times and required 18 major repairs during scanning questionnaires in the 1995 Census. In addition, the unavailability of some machine components, such as MARK HEAD ASSY, OMR AMP PKG, etc., is believed to jeopardize the overall data processing for the 2000 Census.

13. The KNSO tried to find a substitute that is more economic than repurchasing the whole OMR system. Thus, the introduction of OCR was intensively reviewed for the 2000 Census since 1997. Three pilot surveys were carried out to study the operational feasibility of OCR since 1997. The immediate findings in these surveys were as follows:

- scanning speed(questionnaires/minute) of OCR was much lower with 50 than that of OMR with 105;
- recognition rates were quite acceptable marking around 99% in marks and 95% in numerals, but poor in characters around 60%;
- verification of recognition after scanning was recommended in Korean context, which resulted in additional workload of around 23,000 man-days for eye-checking every character and numeral;
- questionnaires by the self-enumeration method are less suitable for OCR due to its poorer accuracy compared with the canvasser method;
- 7 large-size scanners would be needed to meet the target date of data release.

14. Four types of the data capture options available were carefully compared in terms of cost, time and accuracy. The OMR/PC and the OMR/OCR were ruled out at an early stage due to both the out-of-dateness of OMR machines and the complexity of mixed measures. Also, the OCR outsourcing was not selected because of the small size of private enterprises in this field for assuming tasks as large as census. The PC outsourcing in a decentralized manner was evaluated as the most economic and time-saving method in Korean context.

Table 1. Comparison of data-capture methods for the 2000 Census in Korea

	OCR outsourcing	OMR/PC	PC outsourcing	OMR/OCR
Total costs (billion won)	61	55	52	59
Time period required(months)	8	5	5	6
Estimated error rates in data entry(%)	2-10	2	2	2-10

15. The KNSO arrived at a decision to contract out the data entry of the 2000 Census to local firms in sixteen provinces. Thereby, sixteen provincial centers of data processing will be established near the regional offices of the KNSO. Editing staffs will telephone or revisit the households of incorrect questionnaires at each provincial center respectively. This kind of decentralized system of data processing is believed to speed up the data-processing time through placing data centers closer to the source of data. As for the type of PC key-entry, the third pilot survey carried out a comparison study between form-entry and line-entry. Line-entry was preferred to form-entry simply because of higher speed of data capture.

### More accurate

16. The delineation of EAs for the 2000 Census is currently ongoing since October 1999. In this process, the digital maps which were produced by the National Geographic Information System(NGIS) of the National Geography Institute are utilized for the first time. The scale of base maps is 1/1,000 in urban area, 1/5,000 in rural area and 1/25,000 in mountainous area.

17. EA maps will also be produced using this technology. In the 1995 Census, most of the EA maps were drawn manually by staffs in *Ups, Myons and Dongs*, the third level of administrative divisions, even though some raster EA maps were available. Therefore, the computerization of mapping tasks is able to save manpower as well as minimize the duplication and omission in the census.

18. One of the major factors to decrease the accuracy of the 2000 Census in Korea would be the increasing number of daytime-absent households caused due to both the growing participation in social activities among housewives and the increasing numbers of one-person households. As enumerators fail to interview absent households during their working period, these types of households might be omitted in count. The security of enumerators might be threatened by the fact that they have to visit such households several times even late at night. The growing numbers of aged households also draw special attentions among census-takers since their responses are generally lacking in accuracy.

Table 2. Trends of one-person households and aged households in Korea during 1985-1995

1985		1990		1995	
Number ('000)	%	Number ('000)	%	Number ('000)	%

Overall households	9, 571	100.0	11,355	100.0	12,958	100.0
One-person households	661	6.9	1,021	9.0	1,642	12.7
Aged households 1)	677	7.1	963	8.5	1,251	9.7
● One-generation	147	1.5	238	2.1	389	3.0
● Single	115	1.2	193	1.7	349	2.7

Note : 1) Households whose head is 65 and over.

19. The following steps were recommended to tackle the problem of absent households, which were proven useful through several pilot surveys:

- Firstly, the role of supervisors should be significantly strengthened. They have not only to check the content of questionnaires collected by enumerators but also to canvass those households that are not covered by enumerators due to being absent. Every supervisor in the 2000 Census should take charge of 10 enumerators instead of 15 in the 1995 Census. It should be pointed out that supervisors in the 1995 Census played a nominal role in encouraging enumerators to complete their work;
- Secondly, in case of absent households, enumerators should be allowed under certain conditions to collect only some basic items such as sex, age group, marital status, type of household, type of occupancy, type of living quarters and number of rooms used. This information can be tapped from either neighbors or administrative records;
- Thirdly, imputation, as a successive step for the measures mentioned above, should be made for these absent households.

