

Norwegian Citizen Panel

2022, 25th Wave

Methodology report

Øivind Skjervheim

Olav Bjørnebekk

Joachim Wettergreen

Ola Grendal

December 2022

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.....	4
Data Collection.....	4
Recruiting a New Set of Panel Members	4
The Recruitment Process	4
Results of the Recruitment Process - Survey Respondents and Panel Members	5
Responses by Method of Data Collection.....	6
Responses by Existing Panel Members	6
Response of Existing Panel Members Over Time	7
Platforms.....	7
Time Usage	8
Representativity.....	9
Factors Explaining Lack of Representativity	9
The Representativity of the Norwegian Citizen Panel	9
Weighting.....	13
Appendix	15

BACKGROUND

This report describes the procedures of data collection in the 25th 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 i.e. 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;
}
```

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

PREVIOUS WAVES OF RECRUITMENT

Existing panel members were recruited in wave 1, wave 3, wave 8, wave 11, wave 14, wave 16, wave 18, and wave 22. 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 Tax Administration holds the formal responsibility for this registry, but the administration is partly outsourced to the private IT-company Evry. Evry drew the sample on behalf of the Norwegian Citizen Panel after 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.

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. A detailed description of the recruitment in wave 25 follows in the next section.

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 recruitments 4-8 were substantially lower than earlier waves of recruitment. The most important explanation is new restrictions enforced by the Norwegian Tax Administration with regards to how many times the Citizen Panel can contact people in the net sample. 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

RECRUITING A NEW SET OF PANEL MEMBERS

In wave 25, the Norwegian Citizen Panel recruited new panel members. This section gives a detailed description of the sample frame, recruitment process and the results.

THE RECRUITMENT PROCESS

As in the preceding waves of recruitment a gross sample was drawn from the population registry. Evry drew the sample on behalf of the Citizen Panel after the necessary permissions were acquired from the Norwegian Tax Administration.

18 000 people over the age of 18 were randomly drawn from the register. The extracted information was as before a) last name, b) first name, c) address, d) gender, e) telephone number(s) (if available) and, f) year of birth. The sample excluded individuals without a current home address in Norway.

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 participants, c) the time-frame of the project, d) the participants' rights to opt 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.

In order to maximize the response rate, an incentive in the form of three gift cards is included in the project. The values of the gift cards are 8 000 NOK. To enter the lottery respondents were required to join the panel and provide their email 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 invitational letter was posted 31st of October 2022. The second reminder was distributed by SMS or post card. Respondents below 60 years of age registered with a cell phone number received an SMS. Respondents who did not fit this description received a post card reminder. This is different from the first three waves of recruitment. In wave eleven, an experiment was conducted regarding the use of SMS and postcard. That experiment gave the panel more information regarding the effectiveness of different recruitment strategies, and thus gave the opportunity for a more cost-efficient use of reminders in subsequent waves.

Both reminders were sent to respondents who a) had not logged into the survey, or b) had not completed the survey. Respondents were encouraged to join the panel, with reference to the invitation letter. The unique log-in ID provided in the original letter was included in both the post card and the SMS. The SMS reminder also included a direct link to the survey.

The post card was posted the 14th of November, and the SMS was distributed November 17th.

RESULTS OF THE RECRUITMENT PROCESS - SURVEY RESPONDENTS AND PANEL MEMBERS

It is necessary to make a distinction between panel members and survey respondents. We define panel members as respondents who register their e-mail address, regardless of whether they have completed the questionnaire or not. Survey respondents are respondents who have completed a certain share of the questionnaire, regardless of whether they have entered their e-mail address or not.

Of the 17 900⁵ letters that were sent out, 562 were returned, and 10 respondents opted out. 16.5 percent (2,861) of the remaining 17 328 logged on and accessed the survey. 2 342 individuals completed the questionnaire, while 510 exited the questionnaire before completion, though 14.1 percent of these responses are kept as a part of the survey data as these respondents completed a certain amount of the questionnaire before exiting. The remaining 438 incomplete responses were excluded from the data set, due to lack of data.

In sum, after subtracting a few cases where the credentials of the respondent did not match the credentials of the invited, this recruitment wave resulted in 2 405 new survey respondents, a recruitment rate of 13.9 percent. This is slightly lower than what was achieved in recruitment 8 (14.5 %).

