Statistisk sentralbyrå
Statistics Norway

Weighting methodology for the Norwegian Labour Force Survey from 2021 onwards

TALL SOM FORTELLER

Melike Oguz Alper

NOTATER / DOCUMENTS

2023/51

In the series Documents, documentation, method descriptions, model descriptions and standards are published.

Published: 17 November 2023

ISBN 978-82-587-1847-2 (electronic)
ISSN 2535-7271 (electronic)

Symbols in tables	Symbol
Category not applicable Figures do not exist at this time, because the category was not in use when the figures were collected.	.
Not available Figures have not been entered into our databases or are too unreliable to be published.	..
Confidential Figures are not published to avoid identifying persons or companies.	:
Decimal punctuation mark	.

Preface

This document provides the background of the revision that was made in 2021 in the weighting methodology for the Norwegian Labour Force Survey (NLFS), and practical and technical details behind the calculation of monthly and quarterly individual weights and yearly household weights. This revision was required due to several changes occurred in the target population, sampling design, and data collection in 2021, as well as the introduction of a new European legislation for the European Union LFSs.

The previous review in the weighting methodology for the NLFS was carried out following the availability of better quality of register data from 2015, namely A-ordningen. A new weighting methodology was developed as a result of this review that was used as the main method from 2018 until 2021. The NLFS statistics were re-produced back to 2006 to compare the new series against those obtained from the previous methodology developed in 1996.

The project was carried out with the collaboration of Division for Labour Market and Wage Statistics, Division for Methods, and colleagues provided IT-related support.

The author of this document wishes to thank to all those involved in this collaboration, particularly, Camilla Schjetlein Sundt, Håvard Hungnes Lien, Jørn-Ivar Hamre, Erik Herstad Horgen, Kristian Lønø, Anita Cecilie Dal, and Susie Jentoft. Many thanks go also to Xeni Kristine Dimakos, the head of the Division of Methods, for their helpful comments on the first draft of this document.

Statistics Norway, 7 November 2023

Arvid Olav Lysø

Abstract

A new revision in the weighting procedure of the Norwegian Labour Force Survey since the last one made in 2018 was required due to the change in the target age group for employment statistics, the introduction of the new sampling design in 2021 and the new weighting requirements with the new EU legislation for labour force surveys within the EU. Compared to the revision in 2018, the main weighting method via model-calibration has remained unchanged for the age group $15-74$. However, significant changes occurred in the calculation of selection probabilities, which is not straightforward, and initial weights to calibration due to the new sampling design, changes in the data structure and the new weighting requirements by the new EU legislation. A linear calibration method was introduced for those within age groups $0-14$ and $75+$. In order to fulfil the individual- and household-level requirements simultaneously and to ensure same weights for all individuals within any given household, an integrative calibration approach was introduced to be used for the calculation of yearly household weights. Several models were evaluated in this respect, and a final model was chosen by considering robustness and gain in precision.

Contents

1. Introduction

There have been several changes in the Norwegian Labour Force Survey (NLFS) since 2021 (Horgen et al., 2023) in terms of target age group, sampling design, and data collection to comply with the new legislation of the EU LFS (EC, 2019). The key changes with the new legislation are summarised as follows.

The target age group has been changed from $15-74$ to $15-89$ for employment statistics. Besides, the new legislation requires producing weights for all individuals in the sample regardless whether or not their ages are within the target age group 15-89.

A new sampling design for the NLFS has been implemented since 2021 (Jentoft, 2022). The key change in the sampling design happened by selecting individuals directly from a list of individuals in the target population of the NLFS rather than selecting *families*. New stratification variables and a new allocation method of the total sample size to strata have also been introduced with the new sampling design to increase the precision of the estimators of the key statistics.

Proxy interviews have not been used in the data collection of the NLFS since 2021 (Jentoft, 2022). Hence, sample persons answer only for themselves. This may increase non-response rates. Another key change is to introduce the use of *household* concept in alignment with the one used in the EU LFS to produce household statistics, which was not possible to deliver from the NLFS before 2021. Subsampling is implemented for the data collection at household level. A subsample consists of individuals directly selected with respect to the new sampling design and all their corresponding household members. Hence, additional household members will be in the sample via the corresponding *reference person* who is defined to be an adult aged above 14 years and being selected into the sample directly, not via other persons. The calculation of the inclusion probabilities of households will be more complicated than the one that was used for the NLFS before 2021.

These changes have brought about a necessity for revision of the weighting methodology that was applied until 2021 (Oguz-Alper, 2018). This document summarises all the aspects of the monthly and quarterly individual and yearly household weighting procedures that have been applied to the NLFS since 2021.

In Section 2, previous and current sampling designs are briefly presented. In Section 3, weighting methods that were used until 2021 are briefly described. The current individual weighting procedure that has been in use since 2021 is described in detail in Section 4. Calculation of household weights that are used for the production of annual household statistics is explained in detail in Section 5. Finally, Section 6 addresses the revisions made in the weighting for the NLFS and the primary reasons of those revisions once again, and provides an insight for possible future investigation that could result in a new revision in the weighting of the NLFS.

2. Sampling designs between 1996-2020 and from 2021

A nice overview of the history of survey designs of the NLFS including the one currently implemented is given by Jentoft (2022). In the following sections, the sampling designs between 1996-2020 and the one implemented from 2021 are briefly described.

2.1. Between 1996-2020

The sampling design that was used for the NLFS between 1996-2020 is a stratified one-stage cluster sampling (e.g. Hamre and Heldal, 2013), where strata were formed based on NUTS III regions (e.g. 18 counties according to the regional classification used from 2018 to 2020 (see here: The municipality reform and regional changes 2020). [Online; last accessed 6 November 2023]). *Nuclear families*, which are defined as the people living at the same residential address whom are married or registered partners, along with their children or parents, were the clustering units. All individuals belonging to sample families were taken into the sample. Data for labour market were collected for all sample adults aged 15-74 years at the time of the data collection.