20. One of in-field quality controls to be addressed is the utilization of basic common information among apartment households or row households. Enumerators should collect in advance some basic information on housing conditions that are common among those households during the preparatory period. Several items, such as construction year, housing facilities, piped water, type of living quarters, total area of floor space and number of rooms, can be marked on the questionnaires by enumerators before either interviewing respondents or administering questionnaire. This approach is expected not only to improve the accuracy of responses but also to minimize respondent burden.

21. However, as the quality of data in the census depends primarily on the cooperation of respondents, the KNSO gives priority first to the enforcement of public relations. Various mass media will be employed for the effective campaign of the census. The symbol mark, logo and mascot of the 2000 Census have already designed for this purpose. Around 3.3 billion *won*, 4% of the total costs of the 2000 Census, was allocated for public relations.

22. One of the measures to encourage respondents to participate in the census is the construction of respondent-friendly questionnaires. The easier the census form is, the more accurate the census is. The questionnaires of the 2000 Census in Korea will be designed in A4 size with booklet style binding rather than B4 sheet style without the binding as in previous censuses. To expedite respondents' understanding, examples of each item will be placed to the left of the questionnaire to be filled out. These kinds of changes in questionnaire format were proven to improve the accuracy of responses in several pilot surveys.

23. On the other hand, the items of industry and occupation of the census in Korea had been coded up to 3 digits. The coding of industry and occupation is a troublesome task since respondents usually provide insufficient or inappropriate information for the proper classification. According to the results of the fourth pilot survey, 83.6% of industry and 78.6% of occupation were only able to be coded in 3 digits.

24. Special instructions for enumerators on these two items are prepared. Time will be allocated to the explanation of instructions in the training timetable of enumerator for this purpose. In addition, automatic coding of these two items is also being studied to improve coding consistency and accuracy for the first time in the 2000 Census in Korea.

Table 3. Degree of manual classification of industry and occupation  
: in the fourth pilot survey of the 2000 Census, Korea

		1 digit	2 digit	3 digit	4 digit	5 digit
Industry(%)	CC(1)	95.5	93.7	77.0	71.0	64.5
	EC(2)	98.2	96.7	83.6	78.1	71.9
	(2)-(1)	2.7	3.0	6.6	7.1	7.4
Occupation(%)	CC(1)	92.7	91.3	68.8	59.3	53.8
	EC(2)	96.2	95.1	78.6	69.4	63.0
	(2)-(1)	3.5	3.8	9.8	10.1	9.2

Key : a. CC : Classification by the contents in questionnaire only

b. EC : Classification by both the contents in questionnaire and the judgement of coders on their own experience

## More timely

25. The release time of census results has been gradually decreasing. However, the 2000 Census as a census in the Information Age should be provided more timely than other censuses in the past. The KNSO set a goal to release overall data of the 2000 Census not beyond 2001.

26. To reach that goal, a prompt report based on a 2% sample will not be released in the 2000 Census. Instead, 10% of the sample data will be made available nine months earlier. According to the data-processing plan, preliminary counts on population, households and housing by administrative unit based on summary tables are scheduled to be released within this year. Likewise, the report of short-forms accounting for 90% of total households will be available by November 2001 while the report of long-forms carrying 10% of total households by December 2001. Four months and nine months will be advanced forward for the short- and long-forms respectively.

Table 4. Time schedule for data release of the Population and Housing Census in Korea

	1990 Census	1995 Census	2000 Census
Preliminary count	Mar. 1991(4 months)	Mar. 1996(4months)	Dec. 2000(1month)
2% prompt report	Jul. 1991(7 months)	Jul. 1996(7months)	-
Short-form	Dec. 1992(25months)	Feb. 1997(15months)	Nov. 2001(11months)
Long-form	Dec. 1992(25months)	Oct. 1997(22months)	Dec. 2001(13months)

Note : Numbers in parentheses indicate total periods taken from the census date

### More effective

27. The KNSO tries to make effective use of the information technology for more effective and efficient management of census operation. Every staff of the KNSO including 47 regional offices and branches has its own LAN-connected PC which data and documents can be exchanged easily. Thus, the plans and results of various tests and studies on the 2000 Census are shared among all staff members including census planers, census demographers, computer programmers, census managers, field staffs and so on.

28. The homepage of the 2000 Census in Korea has recently opened at <http://www.nso.go.kr/census2000>. This homepage provides not only such basic information as the introduction of census, Q&A, recruitment of enumerators, manuals, glossaries, the results of the 1995 Census but also various forms needed for field operations.