99.8 percent of the respondents who completed the survey submitted their e-mail address. Of the excluded respondents, 16.9 percent entered their e-mail address. In sum, after subtracting respondents with

⁵ 100 of the initial list of people to be recruited were already panel members.

mismatching credentials, 2 458 new panel members were recruited to the Norwegian Citizen Panel, resulting in a panel recruitment rate of 14.2 percent.

Further discussions in this report about new recruits in wave 25 are based on data on survey respondents. As there is an almost perfect overlap between survey respondents and panel members, the descriptions are also valid for the panel members.

RESPONSES BY METHOD OF DATA COLLECTION

Table 2 summarizes the effects of the various stages of data collection. The invitational letter accumulated 1 444 responses, the SMS generated 465 responses, and the postcard 496 responses. Resulting in a cumulative response rate of 13.9 percent.

Table 2: Number of responses and response rates for the new panel members by the various stages of data collection

	Response	Response rate (%)	Cumulative Responses	Cumulative Response Rate (%)
Invitation (31st of October)	1 444	8.3 %	1 444	8.3 %
SMS, reminder (17th of November)	465	2.7 %	1 909	11.0 %
Postcard, reminder (14th of November)	496	2.9 %	2 405	13.9 %

RESPONSES BY EXISTING PANEL MEMBERS

The survey was distributed via email to 25 448 existing panel members on the 27th and 31th of October 2022. In these e-mails, basic information about the Norwegian Citizen Panel was conveyed, and the individual panel members received unique URLs that led to the questionnaire.

The invitation, the first reminder and the second reminder were all distributed via e-mail. The third, and last reminder was, depending on whether the individual panel member had a registered mobile phone number or not, distributed via SMS or e-mail. Prior to wave 25, 51.7 percent of the panel members were registered with a mobile phone number.

Table 3: Responses and response rate for panel members by the different stages of data collection

	Response	Cumulative Responses	Response Rate (%)	Cumulative Response Rate
Invitation (27th and 24th of October)	5 169	5 169	35.1 %	35.1 %
1st reminder (November 4th)	3 610	8 779	24.5 %	59.6 %
2nd reminder – email (November 10th)	1 413	10 192	9.6 %	69.2 %
3rd reminder – email (November 17th)	454	10 646	3.1 %	72.3 %
3rd reminder – SMS (November 17th)	687	11 333	4.7 %	77.0 %

In total, 11 333 existing panel members filled out the questionnaire. A response rate of 35.1 percent was reached between the invitation and the first reminder. Following a pattern observed in earlier waves, the email invitation produced a higher number of respondents than the subsequent reminders. For details on the number of respondents after each reminder, see table 3.

When calculating the response rate, following the methodology from earlier waves, respondents who have not participated in any of the last three waves are excluded. This leaves us with 14 746 eligible respondents. The overall response rate, as reported in table 3, is 77 percent.

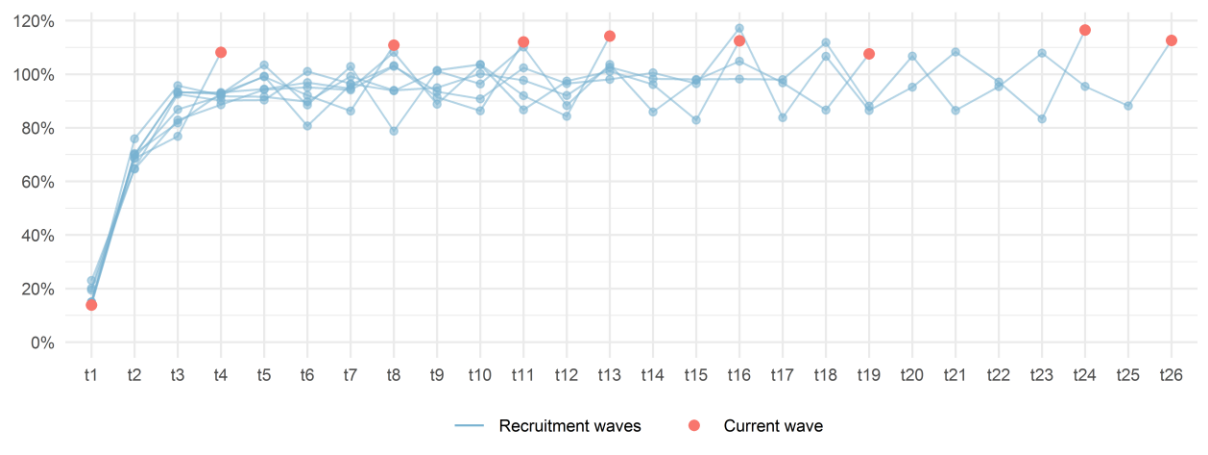
1 135 of the initial invitations were reported as undelivered by Confirmit as spam, which rounds to about 2 percent of the invited panel members. Measures are taken to ensure email deliverability, but are unable to accurately estimate how many of the delivered emails ended up as spam with the recipient.

