# Norwegian citizen panel

## 2013, first wave

## Methodology report

Asle Høgestøl

Øivind Skjervheim

January, 2013

NORSK MEDBORGERPANEL

IDEAS2EVIDENCE
Bygger kunnskap

# NORWEGIAN CITIZEN PANEL – 2013, FIRST WAVE – METHODOLOGY REPORT

## BACKGROUND

This report describes the stages in the process of establishing the Norwegian Citizen Panel, the practical procedures for recruitment of panel members and the first wave of data collection. Furthermore, the report discusses the representativeness of the panel and the designing of weights.

The Norwegian Citizen Panel has been established as a cooperation between the University of Bergen, several institutes at the Faculty of Social Sciences at the University of Bergen and UNI Rokkansenteret.

ideas2evidence are responsible for the recruitment and administration of panel members, as well as the technical solutions regarding data collection and computing.

## BUILDING THE SURVEY

Panel recruitment and the first wave of data collection were conducted simultaneously. Members were recruited per post. A representative sample of 25 000 individuals received an invitation with information about the project and registration instructions. The invitees were asked to answer the survey and provide their e-mail address for all further communication. Detailed information on the survey and the panel recruitment will be given in a later section of this report.

The survey and the panel administration are conducted by use of the web-based research software Confirmit. Confirmit is a so-called "Software-as-a-Service" solution, where all software runs on Confirmit's continuously monitored server park and where survey respondents and developers interact with the system through various web-based interfaces. This ensures high data security and operational stability. The security measures are the most stringent in the industry and Confirmit guarantees 99.7 percent uptime. Ideas2evidence are responsible for implementing the survey in Confirmit on behalf of The Citizen Panel.

The survey was pilot tested twice prior to publishing. The first pilot test was carried out among a group of twelve master students. The test focused mainly on survey's substance and the clarity of the questions. The second pilot test was conducted among 112 bachelor students attending compulsory seminars in methodology. The pilot testing was successful and no major revisions were deemed necessary.

## DRAWING THE SAMPLE

Based on international literature the panel recruitment process was expected to have a success rate of approximately 14 percent[1]. In order to reach our minimum target of 3 500 respondents we therefore had to draw a gross sample of 25 000 individuals.

The sample was drawn from the Norwegian National Population Registry. This register includes everyone born in Norway as well as former and current inhabitants. The Norwegian Tax Administration is responsible for the register, but the administration is partly outsourced to the private IT-company Evry. Evry drew the sample on behalf of the Citizen Panel after the necessary permissions were acquired from the Norwegian Tax Administration.

[1] Rao, Kumar, Olena Kominska and Allan L. McCutcheon: *Recruiting probability samples for a multi-mode research panel with internet and mail components.* Public Opinion Quarterly, Vol. 74, Spring 2010, pp. 68-84

25 000 individuals over the age of 18 were randomly drawn from the register. The extracted data was a) last name, b) first name, c) address, d) gender, and e) age. The sample excluded people with no current home address in Norway.

Furthermore, all individuals over the age of 95 were excluded from the sample. This amounted to 58 persons, thus leaving a net sample of 24 942 individuals before the recruitment process started.

## RECRUITING THE MEMBERS

Panel members were recruited per post in two steps.

First, letters were sent to everyone in the sample. The letters contained the following information; a) a description of the project, b) the Citizen Panel's policy on privacy and measures taken to protect the anonymity of the participators, c) the time-frame of the project, d) the participants' rights to opt out of the panel at any time in the future, e) contact information for the people responsible for the project, f) a unique log-in id and the web address to the panel's web site and g) the estimated time required to complete the survey (20 minutes).

In order to maximize the response rate, an incentive in the form of a travel gift card was included in the project. The value of the gift card was 25 000 NOK. To enter the lottery respondents were required to join the panel and provide their e-mail addresses. Respondents were asked to register on the panel's web site and log into the survey using the unique id-code provided in their personal letter. Information on the lottery was included in all correspondence with respondents.

The letter was posted on the 6th of November 2013.

A reminder post card was sent on the 22nd of November 2013 to those respondents who a) had not logged into the survey, or b) had neither completed the survey nor provided their e-mail address. Respondents were encouraged to join the panel, with reference to the letter sent two weeks prior. The unique log-in id provided in the original letter was also included in the post card.

A reminder e-mail was sent on the 25th of November to those respondents who had provided their e-mail addresses, but had not yet completed the survey.