Total sample size of clusters was disproportionally allocated to counties as such that lesser and more populated counties were, respectively, given larger and smaller sample sizes. The ratio of sampling fraction to the overall sampling fraction for each county according to the regional classification (19 counties) until 2018 is given by (e.g. Vedø and Rafat, 2003, p.7). According to the changes in the regional classification in 2018, counties Sør-Trøndelag and Nord-Trøndelag were merged into a single county called Trøndelag, thus leading to 18 counties in total, and the corresponding ratio was set to 1. The number of counties were reduced from 18 to 11 following the new regional classification being used in 2020. The ratios of sampling fractions by the new 11 counties that were used from the second quarter of 2020 to the fourth quarter of 2020 are given by Table A1.1.

The NLFS is a *rotating panel survey*. Sample units have been followed up for two years, equivalently for eight quarters, since 1996. A new panel, which shall be called *wave sample*, is selected each quarter while the panel which has been in the sample for eight quarters rotates out. Therefore, there is a 87.5% overlap between any two consecutive quarters which leads to better precision for estimation of quarterly change statistics. A sample of any quarter consisted of eight waves that included around $12\,000$ families yielding around $24\,000$ individuals. Families selected are randomly distributed over 13 reference weeks in a given quarter. Hence, the sample size of individuals for each month was around $24\,000 * 4/13 = 7\,385$ or $24\,000 * 5/13 = 9\,231$, depending on how many reference weeks, four or five, there were (e.g. Hamre and Heldal, 2013, p.9).

2.2. From 2021

The new sampling design that has been implemented since 2021 is a stratified simple random sampling, where strata are formed based on NUTS II (6 regions by the new regional classification from 2021 (see Table A1.1)), age groups (15-24, 25-54, 55-66, 67-74, and 75-89) and register based labour market status (employee, unemployed, and other) (Jentoft, 2022, p.12). Note that not all these categories are cross-classified in the construction of strata (Jentoft, 2022, p.22). Sampling frame consists of all individuals aged within 15-89 years. This is due to the change in the definition of population for which information about whether, or not, being employed is collected. A supplementary sample of 15 years olds is selected each quarter and added to the sample of each wave (waves 2-8) due to the attrition in the sample of this age group over time (Jentoft, 2022, p.16).

The total sample size of individuals and the rotation scheme remains unchanged. Sample is allocated to strata using a multivariate optimal sampling allocation method (Bethel, 1989), (see Jentoft (2022) for details). A new wave sample of individuals is uniformly distributed to 13 weeks in a given quarter. All households members, regardless of their ages, whom are linked to the reference sample persons are taken into the sample. However, labour force information is only collected for reference sample persons aged 15-89 years per the reference week for the production of monthly and quarterly labour force statistics.

A set of core questions are asked to all sample reference persons aged $15-89$ years while questions regarding annual, biannual and adhoc modules are asked to only those in waves 2 or 6. Each adult household member aged above 14 years corresponding to reference sample individuals in wave

2 is also invited to participate into the survey (e.g. Horgen et al., 2023, p.25). Household members are specified using a new household register which was established in 2021. The main purpose of collecting data from a sub-sample of household members is to produce annual household statistics such as number of jobless households.

3. Calculation of survey weights until 2021

The weighting methodology that was developed before the one in 2018 was implemented from 1996 until 2018 (e.g. Zhang, 1998; Heldal, 2000; Hamre and Heldal, 2013). Under this methodology, survey weights for the NLFS were calculated in two steps. Following the revision in 2018 (Oguz-Alper, 2018), a one-step approach was implemented for the calculation of the survey weights for the NLFS. This weighting methodology was also implemented in parallel to the old one from 2006 until 2018 to compare the two series obtained from these two weighting methodologies (see e.g. New method gives better LFS figures. [Online; last accessed 6 November 2023]). From 2018 until 2021, only the one developed in 2018 was used for the NLFS. Whether the weighting approach involves one–step or two–steps depends on if the initial weights, which are often the design weights, the inverse of the selection probabilities, are adjusted in a separate step before the calculation of final weights via calibration (e.g. Deville and Särndal, 1992). There are several papers discussing one–step and two–step approaches (e.g. Lundström and Särndal, 1999; Andersson and Särndal, 2016; Haziza and Lesage, 2016), without a consensus about a choice between them.

For a given data set, several plausible methods should be cross-examined in terms of nonresponse bias and variance reduction by using a reference estimator that does not use any auxiliary information neither for nonresponse adjustment nor calibration purposes to increase efficiency. Decision may be made in favour of the one providing a limited loss of efficiency and avoiding spurious adjustments in point estimates (Nguyen and Zhang, 2020). The revision in the estimation method in 2018 was mainly resulted from such an evaluation (Oguz-Alper, 2018) that was carried out to make a better use of administrative data, namely A-ordningen (see e.g. About A-ordningen. [Online; last accessed 6 November 2023]), which were made available from 1 January 2015 with the collaboration of the Norwegian Labour and Welfare Administration (NAV), the Norwegian Tax Administration, and Statistics Norway (SSB).

Oguz-Alper (2018) examined the generalised regression (GREG) (e.g. Cassel et al., 1976; Deville and Särndal, 1992) and (multiple) model–calibration (e.g. Wu and Sitter, 2001; Montanari and Ranalli, 2009) estimators with one–step and two–step variations as well as a post-stratified estimator by using the NLFS data from the first quarter of 2015 to the first quarter of 2017. The estimators were evaluated in terms of point and variance estimates against those obtained from a reference estimator that does not use any auxiliary variable. The methods were tested for the core variables, such as number of employed and unemployed people, number of people outside of labour force, and unemployment rate overall and over domains, such as gender and age groups. It was observed that the two–step approach provided higher variances, in general, without further improvement in reduction of nonresponse bias compared to the one–step approach regardless of that which estimator, GREG, MMC, or PS, was used. Overall, the MMC estimator provided lowest variance estimates among all the others for a given approach, one-step or two–step. Therefore, in the production of the quarterly NLFS statistics, it was decided to use an MMC method to calculate the survey weights.