RESPONSE OF EXISTING PANEL MEMBERS OVER TIME

Comparing the number of wave 25 respondents (11 333), to the number of respondents in the previous wave 24 (10 160), gives an overall wave-to-wave retention rate of 115 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 26 waves of data collection. Depending on when the respondents were recruited, the current wave is highlighted with a red circle. For the respondents recruited in wave 1, the current wave is the 26th data collection (t26). For the respondents recruited in wave 22, the current wave is the fourth data collection (t4).

The wave-to-wave retention rate increases substantially after the first three waves (t1 – t3), before stabilizing around a mean of 95 percent. Across all waves of recruitment, the current wave has a retention rate above 100 percent. In other words, more respondents participated in wave 25 compared to wa ve 24. Higher retention rates in the fall is a pattern we have observed over time. Wave 16, 19 and 22 all have higher retention rates than the data collections carried out in the winter or spring.

Figure 1: Wave-to-wave retention rate



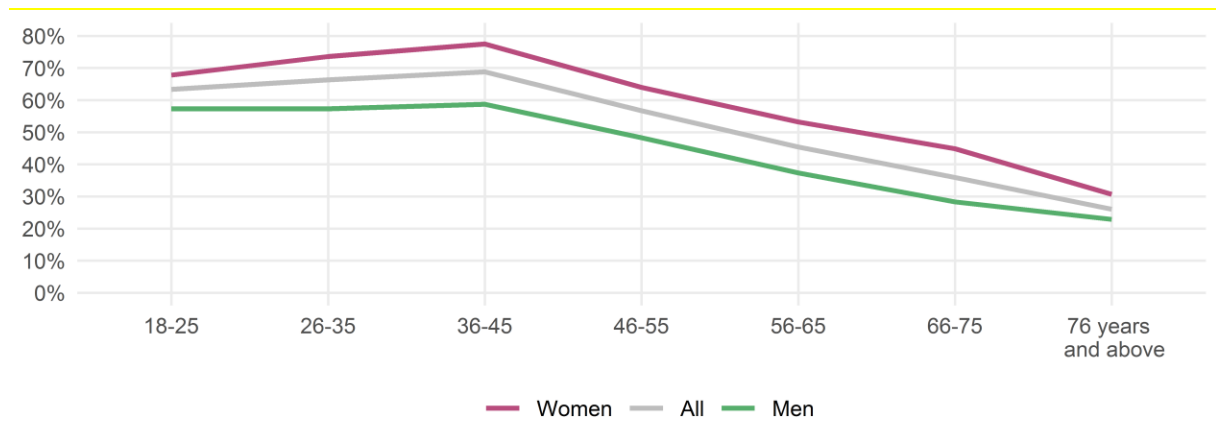
PLATFORMS

The questionnaire was prepared for data input via smart phones, tablets, and other units capable of running web-browsers. In order to enhance the respondents’ experience, the questionnaire is responsive. Meaning that respondents on small devices got a slightly different visual representation 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 grid questions are presented in a table. 48 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 survey respondent. 7 percent of the mobile users did not reach this minimum requirement, compared to 9 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 declines substantially with age.