Figure 1: Completion rate per day

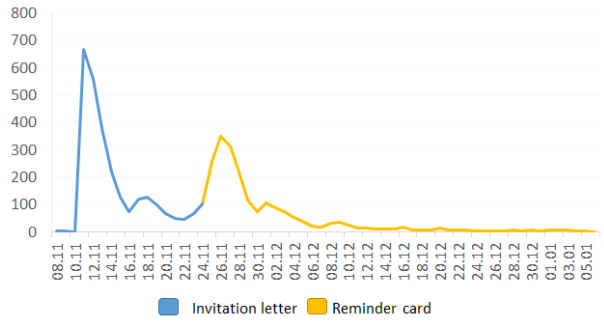
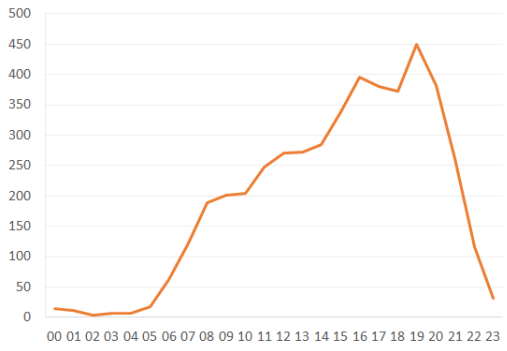



Figure 2: Completion rate per hour




figures 1 and 2 illustrate the daily and hourly respondent completion rate. As figure 1 shows, actual data collection starts on the 11th of November, five days after invitations were posted. 667 respondents completed the survey that day. The completion rate declines steeply throughout the week, before peaking again during the weekend (16th and 17th of November). The reminder card was sent on the 22nd of November and the next peak occurs on the following Monday (25th of November). In other words, response rates were highest on the days when respondents received the invitation and reminder. Many respondents also seem to prefer giving their response during the weekend.

Figure 2 shows that the hourly completion rate increases throughout the day. Approximately 50 percent of the respondents completed the survey between 16:00 and 21:00.

## SURVEY RESPONDENTS AND PANEL MEMBERS

A distinction has to be made between panel members and survey respondents. Panel members are defined as those respondents who enter their e-mail addresses, regardless of whether they choose to complete the questionnaire or not. Survey respondents are defined as respondents who complete a substantial share of the questionnaire.

Of the 24 396 individuals who received letters from the Citizen Survey (546 letters were returned), 21.1 percent (5 163 people) logged into the survey. 4 636 respondents completed the questionnaire and 526 respondents exited the questionnaire before completion. However, 51 percent of the non-complete responses are saved as survey data. The remaining 257 incomplete responses are excluded from the survey, due to lack of data. In consequence, the first wave of the Norwegian Citizen Survey has 4 905 **survey respondents**. This equals a survey response rate of 20.1 percent.

98 percent of the respondents who completed the survey entered their e-mail addresses. Of the total of 526 incomplete responses, 58 percent entered their e-mail addresses. This adds up to a total of 4 870 **panel members** in the Norwegian Citizen Panel, equaling a panel recruitment rate of 20 percent.

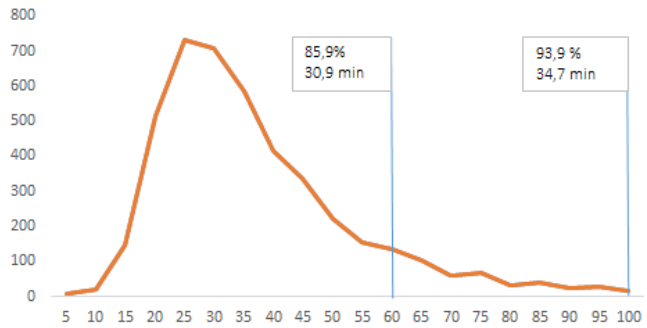
As these figures suggest, there is little reluctance among respondents to providing their e-mail addresses. 94 percent of the panel members who completed the survey, entered their e-mail addresses at the very beginning of the questionnaire. An additional 4 percent provided the address when asked a second time at the very end of the questionnaire.

## NON-RESPONSE

In the large N pilot test, respondents spent an average of 22.3 minutes on the questionnaire. In comparison, survey respondents who completed the questionnaire spent an average of 31 minutes.

