

Norwegian Citizen Panel

2023, 27th Wave

Methodology report

Øivind Skjervheim

Olav Bjørnebekk

Joachim Wettergreen

Ola Grendal

August, 2023

TABLE OF CONTENTS

Background.....	2
Technical Aspects of the Survey	2
Software.....	2
Pilot, Soft Launch, and Distribution	2
Randomization Procedures	2
Previous Waves of Recruitment.....	3
Data Collection.....	4
Responses by Method of Data Collection.....	4
Reduced responderate from @online.no-respondents	5
Response of Existing Panel Members Over Time	6
Platforms.....	6
Time Usage	7
Representativity.....	8
Factors Explaining Lack of Representativity	8
The Representativity of the Norwegian Citizen Panel	8
Weighting.....	12
Appendix	14

BACKGROUND

This report describes the procedures of data collection in the 27th wave of The Norwegian Citizen Panel. Technical aspects of data collection are discussed, along with the representativity of the panel, and how survey weights are calculated.

The Norwegian Citizen Panel (NCP) is one of the main components of Digital Social Science Core Facility (DIGSSCORE) at the University of Bergen. NCP is a collaboration between several departments at the Faculty of Social Sciences at the University of Bergen and NORCE.

ideas2evidence is responsible for the panel recruitment, the administration of the panel, and the technical solutions regarding data collection and computing.

TECHNICAL ASPECTS OF THE SURVEY

SOFTWARE

The surveys are administered through the web-based survey software Conformat. Conformat is a "Software-as-a-Service" solution, where all software runs on Conformat's continuously monitored server park, and where survey respondents and developers interact with the system through various web-based interfaces. This software provides very high data security and operational stability. The security measures are the most stringent in the industry, and Conformat guarantees 99.7 percent uptime. ideas2evidence programs the survey in Conformat on behalf of The Norwegian Citizen Panel.

PILOT, SOFT LAUNCH, AND DISTRIBUTION

The survey went through small-N pilot testing before data collection. In addition, the survey was tested extensively during the development phase by ideas2evidence and the researchers involved in the project.

The pilot testing was regarded as successful, and no major technical revisions were deemed necessary.

The field period started by inviting a random sample of high participation respondents (soft launch). Soft launch is used in order to minimize the consequences if the questionnaire contained technical errors. No technical errors were discovered during soft launch.

RANDOMIZATION PROCEDURES

Each wave of NCP has an extensive use of randomization procedures. The context of each randomization procedure may vary,¹ but they all share some commonalities.

All randomization procedures are executed live in the questionnaire. This means that the randomization takes place while the respondent is in the questionnaire, as opposed to pre-defined randomizations that are uploaded to the questionnaire. All randomizations are independent from another, unless the documentation states otherwise.

The randomization procedures are written in JavaScript. `Math.random()`² is used in combination with `Math.floor()`³. These functions are used to achieve the following:

¹ Some examples: sorting respondents in different thematic subsets, randomly allocate treatment values in experiments, randomize order of an answer list/array, order a sequence of questions by random, ask a given question to a subset of the respondents.

² Please see following resource (or other internet resources): https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/Math/random

³ Please see following resource (or other internet resources): https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/Math/floor

- Randomly select one value from a vector
- Randomly shuffle the contents of an array

The first procedure is typically used to determine a random sample of respondents to i.e. a control group. Say for example we wish to create two groups of respondents: group 1 and group 2. All respondents are randomly assigned the value 1 or 2, where each randomization is independent from one another. When N is large enough these two groups will be of equal size (50/50).

Here is an example of the JavaScript code executed in Confrontit:

```
var form = f("x1");
if(!form.toBoolean()) // If no previous randomization on x1
{
  var precodes = x1.domainValues(); // Copies the length of x1
  var randomNumber : float = Math.random() * precodes.length;
  var randomIndex : int = Math.floor(randomNumber);
  var code = precodes[randomIndex];
  form.set(code);
}
```

The second procedure is typically used when defining the order of an answer list as random. This can be useful for example when asking for the respondent's party preference or in a list experiment. However, since a party cannot be listed twice, the procedure must take into account that the array of parties is reduced by 1 for each randomization.

Here is an example of the JavaScript code executed in Confrontit⁴:

```
Function shuffle(array) {
  var currentIndex = array.length, temporaryValue, randomIndex;
  // While there remain elements to shuffle ...
  while (0 <= currentIndex) {
    // Pick a remaining element ...
    randomIndex = Math.floor(Math.random() * currentIndex);
    currentIndex -= 1;

    // And swap it with the current element.
    temporaryValue = array[currentIndex];
    array[currentIndex] = array[randomIndex];
    array[randomIndex] = temporaryValue;
  }
  return array;
}
```

PREVIOUS WAVES OF RECRUITMENT

Existing panel members were recruited in wave 1, wave 3, wave 8, wave 11, wave 14, wave 16, wave 18, wave 22, and wave 25. All samples were drawn from the National Population Registry of Norway. This registry holds information on everyone born in Norway, as well as former and current inhabitants. The Norwegian Citizen Panel received the sample after all relevant permissions were acquired from the Norwegian Tax Administration.

