United Nations International Seminar
on Population and Housing Censuses:
Beyond the 2010 Round
27-29 November 2012
Seoul, Republic of Korea

SESSION 7: Use of modern technologies for censuses

FACTORS AFFECTING THE DECISION TO PARTICIPATE IN THE INTERNET OPTION FOR THE 2010 CENSUS KOREA

Prepared by

Mr. Eunkoo Lee
Mr. Sungjin Kim
Population Census Division
Statistics Korea

* This document is being reproduced without formal editing.

Factors affecting the decision to participate in the internet option for the 2010 Census of Korea

Eunkoo Lee, Sungjin Kim
Population Census Division
Statistics Korea

Background

The Population and Housing Census of Korea is conducted every five years. In the 2010 Census, respondents were offered an internet option (online response system) to complete the form. Extensive developmental efforts were made to ensure confidentiality and reliability of the online response system. The performance of the 2010 online response system was a huge success. 47.9% of respondent completed their census forms via internet (47.7% and 49.3% for the short and long form respectively). For the 2010 Census round, Korea's internet uptake rate is only second to Canada (54.4%). The success of the internet option in the 2010 Census can be attributed to several factors. The high level of internet penetration in Korea supported by advanced and modern IT infrastructure could be a potential factor. Active promotion of the internet option could be another factor. Statistics Korea worked closely with local governments to promote the internet option. Census advertising campaigns were launched via TV, radio, and etc. This presentation attempts to investigate factors that characterize the respondents who have participated in the internet option for the 2010 Census Korea

Objective

The objective of this presentation is to identify factors that have significantly affected the respondents' decision to complete the census form via internet.

Methods

The 2010 Census questions from the long form (independent variables) were selected to build a logistic regression model. The response variable is binary (0, 1). It is coded 1 if the household participated in the internet option, otherwise it is coded 0. 1,795,984 households who completed the long form in the 2010 Census were selected for this analysis. First, the logit model was built for variables related to households. Then the second logit model was built for variables related to the head of household. The independent variables for the logit model from the perspectives of the household and the head of the household are shown respectively in Table 1 and 2 below. The accessibility of internet line was controlled in the both models to eliminate possible confounding effects.

Results

The results of the fitted logistic regression model from the perspectives of the household suggest the following (refer to table 3 and table 4 for the details of statistical analysis):

- The odds of households in the rural area (versus in the urban area) to participate in the internet option increase by a factor of 1.193

- The odds of one person households (versus one family household) to participate in the internet option increase by a factor of 1.294
- The odds of households with owned or leased residence type residence type (versus monthly rent paid in advance) to participate in the internet option increase by a factor of 1.516 and 1.335 respectively.
- The odds of households with apartment type (versus detached house) to participate in the internet option increase by a factor of 1.510
- The odds of households with internet line (versus without internet line) to participate in the internet option increase by a factor of 12.835
- The odds of households which have children attending elementary, middle and high school are more likely (versus without children attending elementary, middle and high school) to participate in the internet option by a factor of 1.404

The results of the fitted logistic regression model from the perspective of the head of the household suggest the following: To be more precise, the results show the characteristics of the head of the household from the households which have participated in the internet option.

- The odds of households where the head of household have no formal schooling or have graduated elementary school are more likely (versus high school graduates) to participate in the internet option by a factor of 1.770, 1.264 respectively.
- The odds of households where the head of household have university, master and doctorate education are more likely (versus high school graduates) to participate in the internet option by a factor of 1.262, 1.609, 1.684 respectively.
- The odds of households where the head of household actively engage in social activities are more likely to participate in the internet option by a factor of 1.098
- The odds of households where the head of household has been working during the past week of the census date are more likely (versus no work during the past week of the census date) to participate in the internet option by a factor of 1.373
- The odds of households where the head of household has been married or widowed or divorced are more likely (versus never married) to participate in the internet option by a factor of 1.686, 1.537 and 1.194 respectively.
- The odds of households where the head of household's occupation are managers, professionals, and other white collar office workers are more likely (versus agricultural, forestry & fishery workers) to participate in the internet option by a factor of 1.829, 1.591 and 1.549 respectively.

Discussion and conclusion

Using the 2010 census data, we have attempted to find characteristics of the households that have participated in the internet option. Our logistic regression models suggest that households in the rural area are more likely to participate in the internet option which is counterintuitive at first glance. One possible explanation could be the high concentration of one person households with older people in the rural area. Such households lack internet accessibility and they often times need help in completing the census form. In 2010 census, the local governments were active in reaching out to such households in the rural area where they encouraged them to come to the community center to complete the census form via internet with the help of enumerators. It is also expected that people living in the apartments are more likely to participate in the internet option in relative to detached houses since it is

more convenient for internet providers to provide large scale internet access services to people living in the apartments. Another interesting observation is that households with children attending elementary to high school were more likely to participate in the internet option. In Korea, it is mandatory for elementary school students, middle school students and high school students to satisfy certain quota for volunteer community services at school. In 2010 Census, volunteer service points were given to students to encourage these households to participate in the internet option. The results show that such incentives are proven to be effective.