The survey administration software allows for the questionnaire to be completed in several stages. This creates an artificially high average for the time used to complete the survey, as respondents may leave the questionnaire while still being logged on, in order to complete the survey later. The calculated average of 31 minutes, therefore, only includes the 86 percent of the respondents who spent 60 minutes or less to complete the survey.

Figure 3: Survey respondents time usage




The Norwegian Citizen Survey has a high completion rate: 90 percent (4636/5163) of the individuals that accessed the survey, completed it. Considering the high average time use, this is a very satisfying result. If time

use was a source of frustration for the respondents, we would expect the share of incomplete responses to be higher.

Moreover, the 257 respondents who were excluded from the final net sample exited the survey early, indicating that the exit was not caused by fatigue. Only 97 respondents left the questionnaire when asked to enter their e-mail address at the very beginning of the questionnaire.

The loss of respondents who have accessed the survey is at a minimum. The main barrier therefore lies in the initial process of accessing the survey. As long as the survey captures the interest of the respondent, once he/she has accessed the survey the loss will be minimal.

## REPRESENTATIVENESS

This section describes the representativeness of the net sample of survey respondents. Since there is an almost perfect overlap between survey respondents and panel members, the descriptions are also valid for the panel.

There are two main challenges related to non-response and representativeness:

- access to and familiarity with the internet given that the web-based questionnaire was the only available response mode
- the motivation and interest of the respondents

The first challenge is strongly related to the age composition of the survey respondents. Although Norway has a very high computer and internet density, the probability of having an e-mail address and the skills required to access and fill in an on-line questionnaire will normally decrease with increasing age. The second challenge, motivation and interest, is often explained by the respondents' level of education. In addition to age and education, variables of geography and gender are applied to test the representativeness of the survey respondents. The variables have the following categories:

- Age: 19-29 years, 30-59 years, 60 and above
- Highest completed education: no education/elementary school, upper secondary, university/university college.
- Geography: Oslo/Akershus, Eastern Norway, Southern Norway, Western Norway, Trøndelag, Northern Norway.

The sampling frame of the survey is Norwegians above 18, comprising a population of approximately 3.9 million individuals. The distribution of men and women in the population is 49.9 percent and 50.1 percent respectively. The distribution of men and women in the net sample is 50.7 percent men, 49.3 percent women. In other words, men are slightly overrepresented.

*Table 1: Age distribution in the population and the net sample*

| | 18-29 years | 30-59 years | 60 years and above |
|---|---|---|---|
| Population | 20,3 % | 52,1 % | 27,6 % |
| Net sample | 18,0 % | 57,1 % | 24,9 % |

The age distribution in table 1 shows that the age groups of 18-29 years and 60 years and above are underrepresented. There are two different explanations for this bias. The underrepresentation of people above 60 years can mainly be explained by the before mentioned challenge of access to and familiarity with the internet. In addition, the eldest segment of this group will normally have lower response rates in any survey independent of mode. When it comes to the age group 18-29 years, the underrepresentation can partly be explained by the fact that a substantial part of this group is students. For many students the registered mail address often differs from the actual address of residence and it will take time before the mail is redirected to

the correct address. This is also an age group who normally has a busy life with little time to set aside for answering surveys.

*Table 2: Combined distribution of age and gender in the population and the net sample*

| | 18-29 years | | 30-59 years | | 60 years and above | |
|---|---|---|---|---|---|---|
| | Men | Women | Men | Women | Men | Women |
| Population | 10,4 % | 9,9 % | 26,7 % | 25,4 % | 12,8 % | 14,8 % |
| Net sample | 8,5 % | 9,5 % | 28,4 % | 28,8 % | 13,8 % | 11,0 % |

Table 2 shows the combined effects of age and gender on response rates. Compared to the population, the net sample has a slight underrepresentation of men in the youngest age group. Men above 30 years, on the other hand, are slightly overrepresented. For women, the age group 30-59 is overrepresented and the group above 60 overrepresented. Overall, when assessing the dimensions of gender and age, the composition of the net sample corresponds quite well with the composition of the population.