Samples consist of people over the age of 18 who were randomly drawn from the registry. The extracted information was a) last name, b) first name, c) address, d) gender, e) year of birth, and f) phone number (the latter was not included in wave 1). Samples exclude people without a permanent address in Norway.

⁴ Code collected from Mike Bostocks visualization: <https://bost.ocks.org/mike/shuffle/>

Table 1 outlines a short summary of previous recruitment efforts. Note that there are some differences between the recruitment processes. For a detailed description of each recruitment process, please refer to the respective methodology reports.

Table 1: Summary of recruitment processes

	Sample size	Mode	Contacts	Returned letters	Response Rate (%)
Recruitment 1 (wave 1)	25 000	Postal	2	546	20.1 %
Recruitment 2 (wave 3)	25 000	Postal, phone/SMS	4	543	23.0 %
Recruitment 3 (wave 8)	22 000	Postal/SMS	3	479	19.4 %
Recruitment 4 (wave 11)	14 000	Postal/SMS	2	334	15.1 %
Recruitment 5 (wave 14)	14 000	Postal/SMS	2	389	15.0 %
Recruitment 6 (wave 16)	34 000	Postal/SMS	2	994	14.9 %
Recruitment 7 (wave 18)	15 000	Postal/SMS	2	381	14.0 %
Recruitment 8 (wave 22)	23 000	Postal/SMS	2	623	14.5 %
Recruitment 9 (wave 25)	18 000	Postal/SMS	2	562	13.9 %

The response rate of recruitment in wave 11 and onwards were lower than recruitment in previous waves. The most important explanation is new restrictions enforced by the Norwegian Tax Administration with regards to how many times the Citizen Panel can contact the invitees. Respondents in recruitments 4-8 were contacted twice at most. Recruitment 1 also had a maximum of two contact points, but achieved a response rate of 20 percent. One explanation for why we cannot replicate a response rate of 20 percent in recruitments 4-8 might be that NCP did a lot of promotion of the panel through media outlets leading up to and during recruitment 1. Additional promotion of the panel was carried out in relation to the Norwegian Parliamentary election the same fall.

DATA COLLECTION

RESPONSES BY METHOD OF DATA COLLECTION

The survey was distributed to 27 213 panel members on the 2nd and 5th of June 2023 for the softlaunch and main launch respectively. The invitation contained information on the Norwegian Citizen Panel, unique URLs for each panel member that led to the questionnaire, and unique access code which the panel members could use to log in to the survey by accessing a link on www.uib.no/medborger.

The invitation, first reminder, and second reminder were all distributed by e-mail. The third and last reminder was, depending on whether the panel member had a registered mobile phone number or not, distributed via SMS or e-mail. Prior to wave 27, 53.3 percent of the panel members were registered with a mobile phone number.

Table 2: Responses and response rate for panel members by the different stages of data collection

	Response	Cumulative Responses	Response Rate (%)	Cumulative Response Rate
Invitation (2nd/5th of June)	4 505	4 505	29.2 %	29.2 %
First reminder (8th of June)	2 662	7 167	17.2 %	46.4 %
Second reminder (14th of June)	1 182	8 349	7.7 %	54.1 %
Third reminder - email (19th of June)	533	8 882	3.5 %	57.6 %
Third reminder - SMS (19th of June)	1 115	9 997	7.2 %	64.8 %

In total 9 997 existing panel members filled out the questionnaire. A response rate of 29.2 % was achieved between the invitation and the first reminder. Following a pattern observed in previous waves, the initial invitation produced a higher number of respondents than subsequent reminders. See table 2 for further details on number of respondents after reminders.

Using the same methodology as in previous waves for calculating response rate, respondents who have not participated in any of the last three waves are excluded. This leaves us with 15 440 eligible respondents. The overall response rate, as reported in table 2, is 64.8 %.

Approximately 1 300 of the initial invitations were reported as not delivered by Confermit, which rounds to 4 percent. Measures are taken to ensure email deliverability, but are unable to accurately estimate how many of the delivered emails ended up as spam with the recipient.

REDUCED RESPONSERATE FROM @ONLINE.NO-RESPONDENTS

During data collection, it became evident that fewer panel members registered with an online.no email address responded, compared to previous waves. While we expect a drop-off in participation over time, the decline appeared too steep to be caused by organic deterioration.

Telenor, the owner of the email client made changes to their service on the 1st of June 2023. Notably, starting to require payment for the previously free service, and enhanced the built-in spam filter. The change in policy has been poorly received by costumers, and drawn criticism from interest groups for exploiting the elderly, who might have difficulties switching platforms.⁵ We find it likely that these changes affected our ability to reach the affected panel members. Firstly, due to customers abandoning the platform in lieu of other options, and secondly, due to a more invasive spam filter blocking our emails.

In the table below the number of participating panel members with a @online.no-domain can be found for wave 27, wave 26, and wave 24.⁶ Compared to wave 26, we observe 812 fewer responses, a decline of 44 percent.