Calibration is applied to monthly survey data as a result of which monthly weights are obtained. Quarterly statistics may be calculated as a weighted average of monthly statistics. Under the methodology developed in 1996, all three months in a given quarter were treated equally, and thus, the quarterly weights were equal to the monthly weights divided by three. Following the revision in 2018, quarterly estimates of totals were obtained as a weighted average of monthly estimates where weights were defined to be proportional to the number of calender weeks, four or five, in the corresponding months. Hence, monthly weights were multiplied by either $4/13$ or $5/13$ depending on the number of weeks to produce quarterly weights. The calculation of the survey weights until 2021 are briefly presented in the following sections. In Section 3.1, the one developed in 1996 and used until 2018 is discussed. In Section 3.2, the weighting methodology developed in 2018, which was used from 2006 until 2021, is presented. As noted above, the two methodologies were implemented between 2006 and 2018.

3.1. From 1996 to 2018

The estimation methodology developed in 1996 involves a two-step regression estimation method. In the first step, design weights were adjusted within weighting cells which were formed based on five-year age groups from 15 to 74 years, gender, register based employment by industry groups (employee by primary, secondary, or tertiary industries, and not employed). The weighting cells may also be called *response homogeneity groups* (RHG) since the response probabilities are assumed to be homogeneous within these cells. There were $12*2*4 = 96$ RHGs in total. In the second step, a calibration procedure was implemented within each county, as such that the sample estimates for the marginal counts by gender, 12 age groups and four register based employment groups were made to be consistent with the population counts. A slightly modified version of this estimation procedure (e.g. Oguz-Alper, 2018, p.10) was implemented in the production of the NLFS until 2018 (see (Heldal, 2000) for details as well as the information about the accompanying program). As mentioned above, quarterly weights were obtained by dividing monthly weights by three.

3.2. From 2006 to 2021

The revision in 2018 brought about using a one–step MMC approach in the production of the NLFS (Oguz-Alper, 2018). This approach has been implemented since 2006, and it has been used as the main method for the calculation of the survey weights since 2018. The customary GREG estimator (e.g. Deville and Särndal, 1992) assumes a linear relationship between the variable of interest, y-variable, and the explanatory variables, $\mathbf{x}$. However, this assumption may not always hold. For example, it would not be valid when the y-variable is categorical, which is the case for most of the variables in the NLFS. A *model–calibration* (MC) approach (e.g. Wu and Sitter, 2001), where the relationship between y and $\mathbf{x}$ is defined via a *generalised linear model* (GLM), may perform better than the GREG in terms of efficiency if the model is correct.

The *multiple model calibration* (MMC) approach Montanari and Ranalli (2009) is an extension of the MC approach (Wu and Sitter, 2001) that was proposed to remedy some of the drawbacks of the MC. For example, it may not be possible to ensure the consistency between sample estimates and population frequencies by domains, such as gender, age–groups, etc., if the working model is not a linear one. Besides, the standard MC approach requires fitting a separate model for each of the y-variable that leads to more than one set of survey weights, which is undesirable in practice.

Up-to-date monthly statistical register data, which were formed based on several administrative registers including population register, Aa-registeret (Arbeidsgiver- og arbeidstakerregisteret: State register of employers and employees. [Online; last accessed 6 November 2023]) (see also Villund (2009, ch.2.2)) and A-ordningen were used for weighting purposes. Register based employment, which is one of the variables used in the weighting process, was collected from the Aa-registeret before 2015, and this information has been collected from the A-ordningen since 2015. Register files used for estimation differ from those used for the sample selection. Population files included only individuals at the age of 15-74 years since weighting outside of this age group was not required. Individuals that were not available in the population files were excluded from the corresponding monthly sample files. Since a one–step approach was used, calibration was directly applied to the respondent set of individuals.

Monthly individual weights were obtained by minimising a distance function (Deville and Särndal, 1992) under a set of calibration constraints. Calibration weights were trimmed, as such that the minimum and the maximum values of weights would not be lower or higher than 40 and $8\,000$, respectively. Weights outside of the bounds were set to these bounds, and calibration was implemented once again by using a new set of weights as initial weights. The resulting calibration weights were checked if they were within the bounds, and so on. This iterative process was carried out until all weights were within the limits and the calibration constraints were satisfied.

Quarterly statistics were calculated as a weighted average of monthly statistics (Oguz-Alper, 2018), where weights were proportional to the number of weeks of the corresponding months. Yearly average individual-level statistics were obtained as a simple average of the quarterly statistics.

Because one–step MMC has still been used as the main weighting method for the NLFS for the age group $15-74$, methodological details are provided in Section 4.4 rather than here to avoid repetition. The details about the weighting procedure including the calculation of initial weights, which is different from the one described in Section 4.2, for the period of 2006-2020 can be found in Oguz-Alper (2018).

4. Monthly and quarterly individual weighting from 2021

The estimation methodology that has been used as the main method since 2018 has not been changed for the age group 15-74 in general. However, the changes made in the NLFS from 2021 (see Section 1) and the new requirement of producing weights for those outside of the age group 15-74 years have brought about some challenges in the weighting procedure.

At first, the transition from the previous sampling design to the new one (Jentoft, 2022), which shall be called *family sampling* and *person sampling* from now on, respectively, was done gradually, as such that the new panels, or in other words, *wave 1* samples, were selected with respect to the person sampling starting from the first quarter of 2021. Thus, in this quarter, the sample consisted of wave 1 selected from under the new design and waves 2–8 selected from the old sampling design. The sample data has included only those coming from the new sampling design since the fourth quarter of 2022 (see Figure A1.1). This has required to identify from which sampling design the sample units were selected and to calculate the initial weights accordingly.