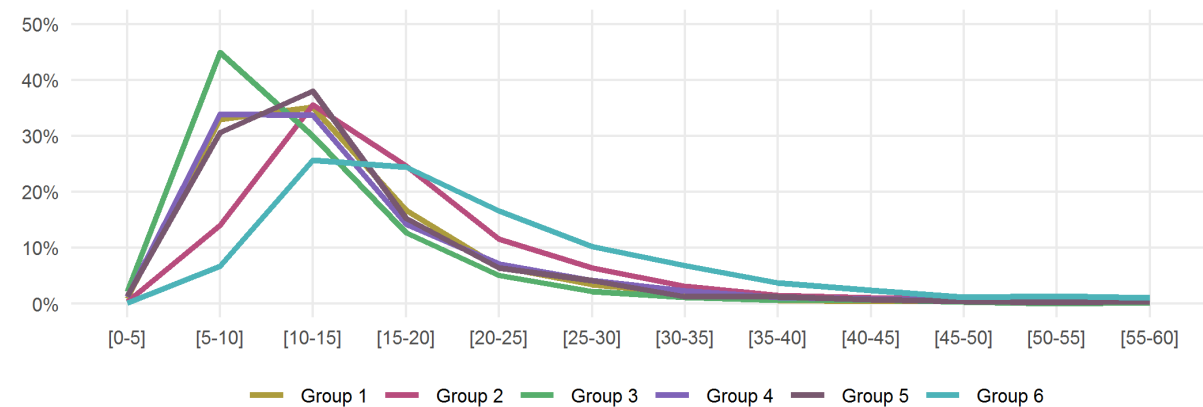
Figure 2: Share of mobile users by gender and age



TIME USAGE

The average respondent used 15.6 minutes to complete the questionnaire. Measuring average time usage poses a challenge, in that 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 consisted of five subsets of questions in addition to questions posed to all respondents. The three subsets of questions were posed to six subgroups of respondents. Group 1-5 is determined by random allocation, while group 6 was reserved for newly recruited respondents.

The time usage of the different groups varies between 12 and 21 minutes. Respondents in group 6 clearly spent the most amount of the time to complete the survey on average, while the respondents in group 3 clearly spent the shortest amount of time.

Table 4: Average time usage (minutes) in each subset

	All	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
All users	15.6	14.0	17.2	12.3	14.2	14.1	20.9
Non-mobile users	16.7	14.9	18.1	13.1	15.5	15.2	22.2
Mobile users	14.3	12.9	16.2	11.4	12.9	12.9	19.4

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 twenty-one, 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 25.

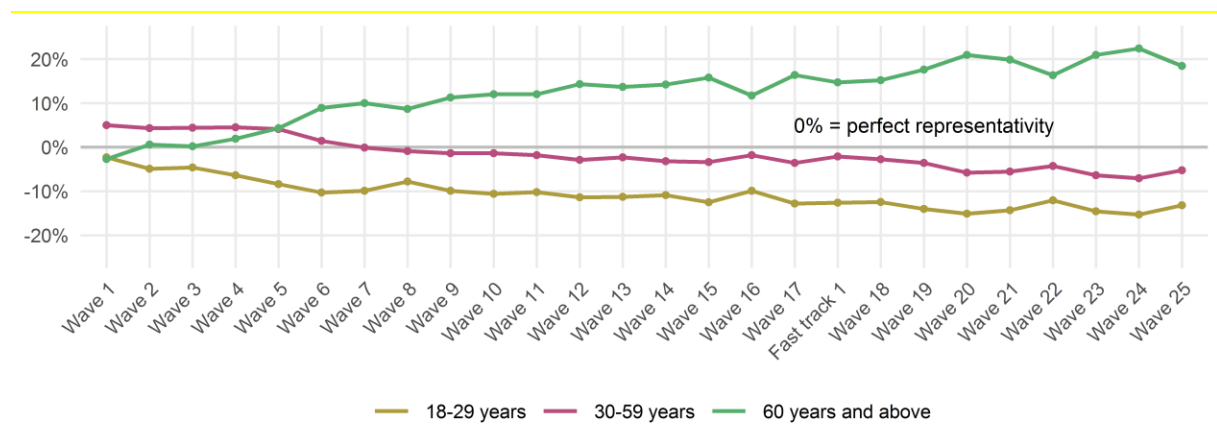
Table 5: Age distribution in the population and the net sample

	18-29 years	30-59 years	60 years and above
Population	19.7 %	50.4 %	29.8 %
Net sample	6.5 %	45.2 %	48.3 %

From the age distribution presented in table 5, we see that 18-29 year olds are underrepresented in the net sample of wave 25. The age group 30-59 years in the net sample is underrepresented compared to the distribution in the population, while respondents aged 60 years and above are overrepresented by some margin.