29. In the Dress Rehearsal Survey(DRS) taken in Nov. 1999, 552 enumerators and 37 supervisors were selected from the application forms on this web site. Sixteen offices of *Ups*, *Myons*, *Dongs* reported the rates of completed

enumeration everyday in the form shown on the census homepage during enumeration period. Another benefit of Internet-based management, which was proven in the DRS, is that it allows more quick and accurate data transmission for the preliminary report.

## More informative

30. The items of the 2000 Census in Korea have to be enlarged for several reasons. First of all, conducted in the years ending in 0, the 2000 Census in Korea would be a regular census that carries more items than a simplified census being conducted in the years ending in 5. Furthermore, the fact that the 2000 Census is taking place at the opening of new millenium is pressing on census planners to make the 2000 Census a more informative census.

31. In fact, there have been requests from policy-makers, researchers and other users for including items to meet the growing demand arisen by socioeconomic changes. Census planners are in a dilemma of how to reconcile the usefulness of census data with increasing respondent burden on public. Taking into account the deteriorating environment in census-taking as stated earlier, conservatives hold that questionnaires should be simple and not requiring a great amount of time for completion.

32. However, the KNSO decided to accommodate user's requests to the highest possible extent in order to enhance the utility of the census data. The number of provisional items of the 2000 Census in Korea is 50 as shown in appendix. The items of the 2000 Census will be finalized by May 2000.

Table 5. Number of items in the Population and Housing Census in Korea

	1970	1975	1980	1985	1990	1995	2000
Total	31	31	45	30	44	28	50
Short-form	15	11	26	30	32	17	20
Long-form	16	20	19		12	11	30
Population	17	22	25	16	21	16	29
Household & housing	14	9	20	14	23	12	21

Note : Number of items in the 2000 Census is provisional

33. Looking at new items in the 2000 Census, the following three categories emerge:

- items related to a knowledge-based and information society: major field of study, duration of current work, mobile phone and beeper, use of PC, use of the Internet, communication facilities;
- items related to social welfare for the elderly: place of child living

nearest to person aged 60 and over, supporter for person aged 60 and over, livelihood source of person aged 60 and over, basic activities of daily living of person aged 60 and over;

- items related to the quality of living conditions: number of automobiles, parking place, piped water facilities, type of drinking water, etc.

## **More diverse and convenient access to data**

34. Better products and services of the census should be available to the public as much as possible so that the costs can be well turned into benefits for the society. Thus, the easy and convenient access to the census data should be ready to meet users' needs.

35. Thanks to the rapid development in computer technology, the dissemination of information has been greatly diversified, ranging from printed materials to computer media and web services. In particular, CD-ROMs and the Internet become as emerging media as PCs come into wide use. According to the estimation of National Computerization Agency, there are 20 PCs per 100 persons as of 1998 while 12 PCs per 100 persons in 1995 in Korea. Meanwhile, Internet users also increased from 366,000 in 1995 to 10,860,000 in 1999.

36. To reflect the rapid computerization and informatization of the public, the KNSO will shift weight from the traditional publication to electronic media and the Internet in terms of data dissemination. In principle, services of three types of media, i.e., publication, CD-ROMs and the Internet, will be provided simultaneously to the public. While publications provide only tables, CD-ROMs and the Internet offer tables in the form of spread sheets together with metadata. Also, on-demand tables as well as individual data will be provided by CD-ROMs.

37. Also, the use of GIS will enhance the usefulness of census data by combining geographic information. When the study on Statistical Districts, in which their boundaries are demarcated by permanent geographic objectives, is finalized, a variety of GIS-based small area statistics will be brought out. Map-based social atlases will also be created.

## ***publications***

38. Even though the role of printed materials as a means of dissemination of information is not what it used to be, publications are still undoubtedly the most important type of media. The KNSO has a plan to increase the number of tables in the publications up to around 300 compared with 152 in the 1995 Census, for the benefit of those who don't own a PC or are not familiar with

computers.

39. Publications for special topics like one-person households, quality of living conditions, social welfare and knowledge-based society will be newly added to the list of publications in the 2000 Census. According to this plan, the total number of volumes will reach 30.

Table 6. Publications of the census reports

	1985	1990	1995
Number of tables	69	124	152
Short-form	48	50	65
Long-form	21	74	87
Number of volumes	17(18)	20(20)	22(23)
Whole country	1 (1)	1 (1)	1 (2)
Provinces	13(13)	13(13)	15(15)
Special topics	3 (4)	6 (6)	6 (6)

Note : 1. Numbers in parentheses indicate the number of books

2. Special topics cover the origin of family name, fertility, and migration in 1985; commuters and students, fertility, migration, economic activity, industry and occupation, and the elderly in 1990; commuters and students, migration, economic activity, industry and occupation, the elderly, and rent in 1995.