Number of respondents with an email associated with the @online.no-domain and share of them using a computer to answer

Wave	Number of respondents	Respondents using a computer to answer
Wave 24	1 683	66.0 %
Wave 26	1 842	65.5 %
Wave 27	1 030	47.7 %

There were approximately 3 500 panel members associated with an email address with the @online.no-domain before wave 27 was fielded. A significant number of these are, however, inactive panel members.

The third and final reminder use both e-mail and SMS as mode of contact, where panel members who have not opened the survey and is registered with a phone number will receive a text instead of an email. After the text reminder was sent, and influx of responses from panel members with an e-mail address from the @online.no-domain was registered, as well as ideas2evidence received several comments, both through email and closing comments to the survey, from panel members with an @online.no-email that they preferred to receive the invitation by email. Panel members that receive a text reminder is likely to use their phone to register their answer in the survey, as the survey can be accessed directly from the reminder. The drop in share of responses registered from a computer, as can be seen in the table above, further underlines the likelihood that the changes from Telenor has affected our capabilities in reaching these respondents through e-mail.

⁵ See for instance: <https://tek.no/nyheter/nyhet/eJy8RO/slakter-telenors-epost-betaling-utnytter-en-saarbar-gruppe>

⁶ We have chosen to compare with wave 26 as it is the most recent wave before wave 27, and has roughly the same population of panel members to draw from (some panel members are likely to have opted out from participating). In addition, we have compared it to the number of respondents in wave 24 as both waves are fielded at the same time during the year, and has the same amount of time since recruitment to the panel, despite that recruitment in wave 25 changed the population.

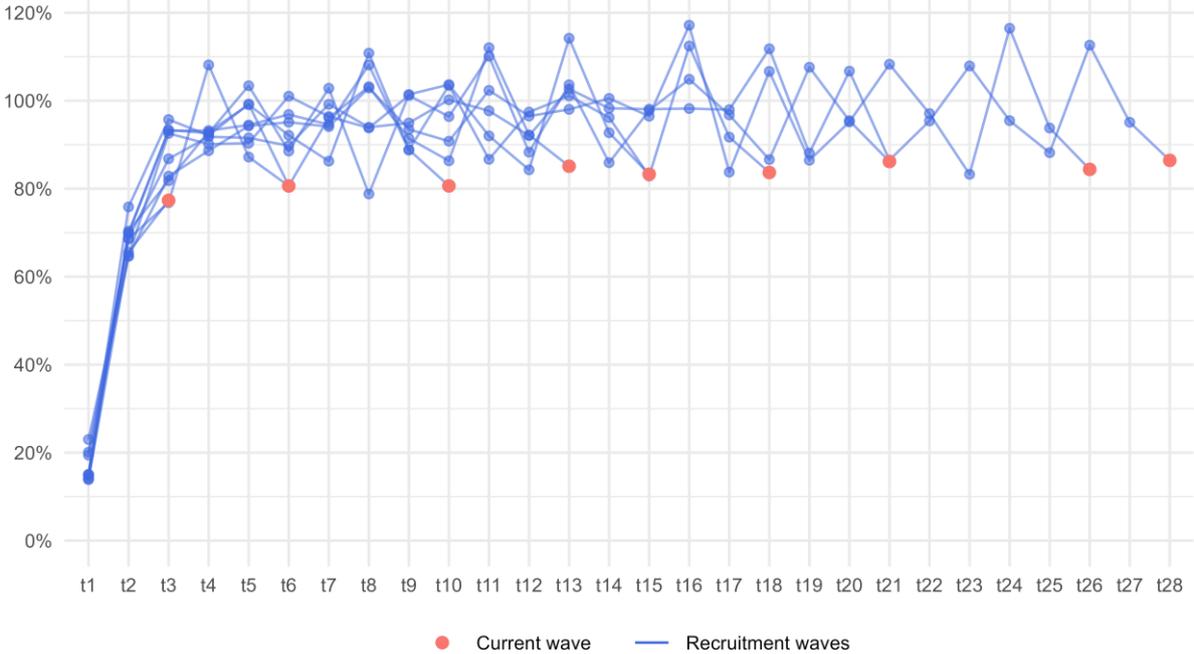
As there were comparatively a low number of respondents with an email associated with the @online.no-domain in wave 27, the drop in share of respondents using a computer to answer, and the number of respondents reaching out to us to explain that they have not seen the initial invitation and reminders sent by email, we assume that there is a relatively large proportion of panel members we have been unable to reach in wave 27, compared to previous waves.

RESPONSE OF EXISTING PANEL MEMBERS OVER TIME

Comparing the number of wave 27 respondents (9 997) to the number of respondents in the previous wave 26 (12 021), gives an overall wave-to-wave retention rate of 83 percent. Figure 1 illustrates each wave of recruitment by individual lines, and shows how many respondents that are preserved for each data collection. NCP has carried out 28 waves of data collection. Depending on when the respondents were recruited, the current wave is highlighted with a red circle. For respondents recruited in wave 1, the current wave is the 28th data collection (t28). For respondents recruited in wave 22, the current wave is the sixth data collection (t6).