As for the characteristics of the head of the household from the households which have participated in the internet option, the head of household with low education level (no schooling and elementary school graduates) showed high level of internet participation in relative to high school graduates. It can be explained that those with low education level are older in age and are more likely to reside in rural areas where they need help from others to complete the census form. This is in agreement with previous observation where the rural area residents were more likely to participate in the internet option in relative to their urban counterparts. On the contrary, those with higher education level (versus high school graduates) were more likely to participate in the internet option. As for the occupation of the head of household, managers, professionals and other white collar office workers were more likely to participate in the internet option in relative to household heads working in the agriculture & fishery sectors.

This presentation discussed potential factors that might influence the respondents' decision to participate in the internet option. The results suggest which specific households should be targeted for internet option based on socioeconomic backgrounds, education, housing type, regions, etc. By identifying such factors, planners of the next Census could utilize the results of this analysis for planning and improving the internet option in the next 2015 Census.

Table 1: Independent variables for the logit model from the perspective of the household

Independent variables	Code	Descriptions
Urban/Rural	0	Urban
	1	Rural
Type of households	1	One-family household
	2	Household consisting of a family and a non-family members
	3	One person household
	4	Household of 5 or less persons who have no blood ties
	5	Household of 6 or more persons who have no blood ties
Type of residence	1	Owned
	2	Lease(no monthly rent)
	3	Monthly rent
	4	Monthly rent paid in advance
	5	Free(official residence, private residence, relative's house)
Type of housing	1	Detached house
	2	Apartment
	3	Row house
	4	Apartment unit in a private house
	5	House within commercial building
	6	Officetel
	7	Hotel, inns and other lodging facility
	8	Dormitory and welfare institution
	9	Shack, vinyl greenhouse
	10	Others
Internet line	0	Available
	1	Not available
Households with students attending elementary, middle or high schools	0	Yes
	1	No

Table 2: Independent variables for the logit model from the perspective of the head of household

Independent variables	Code	Descriptions
Age		Figures
Sex	1	Female
	2	Male
Education	1	No schooling
	2	Elementary school
	3	Middle school
	4	High school
	5	Junior college
	6	University
	7	Master's course
	8	Doctor's course
Social activity participation	1~9	Score
Economic activity (Any work for either pay or profit during the last week)	1	Full time
	2	Part time
	3	Had been working, but was being unemployed for a while
	4	Did not work during the last week
Marital status	1	Never married
	2	Married
	3	Widowed
	4	Divorced
Occupation	1	Managers
	2	Professionals & related workers
	3	Clerks
	4	Service wokers
	5	Sales workers
	6	Skilled agricultural, forestry & fishery workers
	7	Crafts & related trades workers
	8	Plant, machine operators & assemblers
	9	Elementary occupations
	10	Others
Internet line	0	Available
	1	Not available

Table 3: Statistical analysis of the logit model from the perspective of the household

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-2.4615	0.0172	20396.1643	<.0001
Urban/Rural	**1**	1	0.1765	0.00461	1463.5206	<.0001
Type of households	**2**	1	-0.1184	0.0261	20.5915	<.0001
	3	1	0.2576	0.00491	2755.3975	<.0001
	4	1	-0.1923	0.0176	119.0432	<.0001
Type of residence	**1**	1	0.4164	0.0166	626.4551	<.0001
	2	1	0.2893	0.0169	292.8466	<.0001
	3	1	-0.0753	0.0168	20.0415	<.0001
	5	1	-0.0933	0.0183	25.9118	<.0001
Type of housing	**2**	1	0.412	0.00419	9667.5429	<.0001
	3	1	0.0508	0.0108	22.0027	<.0001
	4	1	0.0264	0.0075	12.3487	0.0004
	5	1	-0.1401	0.0146	92.6814	<.0001
	6	1	0.1503	0.0163	85.3966	<.0001
	7	1	0.00265	0.066	0.0016	0.9679
	8	1	0.3136	0.0439	51.1184	<.0001
	9	1	0.0606	0.05	1.4732	0.2248
	10	1	0.3381	0.0252	180.2531	<.0001
Internet line	**0**	1	2.5522	0.0049	270781.551	<.0001
Households with students attending elementary, middle or high schools	**0**	1	0.3394	0.00433	6150.1724	<.0001