At second, unlike the family sampling, household members linked to reference persons are not interviewed unless they are in wave 2. However, regardless of which wave they belong to, all household members aged $0-14$ and $90+$ whom are taken into the sample via corresponding reference persons are used in the weighting procedure in order to fulfil the consistency requirements (EC, 2019) for the population distribution by these age groups. Besides, it is desirable to have equal weights for all individuals belong to the same household regardless of whether, or not, they are reference persons. Hence, it is not straightforward to calculate the selection probabilities as well as the initial weights, which may not necessarily be the design weights.

At third, data is collected for all household members linked to reference persons, if they are in wave 2, to produce annual household statistics. Since proxy interviews are not allowed in the NLFS any longer, not all household members including reference persons respond to the survey which may lead to partial response or nonresponse for any household. This has brought about the question under which situations households would be used for inference.

All these challenges have been taken into account in the construction the new weighting procedure which is described in detail in the following sections. The notation used is provided in Section 4.1.

4.1. Notation

Let U be a population of size N that consist of people aged $15-89$ years or aged $90+$ years and living alone in the scope of the NLFS. Let U be divided into H disjoint strata as such that $U = \cup_{h=1}^{H} U_h$ and $N = \sum_{h=1}^{H} N_h$, where N_h is the population size of stratum h. In practice, two separate sampling frames are used to select samples correspondingly from the sub-population of people aged $15-89$ and the sub-population including people aged $90+$ years and living alone. A random sample of people aged 90 years and over living alone are selected directly since people in this sub-population cannot be in the sample via any other person but themselves.

For simplicity in notation, we shall not use different notations for these sub-populations since both samples are selected with stratified simple random sampling designs, and expressions regarding the calculation of selection probabilities are valid for both. Stratification variables used for these sub-populations are not exactly the same. Region at NUTS II level, register-based labour market status and age are the variables used in the stratification for population of people aged $15-89$ years. For the other sub-population, region at NUTS II level and gender are used in the stratification. In order to maintain the same notation, we shall suppose that being in one sub-population or in the other one is also used in the stratification of U as such that U is first divided into two upper strata each of which is divided into further strata, *design strata*, based on the stratification variables corresponding to each sub-population. Hence, any stratum h, hereafter, will correspond to a combination of one upper stratum and a corresponding design strata.

Let Ω be a population of households. Once a person of age $15-89$ years old in U is selected, all individuals that are linked to sample reference persons due to having the same household membership are taken into the sample regardless of data being collected for all, or not. We assume that each household in Ω corresponds to at least one individual in U. Let β_κ denote all individuals in U that are linked to household $\kappa \in \Omega$. We also assume that each individual in U is linked to one and only one household in Ω. Notice that a household may be linked to individuals from different design strata, but not from different upper strata. Let $\beta_{h\kappa}$ include all individuals in stratum h that are linked to household κ, as such that $\beta_\kappa = \cup_{h=1}^{H} \beta_{h\kappa}$.

A sample of individuals is selected according to a stratified random sampling from population U for the NLFS each quarter. The sample set of individuals newly selected in a given quarter is called wave 1 sample. Let n be the sample size for wave 1 as such that $n = \sum_{h=1}^{H} n_h$, where n_h is the size of the stratum sample selected from U_h with simple random sampling. For each wave of 2–8, an extra sample of individuals aged 15 of size n_{15} is selected with simple random sampling from a population of individuals aged 15 years $U_{15} \subset U$ of size N_{15} to compensate for the attrition due to individuals aged 15 years when selected in wave 1 and being older than 15 in later waves. In practice, at first, a random sample for wave 1 is selected, and then an extra sample of size $7 * n_{15}$ is selected from $U_{15} \setminus s_{0;15}$, where $s_{0;15}$ is the sample of 15-year old individuals selected in wave 1. Let $\beta_{\kappa;15} \subseteq \beta_\kappa$ be a set of 15-year old individuals in U linked to household κ. Let s_0 be the sample that includes both wave 1 sample and any one of the extra samples of 15-year old individuals selected for waves 2–8. Let $\Omega_s \subseteq \Omega$ be a set of sample households added via sample individuals in s_0 and s be the sample of all individuals linked to sample households in Ω_s. Let $a_{i\kappa}$ be a household membership indicator as such that $a_{i\kappa} = 1$ if individual i is a member of household κ, and $a_{i\kappa} = 0$ otherwise. Notice that the sum of $a_{i\kappa}$ over household members for any given household will be greater than or equal to $|\beta_\kappa|$. This is because of that β_κ includes only members in U which does not contain members aged $0-14$, or aged above 89 and not living alone, whereas $a_{i\kappa}$ concerns all members regardless of age.

4.2. Selection probabilities

The sample of the NLFS consisted of both family and person samples until the fourth quarter of 2022, and thus two ways were followed until this time to calculate the selection probabilities given how the unit of interest was selected. Selection probability of any individual i who is a member of household κ is determined by the selection probability of κ regardless of the design. This is mainly due to allowing to use same initial weights in individual and household weighting in the production of different statistics with different frequencies. More detailed explanations are provided in Section 4.2. Let π_κ be the selection probability of household κ defined by

$$\pi_\kappa = \Pr(\kappa \in \Omega_s)\cdot \tag{4.1}$$

Calculations of π_κ under family and person sampling are explained in detail in Sections 4.2 and 4.2, respectively. The selection probability of any individual i is given by

$$\pi_i = \Pr(i \in s) = \pi_\kappa, \text{ given } a_{i\kappa} = 1\cdot \tag{4.2}$$

Family sampling

Exact calculation of selection probabilities was not possible due to some operational reasons such as difficulties in getting the number of family population sizes by strata (i.e. NUTS III) by sample selection dates and building transformation links between the old and the new regional classifications. Besides, all household members linked to sample families were taken into sample which made the calculation of selection probabilities more infeasible due to the lack of linkage between families and households before 2021. Therefore, proxy selection probabilities were calculated for all those selected under the family sampling. The total number of waves that was selected by the family sampling was seven in the first quarter of 2021.