The wave-to-wave retention rate increases substantially after the first three waves (t1 - t3), until it stabilizes around a mean of 95 percent. Across all waves of recruitment, the current wave has a retention rate of below 100 percent. In other words, fewer respondents participated in wave 27 compared to wave 26. That fewer respondents participate in the second wave during the year is not unusual when comparing across the last few years; it is consistently the wave where the least amount of panel members participate.

Figure 1: Wave-to-wave retention rate



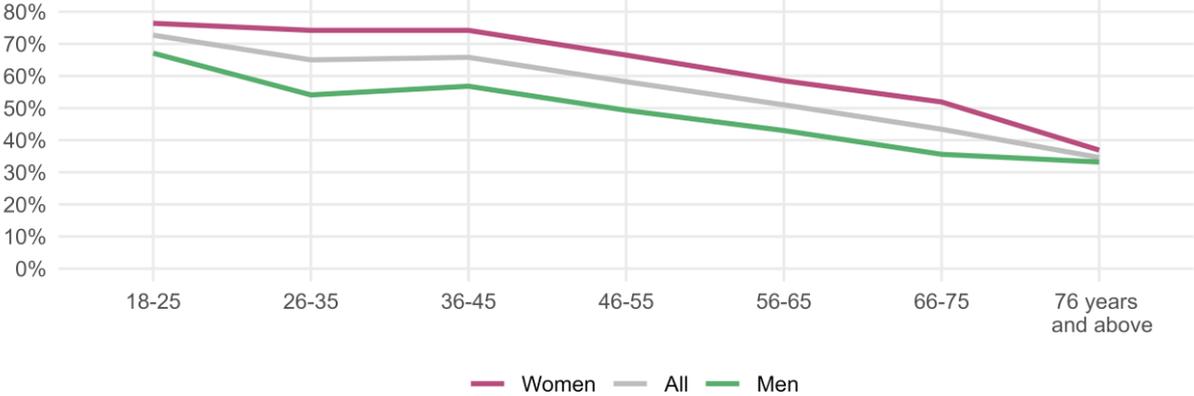
PLATFORMS

The questionnaire was prepared for input via smart phones, tables, and other units capable of running web-browsers. In order to enhance the respondents’ experience, the questionnaire is responsive. Respondents on smaller devices, measured in pixels per inch (PPI), are exposed to slightly different visual representations of some questions. For instance is a question grid presented as a set of individual questions on the same page, which is different from the desktop presentation where it would be presented in a table. 49 percent of all survey respondents that opened the questionnaire used a mobile phone.

A set number of survey questions must be answered for a person to be included as a respondent. 7 percent of the mobile users did not reach this minimum requirement, compared to 15 percent for non-mobile users.

The share of mobile users is high among respondents between 18 and 45 of age. As shown in figure 2, the share of mobile users decline substantially with age.

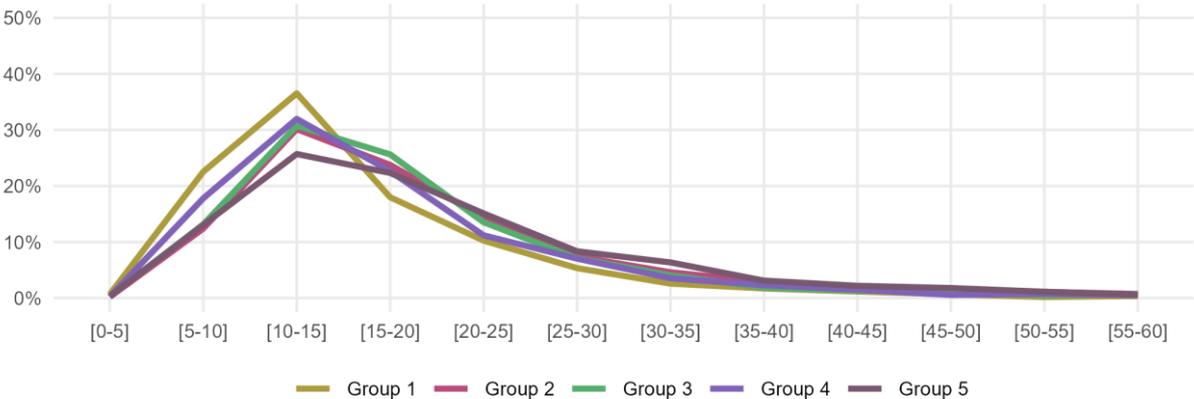
Figure 2: Share of mobile users by gender and age



TIME USAGE

The average respondent used 17.9 minutes to complete the questionnaire. Measuring average time usage is a challenge, as respondents may leave the questionnaire open in order to complete the survey later. This idle time causes an artificially high average for completing the survey. The average therefore includes only the respondents that spent 60 minutes or less completing the survey.

Figure 3: Time usage distribution of survey respondents in subgroups



The questionnaire was divided into five subsets of questions in addition to questions, which were exposed to all respondents, and a further subset indicated as panel group in the data. The different sets were exposed to five subgroups of respondents, determined by random for the main subsets (r27_group in the data), and conditional random allocation (r27_panelgroup). For the latter, respondents who participated and were randomised to either group 3 or 4 in wave 26 were assigned to panel group 1 and 2 respectively. All other respondents were randomly allocated to panel group 3, 4, or 5.