*Table 3: Combined distribution of age, gender and education in the population and the net sample*

| | | Population | | Net sample | |
|---|---|---|---|---|---|
| | | Men | Women | Men | Women |
| No education/elementary school | 18-29 years | 4,5 % | 3,6 % | 1,0 % | 1,5 % |
| Upper secondary education | | 4,0 % | 3,3 % | 4,9 % | 4,3 % |
| University/university college | | 1,9 % | 3,0 % | 2,7 % | 3,7 % |
| No education/elementary school | 30-59 years | 6,4 % | 5,5 % | 2,2 % | 1,8 % |
| Upper secondary education | | 11,9 % | 9,2 % | 10,4 % | 8,3 % |
| University/university college | | 8,4 % | 10,7 % | 15,8 % | 18,7 % |
| No education/elementary school | 60 and above | 3,4 % | 5,2 % | 2,2 % | 2,7 % |
| Upper secondary education | | 6,2 % | 7,0 % | 4,4 % | 3,2 % |
| University/university college | | 3,1 % | 2,6 % | 7,3 % | 5,2 % |

More substantial levels of bias appear when adding education to the equation. Table 3 reveals a systematic underrepresentation of individuals with no education above elementary school. Persons belonging to the two lowest educational groups are systematically underrepresented independent of gender and age. The underrepresentation is particularly strong for young men. As expected, individuals with education from universities or university colleges are systematically overrepresented across all demographic segments.

When it comes to geography (table 4 below) we observe a slight underrepresentation of eastern Norway and northern Norway, and a corresponding overrepresentation of the capital area – the counties of Oslo and Akershus.[2] Young men and women in northern Norway are markedly underrepresented, as are older men and women in the same region. Older women are generally underrepresented throughout the country, except in Oslo and Akershus. The same is true for young men.

Middle-aged men are overrepresented in Oslo/Akershus and western Norway, and slightly underrepresented in the rest of the country. Middle-aged women are generally overrepresented, except in Trøndelag, where they are slightly underrepresented.

[2] A test with smaller geographical units shows that the counties of Sogn og Fjordane, Møre og Romsdal, Nord-Trøndelag, Finmark, Troms and Hedmark are especially underrepresented. While the regions Oslo, Akershus, Hordaland and Rogaland are clearly overrepresented. Most of the underrepresented counties are thus located in the periphery.

*Table 4: Combined distribution of age, gender and geography in the population and the net sample*

| | | Population | | Net sample | |
|---|---|---|---|---|---|
| | | Men | Women | Men | Women |
| Akershus/Oslo | 18-29 years | 2,5 % | 2,6 % | 2,4 % | 2,8 % |
| | 30-59 years | 6,7 % | 6,4 % | 7,7 % | 8,3 % |
| | 60 and above | 2,5 % | 3,0 % | 3,1 % | 3,1 % |
| Eastern Norway | 18-29 years | 2,5 % | 2,4 % | 1,8 % | 2,2 % |
| | 30-59 years | 6,9 % | 6,8 % | 6,2 % | 7,3 % |
| | 60 and above | 3,8 % | 4,5 % | 4,2 % | 2,8 % |
| Southern Norway | 18-29 years | 0,6 % | 0,6 % | 0,5 % | 0,5 % |
| | 30-59 years | 1,5 % | 1,4 % | 1,4 % | 1,6 % |
| | 60 and above | 0,7 % | 0,8 % | 0,8 % | 0,6 % |
| Western Norway | 18-29 years | 2,8 % | 2,6 % | 2,6 % | 2,7 % |
| | 30-59 years | 7,0 % | 6,4 % | 8,1 % | 7,4 % |
| | 60 and above | 3,2 % | 3,7 % | 3,6 % | 2,8 % |
| Trøndelag | 18-29 years | 1,0 % | 0,9 % | 0,9 % | 0,8 % |
| | 30-59 years | 2,2 % | 2,1 % | 2,2 % | 2,1 % |
| | 60 and above | 1,1 % | 1,3 % | 1,1 % | 0,8 % |
| Northern Norway | 18-29 years | 1,0 % | 0,9 % | 0,7 % | 0,6 % |
| | 30-59 years | 2,4 % | 2,3 % | 2,3 % | 2,5 % |
| | 60 and above | 1,3 % | 1,5 % | 0,9 % | 0,7 % |

## WEIGHTING OF THE DATA

To compensate for the observed bias, a set of weights has been calculated. The weights equal the relation between a given strata in the population and the total population, divided by the relation between a given strata in the net sample and the total net sample.[3] This procedure returns values around 1, but above 0. Respondents who are underrepresented will receive a weight above 1 and respondents who are overrepresented a weight below 1. The weights of the different stratums are listed in table 8 in the appendix.