Over time, we have observed a drift away from perfect representativity of age groups (figure 4). While the oldest respondents started out as underrepresented in wave one, they 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. We note that the rate of misrepresentation of age groups peaked with wave 24. Mostly due to newly recruited respondents and partly because of higher participation in wave 25, the representativity statistics is somewhat improved in wave 25.

Figure 4: Representativity of age groups



In table 6, the population and net samples are broken down 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 older men are more overrepresented than women in the same age bracket.

Table 6: 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	10.1 %	9.5 %	25.8 %	24.6 %	14.2 %	15.6 %
Net sample	2.7 %	3.8 %	21.1 %	24.1 %	26.3 %	22.0 %

The inclusion of educational level in table 7 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 7: 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	3.8 %	2.8 %	0.2 %	0.2 %
Upper secondary education		4.1 %	3.1 %	1.4 %	1.6 %
University/university college		2.3 %	3.6 %	1.1 %	1.9 %
No education/elementary school	30-59 years	5.3 %	4.3 %	0.7 %	0.6 %
Upper secondary education		11.0 %	7.7 %	7.1 %	5.3 %
University/university college		9.5 %	12.7 %	13.5 %	18.1 %
No education/elementary school	60 and above	3.2 %	4.4 %	1.9 %	1.3 %
Upper secondary education		7.1 %	7.4 %	8.9 %	6.9 %
University/university college		4.0 %	3.9 %	15.6 %	13.6 %

Respondents who have completed upper secondary education as their highest completed 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

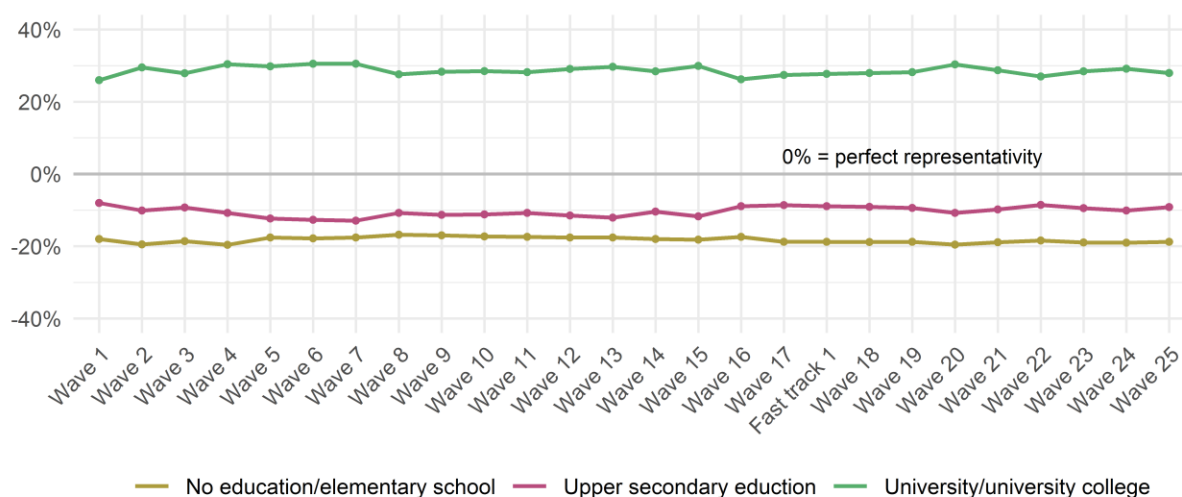


Figure 5 illustrates the representation of education groups since wave one. The general trend is that the highly educated are overrepresented compared to those with less or no education. Except for slight improvements in representativity of the education groups when new respondents are recruited (wave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