Time usage across the different groups varies between 15 and 20 minutes. Respondents in group 5 spent the most time to complete the survey, while group 1 spent the least time.

Table 3: Average time usage (minutes) in each subset

	All	Group 1	Group 2	Group 3	Group 4	Group 5
All users	17.9	15.8	18.8	18.1	17.3	19.7
Non-mobile users	18.8	16.3	19.5	19	18.3	21
Mobile users	17	15.2	18.1	17.1	16.3	18.4

REPRESENTATIVITY

In this section, we describe the representativity of the panel as a whole. First, we will discuss factors explaining representativity. Thereafter we apply demographic variables to present data on representativity by different strata. The data on representativity is the foundation for the section on weighting.

FACTORS EXPLAINING LACK OF REPRESENTATIVITY

There are two main points that can serve as explanations to non-response and lack of representativity when recruiting panel members and maintaining panel members:

- ◆ access to and familiarity with the internet (given that a web-based questionnaire was the only response mode made available)
- ◆ 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 online questionnaire, normally decreases with increasing age. The second challenge, motivation and interest, is often explained by the respondents' level of education. In addition to age and education, we added the variables of geography and gender in order to test the representativity of the survey respondents. The variables have the following categories:

- ◆ Age: 18-29 years, 30-59 years, 60 and above.
- ◆ Highest completed education: no education/elementary school, upper secondary, university/university college.
- ◆ Geography: Oslo, Eastern Norway, Southern Norway, Western Norway, Trøndelag, Northern Norway.

Please note that starting wave 21, the former county of Akershus is part of Eastern Norway, rather than being part of the traditional Akershus/Oslo stratum. This has implications for weighting and representativity analyses, as discussed below.

THE REPRESENTATIVITY OF THE NORWEGIAN CITIZEN PANEL

The sampling frame of the survey equals to the Norwegian population above the age of 18, comprising a population of approximately 4.3 million individuals. Earlier reports have documented a systematic underrepresentation of respondents belonging to the two lowest educational groups, independent of gender and age. The underrepresentation is particularly strong for young men. As expected, individuals with education from universities or university colleges are overrepresented. All of these observations hold true for wave 27.

Table 4: Age distribution in the population and the net sample

	18-29 years	30-59 years	60 years and above
Population	19.4 %	50.2 %	30.4 %
Net sample	4.2 %	43.8 %	52 %

From the age distribution presented in table 4, we see that the two youngest age groups are underrepresented in the net sample of wave 27 when compared to the distribution we find in the population. Inversely, respondents aged 60 years and more are overrepresented in the net sample when compared to the population.

Over time, the panel has exhibited a drift away from perfect representativity with regard to age. As figure 4 shows, the oldest respondents started out as underrepresented in wave 1, but have become increasingly overrepresented over time. The youngest respondents, on the other side, have become increasingly underrepresented. This can be explained by a difference in panel membership loyalty; younger panel members are more likely to stop responding to new NCP waves after having been an active member of the panel. The rate of misrepresentation of age groups peaked with wave 26.

Figure 4: Representativity of age groups

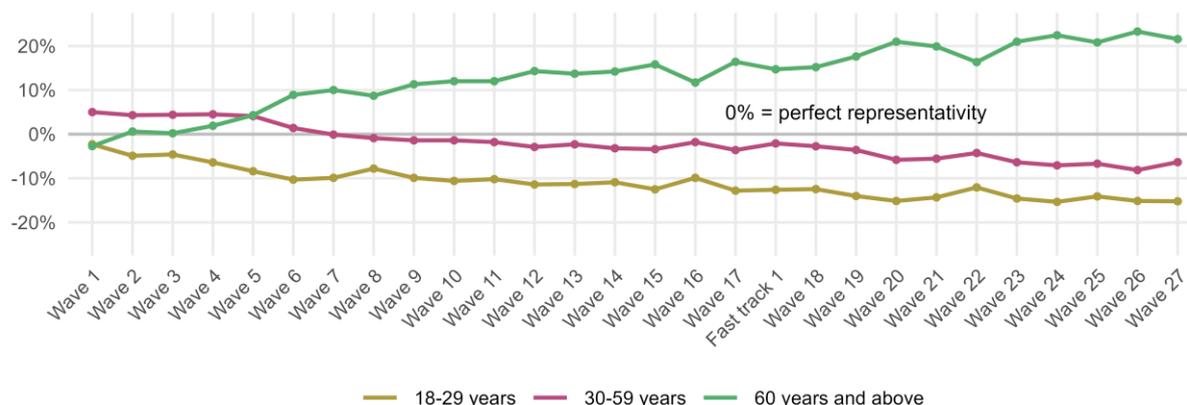


Table 5 breaks down population and the net sample by age and gender. This reveals a slight gender-age interaction in the panel representativity. Younger men are slightly more underrepresented than younger women, while among the oldest in the panel men are more overrepresented than women in the same age bracket.