When calculating the weights, the information regarding the respondent's geographical location, gender and age are based on registry data. These attributes were included in the sample file we received from the Norwegian Population Register. Information regarding the level of education is provided by the respondents when answering the questionnaire. Approximately 9 percent of the net sample did not answer this question. Because of this, two different weights have been calculated:

- Weight 1 based on demographic variables (age, gender and geography)
- Weight 2 combining the demographic variables with education. Respondents with missing data on the education variable are only weighted on demography (the education component of the weight is set to 1 in these cases).

When applied, both weights will provide a weighted N equal to the number of cases in the dataset.

[3] The applied formula for weight $w_i$ for element $i$, in strata $h$ is: $w_i = \frac{N_h/N}{n_h/n}$

We will strongly recommend using weight 2 in any statistical analysis, as this weight provides the most accurate compensation for the various sources of bias in the net sample. An illustration of this is provided in table 5 which shows the effect of weight 2 on the distribution of self-reported level of education in the net sample:

*Table 5: Effect of weight 2 on self-reported level of education*

| | Sample - not weighted | Sample - weighted | Population | Difference between sample and population | Difference between weighted sample and population |
|---|---|---|---|---|---|
| No education/elementary school | 11,3 % | 28,8 % | 28,7 % | -17,4 % | 0,1 % |
| Upper secondary education | 35,4 % | 41,4 % | 41,5 % | -6,1 % | -0,1 % |
| University/university college | 53,3 % | 29,8 % | 29,8 % | 23,5 % | 0,0 % |

Table 6 demonstrates the effects of weight 2 on party affiliation. The survey was conducted a few weeks after the parliamentary election and the respondents were asked for which party they casted their vote.

*Table 6: Effect of weight 2 on party affiliation*

| | Sample - not weighted | Sample - weighted | Election result | Difference been sample and election result | Difference between weighted sample and election result |
|---|---|---|---|---|---|
| The Christian Democratic Party | 5,1 % | 4,5 % | 5,6 % | -0,5 % | -1,1 % |
| The Conservative Party | 28,0 % | 27,6 % | 26,8 % | 1,2 % | 0,8 % |
| The Progress Party | 12,9 % | 15,4 % | 16,3 % | -3,4 % | -0,9 % |
| The Liberal Party | 7,0 % | 6,1 % | 5,2 % | 1,8 % | 0,9 % |
| The Socialist Left Party | 6,4 % | 5,1 % | 4,1 % | 2,3 % | 1,0 % |
| The Centre Party | 4,2 % | 4,3 % | 5,5 % | -1,3 % | -1,2 % |
| The Green Party | 4,1 % | 3,4 % | 2,8 % | 1,3 % | 0,6 % |
| The Labour Party | 28,7 % | 30,2 % | 30,8 % | -2,1 % | -0,6 % |
| Red | 1,9 % | 1,6 % | 1,1 % | 0,8 % | 0,5 % |
| Other | 1,6 % | 1,8 % | 1,8 % | -0,2 % | 0,0 % |

Weight 2 also brings the self-reported party voting of the net sample closer to the election results. The Socialist Left, The Green Party and The Liberal Party are clearly overrepresented in the net sample with 2.3, 1.3 and 1.8 percentage points respectively. The Progress Party and the Labour Party are underrepresented with 3.4 and 2.1 percentage points respectively. The weighted distribution corresponds fairly well with the election results and only in two cases do we observe a difference larger than 1.0 percent points. Firstly, when it comes to The Christian Democrats the weight has in fact increased the difference between the sample and the actual election result (from -0,5 to – 1,1). Secondly, in the case of The Centre Party, applying the weight only has a marginal effect and the weighted result is still 1,2 percent below the election result. Both of these parties have a distinct geographical distribution of votes which partly cuts across the over-simplified geographical stratification used when calculating the weights.

*Table 7: Effect of weight 2 on election turn-out*

| Not weighted | Weighted | Population | Not weighted - population | Weighted - population |
|---|---|---|---|---|
| 85,6 % | 84,4 % | 77,7 % | 7,9 % | 6,7 % |

As could be expected, the self-reported election turn-out in the net sample is higher than the official turn-out in the parliamentary election (85,6 % compared to the official turn-out of 77,7 %). This is partly due to the fact that our net sample is overrepresented by individuals with higher education and an interest in politics. Moreover, as reported by the Norwegian Election Survey Program, Norwegians have a tendency to report that they voted even in cases where they abstained.[4]

Applying weight 2 brings the survey result closer to the official turn-out, but only marginally. A substantial part of the remaining difference is probably caused by the tendency to over-report the turn-out. As much as four percent of the respondents in The Norwegian Election Survey who report that they voted in the 2009 election, did not vote according to the voting registry.[5]

## SURVEY EXPERIMENTS

The first wave of the Citizen Panel Survey includes several survey experiments where different groups of respondents receive questions with slightly different wordings. This was achieved by randomly assigning respondents to groups during the data collection process. In addition, there is also a more permanent split of the respondents in two groups. To reduce the overall time required to answer the survey, some sections of the questionnaire were only presented to one of these groups. For both reasons, the number of respondents who have answered a single question might be substantially lower than the total number of respondents. See the detailed data documentation for further information about this.