Let $\tilde{n}$ and $\tilde{N}$ be the total number of households in the sample selected under the family sampling design and the household population, respectively. Both $\tilde{n}$ and $\tilde{N}$ are fixed for all those selected by the family sampling. The statistical household register data by mid-November 2021, when the family sampling was used for the last time for the NLFS, was used to obtain the number of households $\tilde{N}$. Here, only households that had individuals aged $15-74$ years were included in the register data used since the target population for the family sampling consisted only of individuals aged $15-74$ years. Let κ be a household linked to a reference person i in stratum g belonging to a sample family, usually the oldest person in the family among those aged $15-89$ years, as such that $i \in \beta_\kappa$. The proxy probability of household κ to be included in one of the seven waves under the family sampling is given by

$$\pi_\kappa = R_g \frac{\tilde{n}}{\tilde{N}} \frac{1}{7}, \tag{4.3}$$

where the R_g are the ratios between strata sampling fractions and the overall sampling fractions under the family sampling the values of which are given in Table A1.1. Here, the subscript g refers to the strata used in the family sampling, which are counties (NUTS III regions) given in Table A1.1. The rational behind (4.3) is that the sampling rates R_g for any wave are almost the same as those that could have been obtained if households were selected as the sampling units rather than families. Notice that the rates by strata that were used in the selection of families might be different from the values of R_g given in Table A1.1 if the selection date is before the second quarter of 2020, when the regional classification introduced in 2020 was used for the NLFS for the first time. Nevertheless, the actual rates were not used in the initial weighting for family sampling because of having only the regional classification dated back to 2020 in the sample and population data sets used for the weighting methodology from 2021.

Person sampling

Several statistics at different frequencies are produced from the NLFS. All waves are used in the production of individual-based monthly, quarterly and yearly average statistics while only waves 2 and 6 are used for the production of annual, biennial, and eight-yearly structural variables, and only wave 2 is used for annual household statistics (e.g. Horgen et al., 2023, pp.24-25). Selection probabilities can be calculated separately for individuals and households depending on when they will be used for inference. Other household members aged $15-89$ years are not used in the estimation of individual-based statistics, and data is collected for such members only if they are in wave 2. This brings about the use of different sets of initial weights, which are essentially the inverse of the selection probabilities, for the productions of individual and household statistics. Besides, individual and household weights will only differ from each other for those whom are selected directly since other household members are only selected via sample reference persons, and thus they will take the same probability as the household they belong to in any case. This means that any two individuals in the same design strata may get different selection probabilities just because one of them is selected directly and the other one is selected via a sample reference person in the same household. To avoid such differences as well as for practical reasons, it has been decided to calculate the household selection probabilities, and to use these probabilities also as individual selection probabilities as such that all members in a given household take the same selection probability as the household itself. In this way, it is ensured that the same initial weights are used for all household members in any given household.

A household $\kappa \in \Omega$ is selected if at least one household member $i \in \beta_\kappa$ is selected in s_0. Thus the selection probability of household κ can be written as follows.

$$\pi_\kappa = 1 - \bar{\pi}_\kappa,$$

where $\bar{\pi}_\kappa$ is the *exclusion probability* of household κ defined by

$$\begin{aligned} \bar{\pi}_\kappa &= \Pr(\text{none of the individuals in } \beta_\kappa \text{ selected in } s_0) \\ &= \binom{N_{15} - \hat{n}_{15} - |\beta_{\kappa;15}|}{n_{15}} \binom{N_{15} - \hat{n}_{15}}{n_{15}}^{-1} \prod_{h=1}^{H} \left\{ \binom{N_h - |\beta_{h\kappa}|}{n_h} \binom{N_h}{n_h}^{-1} \right\}, \end{aligned} \tag{4.4}$$

where $|\bullet|$ denotes the size of a set $\bullet$, $\binom{a}{x} = a!/\big(x!(a-x)!\big)$ is the combination operator, and $\hat{n}_{15} = \sum_{h=1}^{H} n_h N_{h;15}/N_h$, where $N_{h;15}$ is the number of 15-year old individuals in U_h, is the expected number of 15-year old individuals to be selected in wave 1 sample by the stratified simple random sampling described in Section 4.1. When both a and x values are large as such that the number of sets of size x with distinct values from a is too large (e.g. 10^{308} in statistical software R (R Core Team, 2021)), the computation of the combinations in (4.4) may become infeasible. In practice, this can be handled with by computing the following which is essentially equivalent to (4.4).

$$\bar{\pi}_\kappa = \prod_{v=0}^{\max(0,|\beta_{\kappa;15}|-1)} \left(1 - \mathbb{I}(|\beta_{\kappa;15}| > 0) \frac{n_{15}}{N_{15} - \hat{n}_{15} - v}\right) \prod_{h=1}^{H} \prod_{b=0}^{\max(0,|\beta_{h\kappa}|-1)} \left(1 - \mathbb{I}(|\beta_{h\kappa}| > 0) \frac{n_h}{N_h - b}\right),$$

where $\mathbb{I}(a)$ is an indicator function that takes value one if a holds, and zero otherwise.

The exclusion probability of a one–person household including an individual aged $90+$ is given by $\binom{N_h - 1}{n_h}/\binom{N_h}{n_h}$, which leads to a selection probability of n_h/N_h, since for such households, we have $|\beta_{\kappa;15}| = 0$, $|\beta_{h\kappa}| = 1$ and $|\beta_{h'\kappa}| = 0$ for any $h' \neq h$ given $\beta_\kappa \in U_h$.

4.3. Initial weights to calibration

A set of initial weights are specified for all individuals in s depending on their selection probabilities that are defined by (4.2), response status, membership of the population by the *reference date*, which is set to the middle of the reference month corresponding to their reference weeks, and whether being directly selected in s_0 if they are in age group $15 - 89$ by the corresponding reference week. Here, the reference week for an individual is defined to be one of the calender weeks for which the data collection is planned to take place for the corresponding reference person in the household of which the individual is a member. The reference month is defined to be the month that includes the reference week of the person of interest. Monthly population register data being up-to-date by the middle of months are used to determine the out-of-scope units.