Table 5: 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	9.9 %	9.4 %	25.6 %	24.6 %	14.6 %	15.9 %
Net sample	1.8 %	2.4 %	21.2 %	22.7 %	27.6 %	24.4 %

The inclusion of educational level in table 6 reveals a systematic underrepresentation of respondents with little or no education, independent of age and gender. The underrepresentation is present in all age brackets, but is especially strong for young respondents.

Table 6: 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	3.5 %	2.7 %	0.1 %	0.1 %
Upper secondary education		4.1 %	3 %	0.7 %	0.6 %
University/university college		2.4 %	3.7 %	1 %	1.7 %
No education/elementary school	30-59 years	5.1 %	4 %	0.5 %	0.4 %
Upper secondary education		10.8 %	7.4 %	6.6 %	4.5 %
University/university college		9.7 %	13.2 %	14.2 %	17.8 %
No education/elementary school	60 and above	3.1 %	4.2 %	1.8 %	1.3 %
Upper secondary education		7.2 %	7.4 %	9.2 %	7.4 %
University/university college		4.2 %	4.3 %	16.7 %	15.4 %

Respondents who have completed upper secondary education as their highest completed level of education are underrepresented in all groups, except for men with upper secondary education aged 60 years or above. Those who have university or university college education are clearly overrepresented in the two oldest age brackets, irrespective of gender.

Figure 5: Representativity of education groups

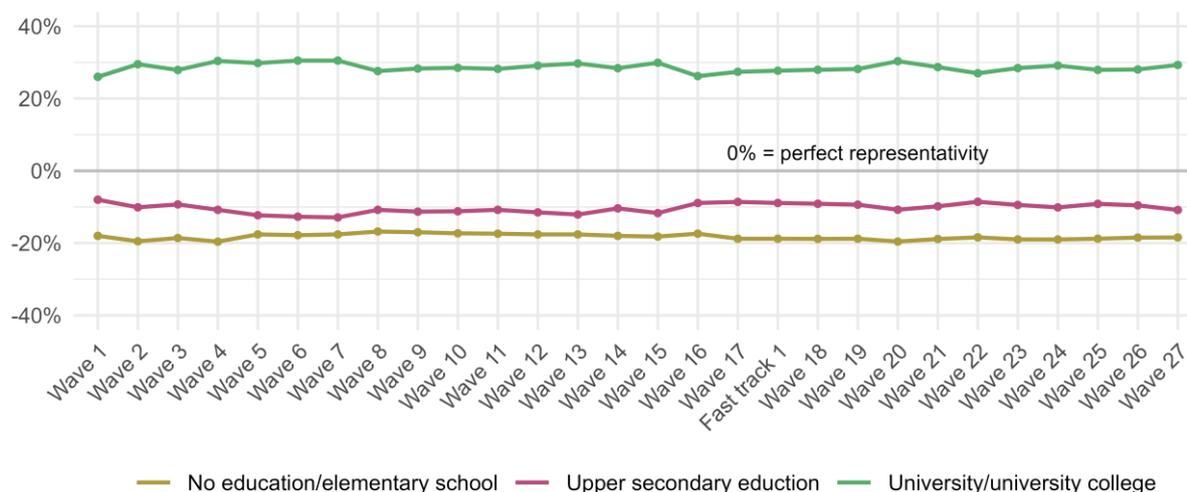


Figure 5 illustrates the representation of education groups since wave one. The general trend is that the highly educated are overrepresented compared to those with less or no education. Except for slight improvements in representativity of the education groups when new respondents are recruited (wave 1, 3, 8, 11, 14, 16, 18, 22 and 25), the overall pattern has remained stable throughout all waves.

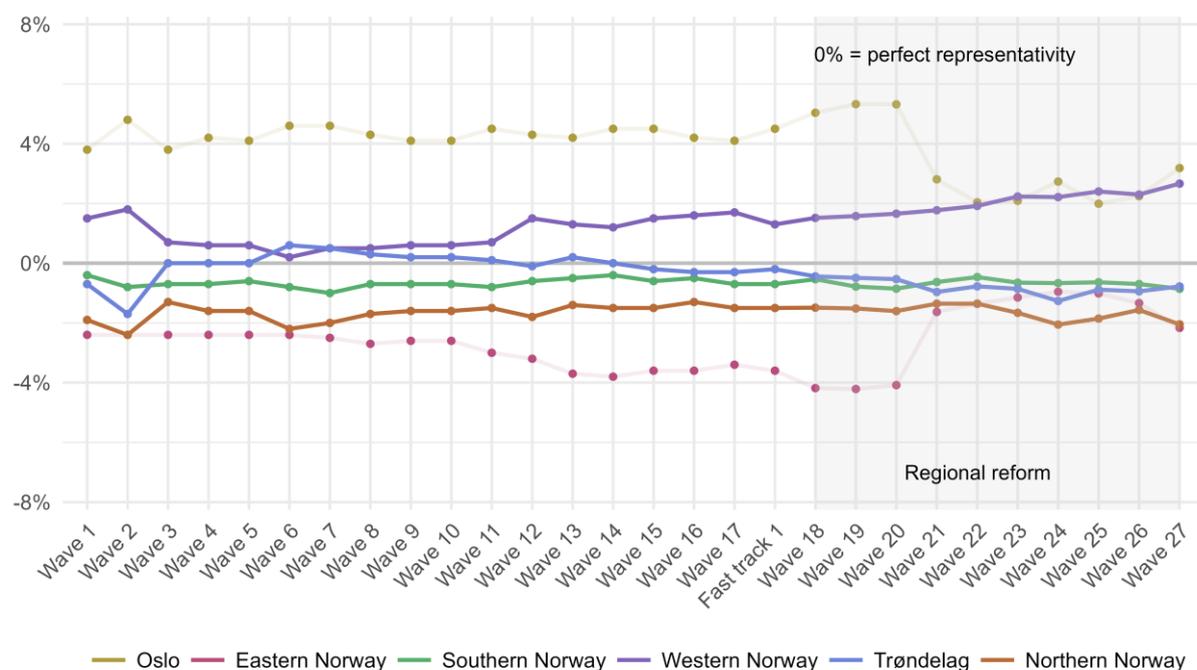
As for geography (table 7), we observe that the representation of panel members living in Trøndelag, Eastern Norway, and Southern Norway are nearly on level with the population, while Oslo and Western Norway is overrepresented. Respondents from Northern Norway meanwhile are underrepresented.