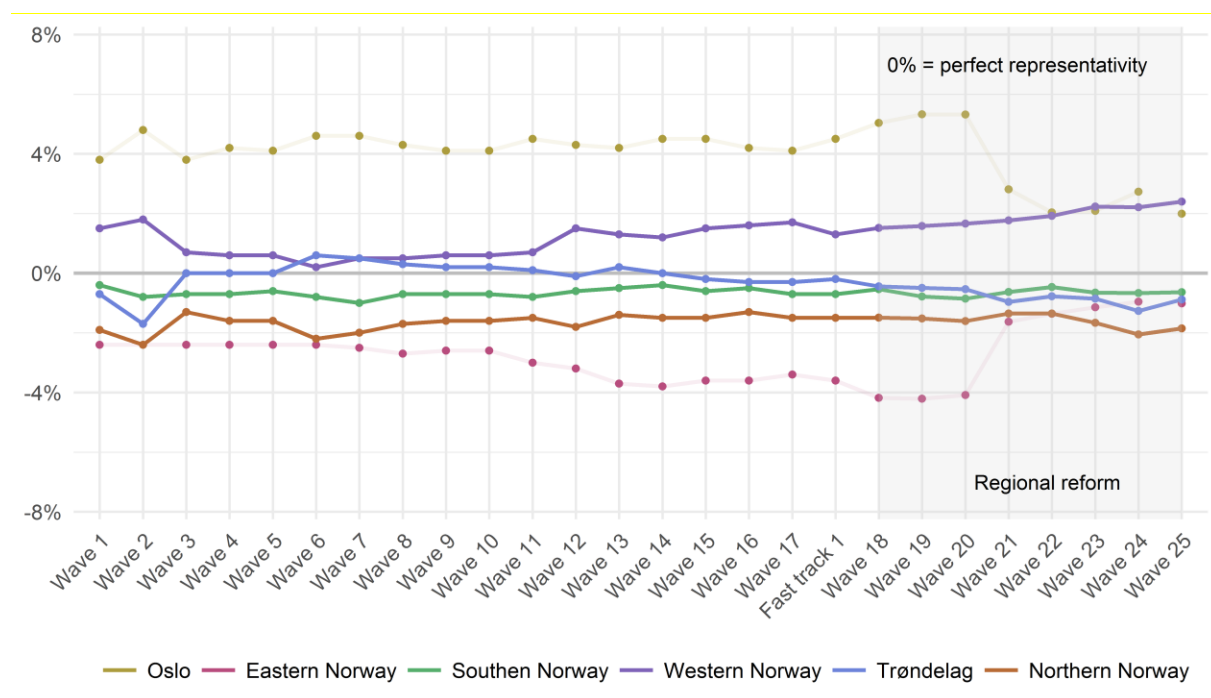
With regard to geography, (table 8) we observe that the representation of panel members living in Trøndelag, Eastern Norway, and Southern Norway are nearly on level with the population, while Oslo and Western Norway is overrepresented. Respondents from Northern Norway meanwhile are underrepresented.

Table 8: 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	1.5 %	1.6 %	3.1 %	0.4 %	0.7 %	1.2 %
	30-59 years	3.8 %	3.5 %	7.3 %	3.7 %	4.4 %	8.1 %
	60 and above	1.3 %	1.5 %	2.8 %	3.0 %	2.9 %	5.9 %
	In total	6.5 %	6.6 %	13.2 %	7.1 %	8.0 %	15.1 %
Eastern Norway	18-29 years	3.4 %	3.1 %	6.5 %	0.7 %	1.3 %	2.0 %
	30-59 years	9.7 %	9.4 %	19.1 %	6.9 %	8.4 %	15.3 %
	60 and above	5.8 %	6.4 %	12.2 %	10.6 %	8.8 %	19.4 %
	In total	18.8 %	18.9 %	37.8 %	18.2 %	18.5 %	36.7 %
Southern Norway	18-29 years	0.6 %	0.5 %	1.1 %	0.1 %	0.2 %	0.3 %
	30-59 years	1.4 %	1.4 %	2.8 %	1.0 %	1.2 %	2.3 %
	60 and above	0.8 %	0.9 %	1.7 %	1.3 %	1.2 %	2.4 %
	In total	2.8 %	2.8 %	5.6 %	2.4 %	2.6 %	5.0 %
Western Norway	18-29 years	2.6 %	2.4 %	5.0 %	0.9 %	0.9 %	1.9 %
	30-59 years	6.6 %	6.2 %	12.8 %	6.3 %	6.3 %	12.6 %
	60 and above	3.6 %	3.9 %	7.5 %	7.3 %	5.8 %	13.1 %
	In total	12.8 %	12.5 %	25.2 %	14.6 %	13.0 %	27.6 %
Trøndelag	18-29 years	1.1 %	1.0 %	2.1 %	0.3 %	0.4 %	0.7 %
	30-59 years	2.2 %	2.1 %	4.3 %	1.8 %	2.0 %	3.7 %
	60 and above	1.3 %	1.4 %	2.6 %	2.0 %	1.6 %	3.6 %
	In total	4.6 %	4.4 %	9.0 %	4.1 %	4.0 %	8.1 %
Northern Norway	18-29 years	1.0 %	0.9 %	1.9 %	0.2 %	0.3 %	0.5 %
	30-59 years	2.2 %	2.1 %	4.3 %	1.3 %	1.9 %	3.2 %
	60 and above	1.5 %	1.5 %	3.0 %	2.1 %	1.7 %	3.8 %
	In total	4.7 %	4.5 %	9.2 %	3.7 %	3.8 %	7.5 %

Respondents aged 60 years and above are overrepresented in all parts of the country, especially in Oslo and Western Norway. Conversely, young people aged 18-29 years are underrepresented in all regions.

