

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

Eunkoo Lee

Population Census Division, Statistics Korea, [ex1155@naver.com](mailto:ex1155@naver.com)

Sungjin Kim\*

Population Census Division, Statistics Korea, [saintgenie@naver.com](mailto:saintgenie@naver.com)

### Abstracts

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. The performance of the 2010 online response system was a huge success. 47.9% of respondent completed their census forms via internet. 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 paper attempts to investigate factors that characterize the respondents who have participated in the internet option for the 2010 Census Korea. 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. About 1.79 million households who completed the long form in the 2010 Census were selected for this analysis. The first logit model was built for variables related to households. Then the second logit model was built for variables related to the head of household. This presentation would discuss 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.

Key words : population and housing census, logistic regression model, response rate of internet option.

### 1. 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

## 2. Objective

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

## 3. 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.

## 4. 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)

- 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.

## 5. 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 workers
	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: Odds Ratio Estimates of the logit model from the perspective of the household

Comparison	Most likely to participate in the internet option	Odds Ratio
Rural VS Urban	Rural	1.193
One person household VS One family household	One person	1.294
Owned or leased residence type VS monthly rent paid in advance	Owned or leased residence type	1.516, 1.335 respectively
Apartment VS Detached house type	Apartment	1.510
Household with internet line VS Household without internet line	Household with internet line	12.835
Household with students attending elementary to high school VS otherwise	Household with students attending elementary to high school	1.404

Table 4: Odds Ratio Estimates of the logit model from the perspective of the head of household

Comparison	Most likely to participate in the internet option	Odds Ratio
No school & elementary school VS High school	No school & elementary school	1.770, 1.264 respectively
University, master, Ph.D VS High school	University, master, Ph.D	1.262, 1.609, 1.684 respectively
Work VS No work	Work	1.373
Married, widowed, divorce VS Never married	Married, widowed, divorce	1.686, 1.537, 1.194 respectively
White collar office workers VS agricultural, forestry & fishery workers	White collar office workers	1.829, 1.591, 1.549 respectively