[4] Berglund, Frode, Ingvild S. Reymert og Bernt Aardal (2011). *Valgundersøkelsen 2009. Dokumentasjonsrapport.* Statistisk Sentralbyrå, Oslo – Kongsvinger.
[5] ibid.

## APPENDIX

*Table 8: Weights applied to different stratums (weight 2)*

| | | | Men | Women | | | | Men | Women |
|---|---|---|---|---|---|---|---|---|---|
| Oslo/Akershus | 18-29 years | No education/elementary school | 4,55 | 2,19 | Western Norway | 18-29 years | No education/elementary school | 3,74 | 2,96 |
| | | Upper secondary education | 0,75 | 0,68 | | | Upper secondary education | 0,70 | 0,77 |
| | | University/university college | 0,62 | 0,72 | | | University/university college | 0,81 | 0,67 |
| | 30-59 years | No education/elementary school | 3,59 | 4,85 | | 30-59 years | No education/elementary school | 2,64 | 3,14 |
| | | Upper secondary education | 1,04 | 0,93 | | | Upper secondary education | 1,00 | 1,13 |
| | | University/university college | 0,53 | 0,56 | | | University/university college | 0,48 | 0,56 |
| | 60 and above | No education/elementary school | 1,11 | 1,25 | | 60 and above | No education/elementary school | 1,56 | 1,71 |
| | | Upper secondary education | 1,47 | 2,27 | | | Upper secondary education | 1,36 | 2,07 |
| | | University/university college | 0,45 | 0,41 | | | University/university college | 0,38 | 0,50 |
| Eastern Norway | 18-29 years | No education/elementary school | 6,70 | 1,91 | Trøndelag | 18-29 years | No education/elementary school | 5,76 | 1,68 |
| | | Upper secondary education | 0,94 | 0,81 | | | Upper secondary education | 0,73 | 0,83 |
| | | University/university college | 0,78 | 0,96 | | | University/university college | 0,71 | 0,96 |
| | 30-59 years | No education/elementary school | 2,92 | 2,62 | | 30-59 years | No education/elementary school | 2,08 | 2,98 |
| | | Upper secondary education | 1,33 | 1,09 | | | Upper secondary education | 1,15 | 1,64 |
| | | University/university college | 0,56 | 0,57 | | | University/university college | 0,64 | 0,68 |
| | 60 and above | No education/elementary school | 1,62 | 2,27 | | 60 and above | No education/elementary school | 1,68 | 1,95 |
| | | Upper secondary education | 1,31 | 2,24 | | | Upper secondary education | 1,36 | 2,25 |
| | | University/university college | 0,41 | 0,52 | | | University/university college | 0,46 | 0,68 |
| Southern Norway | 18-29 years | No education/elementary school | 11,58 | 3,13 | Northern Norway | 18-29 years | No education/elementary school | 3,22 | 5,60 |
| | | Upper secondary education | 0,85 | 0,65 | | | Upper secondary education | 1,45 | 1,01 |
| | | University/university college | 0,79 | 3,26 | | | University/university college | 0,56 | 1,29 |
| | 30-59 years | No education/elementary school | 4,67 | 2,87 | | 30-59 years | No education/elementary school | 2,62 | 2,57 |
| | | Upper secondary education | 1,37 | 1,28 | | | Upper secondary education | 1,38 | 1,05 |
| | | University/university college | 0,61 | 0,50 | | | University/university college | 0,53 | 0,66 |
| | 60 and above | No education/elementary school | 1,57 | 1,98 | | 60 and above | No education/elementary school | 2,28 | 4,98 |
| | | Upper secondary education | 1,54 | 2,38 | | | Upper secondary education | 2,23 | 2,29 |
| | | University/university college | 0,37 | 0,63 | | | University/university college | 0,66 | 0,82 |