Odds Ratio Estimates				
Effect		**Point Estimate**	**95% Wald Confidence Limits**	
Rural vs urban		1.193	1.182	1.204
Type of households	**2 vs 1**	0.888	0.844	0.935
	3 vs 1	1.294	1.281	1.306
	4 vs 1	0.825	0.797	0.854
Type of residence	**1 vs 4**	1.516	1.468	1.567
	2 vs 4	1.335	1.292	1.38
	3 vs 4	0.927	0.897	0.959
	5 vs 4	0.911	0.879	0.944
Type of housing	**2 vs 1**	1.51	1.498	1.522
	3 vs 1	1.052	1.03	1.075
	4 vs 1	1.027	1.012	1.042
	5 vs 1	0.869	0.845	0.894
	6 vs 1	1.162	1.126	1.2
	7 vs 1	1.003	0.881	1.141
	8 vs 1	1.368	1.256	1.491
	9 vs 1	1.063	0.963	1.172
	10 vs1	1.402	1.335	1.473
Internet line	**0 vs 1**	12.835	12.713	12.959
Households with students attending elementary, middle or high schools	**0 vs 1**	1.404	1.392	1.416

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	551188	19	<.0001
Score	502191.1	19	<.0001
Wald	393592.2	19	<.0001

Table 4: Statistical analysis of the logit model from the perspective of the head of household

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-3.2492	0.0302	11562.265	<.0001
Age		1	0.00353	0.000196	322.4428	<.0001
Sex	**1**	1	0.0501	0.00533	88.5739	<.0001
Education	**1**	1	0.571	0.0101	3201.3349	<.0001
	2	1	0.2344	0.00699	1125.2611	<.0001
	3	1	-0.2076	0.00657	997.016	<.0001
	5	1	0.3371	0.00629	2876.0854	<.0001
	6	1	0.2326	0.00548	1800.9212	<.0001
	7	1	0.4757	0.0104	2075.0154	<.0001
	8	1	0.5209	0.0181	828.8601	<.0001
Social activity participation		1	0.0931	0.00246	1435.3094	<.0001
Economic activity	**1**	1	0.3173	0.0265	143.4575	<.0001
	2	1	0.7359	0.0298	609.1941	<.0001
	3	1	1.0762	0.0302	1267.5649	<.0001
Marital status	**2**	1	0.5226	0.00631	6858.2394	<.0001
	3	1	0.43	0.0094	2094.1505	<.0001
	4	1	0.1769	0.00885	399.5691	<.0001
Occupation	**1**	1	0.6039	0.014	1852.5285	<.0001
	2	1	0.4645	0.00926	2517.2644	<.0001
	3	1	0.4377	0.00916	2283.9302	<.0001
	4	1	-0.1235	0.0102	146.9881	<.0001
	5	1	-0.0914	0.00938	94.7816	<.0001
	7	1	0.0693	0.00928	55.7637	<.0001
	8	1	0.1503	0.00872	296.6111	<.0001
	9	1	-0.0763	0.00947	64.8977	<.0001
	10	1	0.6094	0.0271	506.9099	<.0001
Internet line	**0**	1	2.6731	0.00533	251852.551	<.0001

Odds Ratio Estimates				
Effect		**Point Estimate**	**95% Wald Confidence Limits**	
Age		1.004	1.003	1.004
Sex	**1 vs 2**	1.051	1.041	1.062
Education	**1 vs 4**	1.77	1.735	1.805
	2 vs 4	1.264	1.247	1.282
	3 vs 4	0.813	0.802	0.823
	5 vs 4	1.401	1.384	1.418
	6 vs 4	1.262	1.248	1.275
	7 vs 4	1.609	1.576	1.642
	8 vs 4	1.684	1.625	1.744
Social activity Participation		1.098	1.092	1.103
Economic activity	**1 vs 4**	1.373	1.304	1.447
	2 vs 4	2.087	1.969	2.213
	3 vs 4	2.933	2.765	3.112
Marital status	**2 vs 1**	1.686	1.666	1.707
	3 vs 1	1.537	1.509	1.566
	4 vs 1	1.194	1.173	1.214
Occupation	**1 vs 6**	1.829	1.78	1.88
	2 vs 6	1.591	1.563	1.62
	3 vs 6	1.549	1.522	1.577
	4 vs 6	0.884	0.866	0.902
	5 vs 6	0.913	0.896	0.93
	7 vs 6	1.072	1.052	1.091
	8 vs 6	1.162	1.142	1.182
	9 vs 6	0.927	0.909	0.944
	10 vs 6	1.839	1.744	1.94
Internet line	**0 vs 1**	14.485	14.335	14.637

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	564996.4	26	<.0001
Score	510480.8	26	<.0001
Wald	392240.3	26	<.0001