### **computer media**

40. A series, consisting of 22 volumes, was condensed into a single CD-ROM in 1997. Since tables are recorded in EXCEL format, it is serviceable for researchers to compile their own progress and produce data files for further statistical analysis with SAS or SPSS. In addition, it was designed to provide users with some additional functions such as simple calculation and graphing, etc.

41. The production of CD-ROMs was of benefit to both producer and user in terms of cost. The cost of 1,650 CD-ROMs was less than 50 million *won* which was one-seventh of that for the 1,500 series, which consisted of 22,650 books. Therefore, a CD-ROM sells for 25,000 *won* whereas the price of a series of printed publications is 320,000 *won*. This explains why the quantity of CD-ROMs should be increased.

42. CD-ROMs are also used to provide on-demand tabulation, which it has been requested around 30 times every year in case of the 1995 Census. Besides, individual data based on 2% and 10% sample data after eliminating the identifiable data of a confidential nature are also provided in CD-ROMs on a sale basis.

### **Web services**

43. The KNSO established the STAT-KOREA at <http://www.stat.go.kr> in Nov. 1999, which provides one-stop service of statistical information covering 401 kinds of statistics produced by 123 agencies in Korea. Since statistical agencies use the standardized DB software provided by the KNSO, users can navigate through statistical data more easily.

44. Before the advent of the STAT-KOREA, statistical data were provided to the public through the Korea Statistical Information System(KOSIS), which contains 5.3 million series of statistical data. However, it requires a Telnet emulator in order to gain access to this database. In this Telnet approach, the census data were retrieved 32,081 times, which accounts for 7.8% of the total visits to the KOSIS in 1999. Incorporated into the STAT-KOREA, the census data can be easily combined by users with other databases in this data warehouse.

Table 7. Number of visits to the KOSIS

	Total	Population and Housing Census					
		Sub-total	Pop.	Household	Housing	Economic activity	Others
January	23,398	3,181	2,014	248	340	429	150
February	18,950	2,353	1,504	144	243	357	105
March	35,404	3,703	2,052	281	428	631	311
April	38,706	3,515	1,509	291	539	753	423
May	44,292	3,254	1,491	268	402	730	363
June	47,307	3,337	1,483	302	567	638	347
July	33,993	2,715	1,237	261	456	493	268
August	36,283	2,637	1,203	277	376	528	253
September	39,749	2,594	1,147	297	389	560	201
October	42,353	2,732	1,253	265	345	562	307
November	28,645	1,368	595	140	186	323	124
December	17,588	692	285	73	122	131	81
Total (%)	406,668 (100.0)	32,081 (7.9)	15,773	2,847	4,393	6,135	2,933

## Concluding remarks

45. The 2000 Census in Korea should be carried out in ways to cope with not only the deteriorating environment of census-taking but also the growing demand for higher quality products and services of the results. The KNSO endeavors to make the 2000 Census *more economic, more accurate, more timely, more effective, more informative, and more convenient*. It is widely believed that only a successful 2000 Census will secure the census in 2005 and onwards in Korea.

<Appendix> Provisional items for the 2000 Census, Korea

	Total	Contents
--	-------	----------

Items of the 1995 Census which will carry on in the 2000 Census	28	<p>&lt;short-form&gt;  1) name, 2) relationship to the head of household, 3) sex, 4) age, 5) educational attainment, 6) place of birth, 7) marital status, 8) type of household, 9) number of rooms used, 10) type of kitchen, toilet and bathroom, 11) type of tenure, 12) type of living quarters, 13) total area of floor space, 14) total area of housing site, 15) total number of rooms, 16) period of construction, 17) number of housing facilities installed</p> <p>&lt;long-form&gt;  18) place of residence 5 years ago, 19) commuting status, 20) place of work or school, 21) mode of transport to work or school, 22) travel time to the place of work or school, 23) economic activity status, 24) employment status, 25) industry, 26) occupation, 27) duration of occupancy, 28) rent</p>
Items of the 1995 Census which will be dropped in the 2000 Census	1	<p>&lt;short-form&gt;  1) religion</p>
Items added newly to the 2000 Census	22	<p>&lt;short-form&gt;  1) origin of family name, 2) type of detached house, 3) type of occupancy</p> <p>&lt;long-form&gt;  4) major field of study, 5) way of caring for a child, 6) place of residence 1 year ago, 7) use of PC, 8) use of the Internet, 9) mobile phone and beeper, 10) duration of current work, 11) children ever born, 12) place of child living nearest to person aged 60 and over, 13) livelihood source of person aged 60 and over, 14) supporter for person aged 60 and over, 15) basic activities of daily living of person aged 60 and over, 16) fuel used for cooking, 17) type of heating, 18) piped water facilities, 19) type of drinking water, 20) communication facilities, 21) number of automobiles, 22) parking place</p>