Table 7: Combined distribution of age, gender and geography in the population and the net sample

		Population			Net sample		
		Men	Women	Total	Men	Women	Total
Oslo	18-29 years	1.4%	1.6%	3%	0.4 %	0.7 %	1.1 %
	30-59 years	3.6%	3.5%	7.1%	3.9 %	4.2 %	8.1 %
	60 and above	1.3%	1.5%	2.9%	3.4 %	3.6 %	7 %
	In total	6.4%	6.6%	13%	7.7 %	8.5 %	16.2 %
Eastern Norway	18-29 years	3.3%	3.1%	6.4%	0.3 %	0.7 %	1 %
	30-59 years	9.7%	9.5%	19.2%	6.8 %	7.8 %	14.7 %
	60 and above	5.9%	6.5%	12.5%	11 %	9.3 %	20.3 %
	In total	19%	19.1%	38.1%	18.1 %	17.8 %	35.9 %
Southern Norway	18-29 years	0.6%	0.5%	1.1%	0.1 %	0.1 %	0.2 %
	30-59 years	1.4%	1.4%	2.8%	1 %	1.1 %	2.1 %
	60 and above	0.8%	0.9%	1.8%	1.3 %	1.2 %	2.5 %
	In total	2.8%	2.8%	5.7%	2.4 %	2.4 %	4.8 %
Western Norway	18-29 years	2.5%	2.3%	4.9%	0.6 %	0.4 %	1.1 %
	30-59 years	6.5%	6.1%	12.6%	6.2 %	6 %	12.2 %
	60 and above	3.7%	4%	7.6%	7.7 %	6.8 %	14.5 %
	In total	12.7%	12.4%	25.1%	14.5 %	13.2 %	27.8 %
Trøndelag	18-29 years	1.1%	1%	2.1%	0.2 %	0.4 %	0.6 %
	30-59 years	2.2%	2.1%	4.2%	1.8 %	1.9 %	3.7 %
	60 and above	1.3%	1.4%	2.7%	2.1 %	1.8 %	3.9 %
	In total	4.5%	4.4%	9%	4.1 %	4.1 %	8.2 %
Northern Norway	18-29 years	1%	0.9%	1.9%	0.1 %	0.2 %	0.3 %
	30-59 years	2.2%	2.1%	4.2%	1.4 %	1.6 %	3 %
	60 and above	1.5%	1.5%	3%	2.1 %	1.7 %	3.8 %
	In total	4.6%	4.5%	9.2%	3.7 %	3.4 %	7.1 %

Respondents aged 60 years and above are overrepresented in all parts of the country, which is especially so for Oslo and Western Norway. Conversely, respondents aged 18-29 years are underrepresented in all regions.

Figure 6: Representativity of regions



For wave 21, population data stratified on the new regions was available for the first time since the regional reform of 2020. While this data eliminates some small uncertainty in the representativity analyses⁷, it also introduces a break in time series for Oslo (previously including Akershus) and Eastern Norway (now including Akershus). Eastern Norway now makes up almost one fourth of the population, the diversity of which we are no longer able to account for in full in the respondents' weights. Compared to age and education, geography does, however, not seem to be a strong determinant of survey participation. Apart from effects from the regional reform, the geographic representativity is more or less stable over time.

WEIGHTING

To compensate for the observed biases, we have calculated a set of weights. The weights are equal to 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.⁸ This procedure returns values around 1, but above 0. Respondents belonging to a stratum that is underrepresented will receive a weight above 1 and respondents belonging to an overrepresented stratum will receive a weight below 1. We have listed the weights of the different strata in table 10 in the appendix.