Let $s_w \subset s$ denote sample individuals that shall be used in the calibration for individual weighting. All individuals in $s \setminus s_w$ take zero initial weights, where $s \setminus s_w$ is defined to be a set of individuals fulfilling any one of the following criteria:

(a) Not a member of the population in the scope of the NLFS by the reference date, or

(b) Being a reference person aged $15{-}89$ years when selected in s_0, but not in the age group $15{-}89$ by the reference week, or

(c) Being a reference person aged $15 - 89$ and have not responded to the survey, or

(d) Being a household member taken into s via a sample reference person and in the age group $15 - 89$ by the reference week, or

(e) Being a household member taken into s via a sample reference person who fulfils any one of (a), (b), or (c).

The criterion (a) refers to out-of-scope units by the reference date regardless of age and whether being a reference person or other household member. The criterion (b) is about that survey data collection is not carried out for sample reference persons once they become above 89 years old. Such individuals and any corresponding household members are removed from the sample data, and thus, they are not used in the weighting process for the age group $90+$ either. The criterion (c) is related to non-response. Because survey data is not collected for other household members in the age group $15 - 89$ for the production of individual level statistics, their initial weights are set to zero in the monthly (or quarterly) individual weighting. By the criterion (e), all other household members are given zero initial weights if their correspond reference persons take zero weights.

For family sampling, all individuals in a family in the age group $15 - 74$ by the date of selection are considered as reference persons since data is collected for all of them as long as they keep being in the reference population and in the age group $15 - 89$ by the reference date. All other individuals not belonging to the family but in the same household as the family are considered as other household members. Hence, the criteria (b)–(e) should be evaluated accordingly for all those coming from the family sampling.

The initial weights for individual weighting are given by

$$d_i = \begin{cases} \pi_i^{-1} & \text{if } i \in s_w, \\ 0 & \text{if } i \in s \setminus s_w, \end{cases} \tag{4.5}$$

where the π_i are defined by (4.2).

4.4. Reweighting with calibration

Calibration is a technique that makes the sample based estimates of totals agree with known population totals (Deville and Särndal, 1992). The final (calibration) weights are obtained by minimising a pre-defined distance function between initial (design weights or adjusted design weights for nonresponse) and the final (calibration) weights over given a set of calibration constraints. The precision increases if the calibration model is correctly specified. Nonresponse bias may also be reduced provided that auxiliary variables are highly associated with both outcome variables and nonresponse (e.g. Little and Vartivarian, 2005; Zhang et al., 2013).

The target population for the European LFSs was defined as “all persons aged 15 years or more and usually residing in the territory of the country at the time of data collection (e.g. EC, 2016, p.5). On the other hand, the target population for the NLFS was defined as all persons aged 15-74 years and usually residing in the territory of Norway, including persons not living in private households. The new regulation (EC, 2019) extends the target population by including all persons usually residing in private households in the territory of the country at the time of data collection. Besides, it is required that the sample distribution shall agree with the population distribution by sex, five–year age groups and region (NUTS II) in the quarterly weighting. Therefore, all individuals in s_w are used in the calibration process regardless of their age unlike the weighing process before 2021.

The weighting for the NLFS from 2021 onwards has been carried out separately for two groups: i) individuals within age group $15-74$ and ii) individuals outside of age group $15-74$, that is within $0-14$ or $75+$. The MMC approach, being used for the NLFS as the main weighting method since 2018, is implemented for the first group while a linear calibration method is used for the second group. Here, ages by the reference weeks are used to assign the sample individuals in s_w to these two groups.

Calibration is applied to monthly sample data as before 2021. Hence, there is no wave-level calibration procedure for the production of the NLFS statistics. Therefore, monthly sample data used in calibration includes data from eight waves. Recall that s_w corresponds to one of the eight waves in a given quarter. Let $s_w^{m_\ell} \subset s_w$ be a wave sample data of month m_ℓ, where $\ell \in \{1,2,3\}$, that shall be used in the weighting for a given quarter. Let $\mathcal{S}_w^{m_\ell}$ be the sample data of month m_ℓ that is obtained as the union of the wave samples of the same month, $s_w^{m_\ell}$, that contain sample units taking non-zero weights. Thus $\mathcal{S}_w^{m_\ell}$ contains sample units from all of the eight waves crossing over a given month m_ℓ. We can write that $\mathcal{S}_w^{m_\ell} = \mathcal{S}_{w;I}^{m_\ell} \cup \mathcal{S}_{w;II}^{m_\ell}$, where $\mathcal{S}_{w;I}^{m_\ell}$ and $\mathcal{S}_{w;II}^{m_\ell}$ correspond to age group $15-74$ and age groups outside of $15-74$ (i.e. $0-14$ and $75+$), respectively.

MMC for age group 15-74

Let $U_{1574} \subset U$ be the population including people aged within $15-74$ by the reference month. Monthly individual calibration weights for month m_ℓ, denoted by $w_i^{m_\ell}$, are obtained by minimising the distance function (e.g. Deville and Särndal, 1992) given by

$$\sum_{i \in \mathcal{S}_{w;I}^{m_\ell}} \frac{(w_i^{m_\ell} - d_i)^2}{d_i q_i}, \quad q_i = 1, \tag{4.6}$$

where the d_i are the initial weights defined by (4.5), with respect to the calibration constraint given by

$$\sum_{i \in \mathcal{S}_{w;I}^{m_\ell}} w_i^{m_\ell} \boldsymbol{\eta}_i = \sum_{i \in U_{1574}} \boldsymbol{\eta}_i, \tag{4.7}$$