Figure 6: Representativity of regions



For wave twenty-one, 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 analyses⁶, 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.⁷ 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. 1.1 percent of the 25th wave net sample have not answered the question about level of education. Because of this, two 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).

The variables have the following categories:

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

⁶ Note that Oslo (including Akershus), and Eastern Norway diverge in wave eighteen, due to the regional reform implemented 1st of January 2020.

⁷ The applied formula for weight w_i for element i , in strata h is: $w_i = \frac{N_h/N}{n_h/n}$

Consequently, the weights might be less precise for some sub-groups. Note that the dataset is provided with necessary information⁸ 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 provides the most accurate compensation for the various sources of bias in the net sample.

Table 9 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 9: 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	4.9 %	23.5 %	23.7 %	-18.8	-0.2
Upper secondary education	31.2 %	40.5 %	40.3 %	-9.1	0.1
University/university college	63.9 %	36.1 %	35.9 %	27.9	0.1

⁸ See columns r25_Weight1stratapop and r25_Weight2stratapop

Table 10: Weights applied to different strata (weight 2)

		Men	Women			Men	Women		
Oslo	18-29 years	No education/elementary school	7.7	Western Norway	18-29 years	No education/elementary school	10.8	9.1	
		Upper secondary education	10.2			9.3	Upper secondary education	6.1	5.9
		University/university college	2.2			3.1	University/university college	1.2	3.3
	30-59 years	No education/elementary school	3.7		1.7	30-59 years	No education/elementary school	2.2	2.1
		Upper secondary education	1.5		1.4		Upper secondary education	1.3	1.4
		University/university college	0.8		0.8		University/university college	0.7	1.0
	60 and above	No education/elementary school	2.4		1.7	60 and above	No education/elementary school	1.9	2.0
		Upper secondary education	0.7		0.6		Upper secondary education	0.6	0.7
		University/university college	0.3		0.3		University/university college	0.2	0.3
Eastern Norway	18-29 years	No education/elementary school	25.8	16.8	Trøndelag	18-29 years	No education/elementary school	58.5	7.5
		Upper secondary education	8.0	9.2			Upper secondary education	10.6	5.9
		University/university college	1.8	3.0			University/university college	1.5	6.5
	30-59 years	No education/elementary school	3.6	1.9		30-59 years	No education/elementary school	2.4	2.3
		Upper secondary education	1.6	1.5			Upper secondary education	1.7	1.8
		University/university college	0.8	1.1			University/university college	0.9	1.3
	60 and above	No education/elementary school	2.4	1.7		60 and above	No education/elementary school	1.7	1.5
		Upper secondary education	0.8	0.7			Upper secondary education	0.7	0.7
		University/university college	0.2	0.3			University/university college	0.3	0.3
Southern Norway	18-29 years	No education/elementary school	27.5	21.7	Northern Norway	18-29 years	No education/elementary school	16.7	12.3
		Upper secondary education	18.9	8.7			Upper secondary education	5.6	5.1
		University/university college	3.6	3.9			University/university college	1.8	3.6
	30-59 years	No education/elementary school	8.5	1.5		30-59 years	No education/elementary school	3.6	2.1
		Upper secondary education	1.5	1.4			Upper secondary education	2.3	1.3
		University/university college	1.1	1.4			University/university college	0.9	1.4
	60 and above	No education/elementary school	3.1	3.0		60 and above	No education/elementary school	2.4	3.4
		Upper secondary education	0.8	0.8			Upper secondary education	0.8	0.8
		University/university college	0.3	0.3			University/university college	0.3	0.3