When calculating the weights, information regarding the respondents' geographical location, gender and age is based on registry data. Information on these variables was included in the sample file we received from the Norwegian National Registry. Information regarding the level of education is collected from NCP surveys. 1 percent of the 26th wave net sample have not answered the question about level of education. Four different weights have been calculated:

- ◆ **Weight 1** is based on demographic variables only (age, gender and geography)
- ◆ **Weight 2** combines 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 in these cases set to 1).
- ◆ **Weight 4** is similar to weight 2, but highest level of education is recoded into 1) upper secondary and lower, and 2) university/university college
- ◆ **Weight 5** is based on gender and highest level of completed education which is recoded into 1) upper secondary and lower, and 2) university/university college. Weight values are capped ($x \geq .2$ and $x \leq 5$).

The variables have the following categories:

- ◆ **Age:** 18-29 years, 30-59 years, 60 and above.
- ◆ **Gender:** Male and female.
- ◆ **Highest completed education:**
 - **Weight 2:** no education/elementary school, upper secondary, university/university college.

⁷ Note that Oslo (including Akershus), and Eastern Norway diverge in wave eighteen, due to the regional reform implemented 1st of January 2020.

⁸ The applied formula for weight w_i for element i , in strata h is: $w_i = \frac{N_h/N}{n_h/n}$

- **Weight 4 and 5:** upper secondary and lower, university/university college.
- ◆ **Geography:** Oslo, Eastern Norway, Southern Norway, Western Norway, Trøndelag, Northern Norway.

The method for calculating weights is the same as in previous waves, yet the stratification method for geographic regions changed from wave 21 onwards as new population data based on the region reform that came into effect in January 2020 became available.

When applied, both weights will provide a weighted N equal to the number of cases in the dataset. In other words, the weights are calculated using the whole dataset. NCP has an extensive use of (randomized) sub-groups, which might alter the demographic profile of the sub-group compared to the whole dataset. Consequently, the weights might be less precise for some sub-groups. Note that the dataset is provided with necessary information⁹ to calculate custom weights if needed, following the procedure described above.

As discussed above, level of education is the greatest source of observed bias. Therefore, weight 2 might provide an accurate compensation for the various sources of bias in the net sample.

Table 8 shows the effects of weight 2 on the distribution of self-reported level of education in the net sample. As we can observe, the weight gives the sample a distribution close to the population. It is, however, important to stress that the unweighted distribution is far from ideal, with a clear underrepresentation of people with low levels of education.

Table 8: 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	4.2 %	21.6 %	22.6 %	-18.4	-1
Upper secondary school	29 %	40.4 %	39.9 %	-10.8	0.5
University/university college	66.8 %	38 %	37.5 %	29.3	0.5

⁹ See columns r27_Weight1stratapop and r27_Weight2stratapop

APPENDIX

Table 9: Weights applied to different strata (weight 2)

		Men	Women			Men	Women		
Oslo	18-29 years	No education/elementary school	36.8		Western Norway	No education/elementary school	20.6	31.7	
		Upper secondary education	7.3	3.6		Upper secondary education	3.9	4.7	
		University/university college	1.8	1.7		University/university college	1.9	3.6	
	30-59 years	No education/elementary school	15.9	12.6		30-59 years	No education/elementary school	8.9	10.5
		Upper secondary education	1.5	1.7			Upper secondary education	1.4	1.4
		University/university college	0.7	0.6			University/university college	0.6	0.7
	60 and above	No education/elementary school	1.6	3.1		60 and above	No education/elementary school	1.4	3.2
		Upper secondary education	1	0.7			Upper secondary education	0.7	0.8
		University/university college	0.2	0.3			University/university college	0.2	0.2
Eastern Norway	18-29 years	No education/elementary school	42.3	47.4	Trøndelag	No education/elementary school	41.3	16	
		Upper secondary education	9.5	8.1		Upper secondary education	3.9	3.1	
		University/university college	5	2.1		University/university college	3.1	1.6	
	30-59 years	No education/elementary school	9.9	8.9		30-59 years	No education/elementary school	9.5	4.8
		Upper secondary education	1.8	1.8			Upper secondary education	2.2	2
		University/university college	0.8	0.8			University/university college	0.6	0.7
	60 and above	No education/elementary school	1.9	2.9		60 and above	No education/elementary school	1.7	4.6
		Upper secondary education	0.8	1.1			Upper secondary education	0.9	1.3
		University/university college	0.3	0.3			University/university college	0.3	0.3
Southern Norway	18-29 years	No education/elementary school		15.7	Northern Norway	No education/elementary school			
		Upper secondary education	8.5	6.3		Upper secondary education	6.7	3.7	
		University/university college	4	3.8		University/university college	3.2	2.8	
	30-59 years	No education/elementary school	27.4	24.5		30-59 years	No education/elementary school	17	11.7
		Upper secondary education	1.6	1.6			Upper secondary education	2.1	1.7
		University/university college	0.8	0.8			University/university college	0.7	0.9
	60 and above	No education/elementary school	3.2	2.9		60 and above	No education/elementary school	2.1	6.1
		Upper secondary education	1.1	1.5			Upper secondary education	0.9	1.4
		University/university college	0.3	0.3			University/university college	0.3	0.3