where $\boldsymbol{\eta}_i = \{\mu(\mathbf{z}_i;\boldsymbol{\theta})^\top, \mathbf{x}_{i;I}^\top\}^\top$ is a vector of variables that consists of $\mu(\mathbf{z}_i;\boldsymbol{\theta})$, which is a known function of $\mathbf{z}_i$ and $\boldsymbol{\theta}$, and a vector of auxiliary variables, $\mathbf{x}_{i;I}$, available from register data. Here, the relationship between the variable of interest y and a set of covariates $\mathbf{z}$ is described by a model defined by

$$E_\xi(y_i|\mathbf{z}_i) = \mu(\mathbf{z}_i;\boldsymbol{\theta}), \quad V_\xi(y_i|\mathbf{z}_i) = v_i\sigma^2, \tag{4.8}$$

where $\boldsymbol{\theta}$ and σ^2 are unknown model parameters, v_i is a known function of $\mathbf{z}_i$ and $\boldsymbol{\theta}$, and E_ξ and V_ξ are, respectively, the expectation and the variance with respect to the model (4.14). Since $\boldsymbol{\theta}$ is unknown, a sample estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is replaced in $\mu(\mathbf{z}_i;\boldsymbol{\theta})$. Hence, $\hat{\boldsymbol{\eta}}_i = \{\hat{\mu}(\mathbf{z}_i;\boldsymbol{\theta})^\top, \mathbf{x}_{i;I}^\top\}^\top$ is used in the calibration constraint (4.7) in practice. The resulting weights obtained from the minimisation problem is given by

$$w_i^{m_\ell} = d_i \left\{ 1 + \hat{\boldsymbol{\eta}}_i^\top \left(\sum_{i \in \mathcal{S}_{w;I}^{m_\ell}} d_i \hat{\boldsymbol{\eta}}_i \hat{\boldsymbol{\eta}}_i^\top \right)^{-1} \left(\sum_{i \in U_{1574}} \hat{\boldsymbol{\eta}}_i - \sum_{i \in \mathcal{S}_{w;I}^{m_\ell}} d_i \hat{\boldsymbol{\eta}}_i \right) \right\}. \tag{4.9}$$

A constraint on the ratio between initial and calibration weights is also added to the calibration process as such that the weights (4.9) also satisfy with the following condition:

$$\frac{1}{3}\frac{\bar{w}^{m\ell}}{\bar{d}} \leq \frac{w_i^{m_\ell}}{d_i} \leq 3\frac{\bar{w}^{m\ell}}{\bar{d}}, \tag{4.10}$$

where $\bar{w}^{m\ell} = \sum_{i\in\mathcal{S}_{w;I}^{m_\ell}} w_i^{m\ell}/|\mathcal{S}_{w;I}|$ and $\bar{d} = \sum_{i\in\mathcal{S}_{w;I}^{m_\ell}} d_i/|\mathcal{S}_{w;I}|$ are the sample means of $w_i^{m_\ell}$ and d_i, respectively, based on the sample set $\mathcal{S}_{w;I}^{m_\ell}$ the size of which is denoted by $|\mathcal{S}_{w;I}^{m_\ell}|$. The left side of the inequality (4.10) is replaced with $1/3$ in situations where $\bar{w}^{m\ell}/\bar{d} > 3$. If $\bar{w}^{m\ell}/\bar{d} < 1/3$, then the right side of (4.10) is replaced with $infinity$ which leads to no constraint on the upper bound of the weight ratio.

The calibration weights (4.9) ensure the consistency between sample estimates and the population totals of the auxiliary variables $\mathbf{x}_I$ as a result of the calibration constraint (4.7), as such that $\sum_{i\in\mathcal{S}_{w;I}^{m_\ell}} w_i^{m_\ell}\mathbf{x}_{i;I} = \sum_{i\in U_{1574}} \mathbf{x}_{i;I}$. The MMC estimator of the population total of variable y for month m_ℓ, using the weights (4.9), is given by

$$\hat{Y}_{mmc;I}^{m_\ell} = \sum_{i\in\mathcal{S}_{w;I}^{m_\ell}} d_i y_i + \left(\sum_{i\in U_{1574}} \hat{\boldsymbol{\eta}}_i - \sum_{i\in\mathcal{S}_{w;I}^{m_\ell}} d_i\hat{\boldsymbol{\eta}}_i\right)\hat{\boldsymbol{\beta}}_I,$$

where

$$\hat{\boldsymbol{\beta}}_I = \left(\sum_{i\in\mathcal{S}_{w;I}^{m_\ell}} d_i\hat{\boldsymbol{\eta}}_i\hat{\boldsymbol{\eta}}_i^\top\right)^{-1} \sum_{i\in\mathcal{S}_{w;I}^{m_\ell}} d_i\hat{\boldsymbol{\eta}}_i y_i. \tag{4.11}$$

The key variable of labour force surveys is the labour market status, which can be grouped into three categories, as such that *employed*, *unemployed* and *outside of labour force*. A linear relationship between the outcome variable y, which shall be called *ilostat* from now on, and the explanatory variables $\mathbf{z}$ would not be plausible in this case. For the NLFS, a *multinomial logistic regression* model has been used for the target age group 15-74 years to describe the relationship between y and $\mathbf{z}$. There has not been fitted a separate GLM model for each variable of interest in the NLFS due to practical reasons. We may still gain in efficiency and reduce nonresponse bias if other variables of interest is highly associated with the key variable of interest, which is the employment status in three categories (*ilostat*). The $\mathbf{z}$ variables used in the multinomial logistic model are given as follows.

$$\begin{aligned}\mathrm{MnomLog}_{1574}(ilostat) \quad \sim \quad & gender * [age\,(13) + regemp\,(7) + \{age\,(11) * regemp\,(2)\} \\ & + \{regemp\,(2) * marstat\,(2) * age\,(2)\}] + education\,(3) \\ & + county\,(11) + hhsize\,(3) + scheme\,(3) + country\,(3)\end{aligned} \tag{4.12}$$

where '$*$' refers to all–way interactions including main effects (see Table A1.2 for variable descriptions). Let $\hat{p}_e$ $\hat{p}_u$ and $\hat{p}_o$ be the predicted probabilities for individuals to be employed, unemployed, or outside of labour force, respectively, using the multinomial logistic regression model (4.12). We have $\hat{p}_e + \hat{p}_u + \hat{p}_o = 1$ for each individual. These predicted probabilities have been used in the MMC model for $15 - 74$ in addition to a set of auxiliary variables $\mathbf{x}_I$ given as follows.

$$\begin{aligned}\mathrm{MMC}_{1574} \quad := \quad & age\,(12) \times gender + age\,(8) + [age\,(3) \times gender \times \{\hat{p}_e + \hat{p}_u + \hat{p}_o\}] \\ & + [region\,(6) \times \{\hat{p}_e + \hat{p}_u + \hat{p}_o + age\,(3) + gender\}] \\ & + regemp\,(4) \times gender + regemp\,(2) \times country\,(3),\end{aligned} \tag{4.13}$$

where '$\times$' refers to the two–way interactions (see Table A1.2 for variable descriptions).

The only difference between the models given here and those given by Oguz-Alper (2018, pp.42-43) exists in region related variables, that is, $region$ and $county$. This is because of the changes in regional classifications in 2018 and 2020 (see Section 2).

Linear calibration for age groups 0-14 and 75+

Let V be the population including people aged outside of $15 - 74$ by the reference month in the scope of the NLFS. We have $U \cap V = U \setminus U_{1574} = U_{7589} \cup U_{90+}$, where U_{90+} corresponds to those above 89 who lives alone (see Section 4.1). Hence, the target population for the NLFS can be given by $U_{1574} \cup V$. A linear calibration model is used for those outside of $15 - 74$ by the reference week. Monthly individual calibration weights for month m_ℓ are obtained by minimising the distance function given by (4.6) where $\mathcal{S}_{w;I}^{m_\ell}$ is replaced by $\mathcal{S}_{w;II}^{m_\ell}$ under the calibration constraint given by (4.7) where $\mathcal{S}_{w;I}^{m_\ell}$, U_{1574}, and $\boldsymbol{\eta}_i$ are replaced, respectively, with $\mathcal{S}_{w;II}^{m_\ell}$, V, and $\mathbf{x}_{i;II}$. A linear relationship is assumed between the variable of interest y and a set of covariates $\mathbf{x}_{II}$ under linear calibration (e.g. Deville and Särndal, 1992). described by a linear model defined by

$$E_\xi(y_i|\mathbf{x}_{i;II}) = \mathbf{x}_{i;II}^\top \boldsymbol{\beta}_{II}, \quad V_\xi(y_i|\mathbf{x}_{i;II}) = v_i\sigma^2, \tag{4.14}$$

where $\boldsymbol{\beta}_{II}$ and σ^2 are unknown model parameters, and v_i is a known function of $\mathbf{x}_{i;II}$ and $\boldsymbol{\beta}_{II}$. The parameter $\boldsymbol{\beta}_{II}$ is unknown that can be estimated in a similar way to $\boldsymbol{\beta}_I$ by replacing $\mathcal{S}_{w;I}^{m_\ell}$ and $\hat{\boldsymbol{\eta}}_i$, respectively, with $\mathcal{S}_{w;II}^{m_\ell}$ and $\mathbf{x}_{i;II}$ in (4.11). The calibration weights for those in $\mathcal{S}_w^{m_\ell}$ who are aged outside of $15 - 74$ are obtained as follows.

$$w_i^{m_\ell} = d_i \left\{ 1 + \mathbf{x}_{i;II}^\top \left(\sum_{i \in \mathcal{S}_{w;II}^{m_\ell}} d_i \mathbf{x}_{i;II} \mathbf{x}_{i;II}^\top \right)^{-1} \left(\sum_{i \in V} \mathbf{x}_{i;II} - \sum_{i \in \mathcal{S}_{w;II}^{m_\ell}} d_i \mathbf{x}_{i;II} \right) \right\}. \tag{4.15}$$

A lower bound for the ratio between initial and calibration weights is also added to the calibration process for the age group outside of $15 - 74$ that can be defined in the same way as the left side of (4.10), except from that $\bar{w}^{m\ell}$ and $\bar{d}$ are now the sample means of $w_i^{m_\ell}$ and d_i, respectively, based on the sample set $\mathcal{S}_{w;II}^{m_\ell}$. Similarly, lower threshold is replaced with $1/3$ whenever $\bar{w}^{m\ell}/\bar{d} > 3$.

The calibration weights (4.15) ensure the consistency between sample estimates and the population totals of the auxiliary variables $\mathbf{x}_{II}$ as a result of the calibration constraint, as such that $\sum_{i \in \mathcal{S}_{w;II}^{m_\ell}} w_i^{m_\ell} \mathbf{x}_{i;II} = \sum_{i \in V} \mathbf{x}_{i;II}$. The linear calibration estimator, which may also be called *regression estimator*, of the population total of variable y corresponding to those aged outside of $15 - 74$ for month m_ℓ, using the weights (4.15), is given by

$$\hat{Y}_{reg;II}^{m_\ell} = \sum_{i \in \mathcal{S}_{w;II}^{m_\ell}} d_i y_i + \left(\sum_{i \in V} \mathbf{x}_{i;II} - \sum_{i \in \mathcal{S}_{w;II}^{m_\ell}} d_i \mathbf{x}_{i;II} \right) \hat{\boldsymbol{\beta}}_{II}.$$

As mentioned before, the total sample data included sample units from the family and person sampling until the fourth quarter of 2022. Because, at least half of the total sample data consisted of the sample units selected under the family sampling in 2021 and no individuals aged $75 - 89$ were selected under this design, the number of observations for age group $75 - 89$ was not enough to consider this age group as a separate calibration domain for the NLFS 2021 data. Therefore, variable age was not divided into age groups for those above 74 for 2021. Starting from 2022, two age groups have been used: $75 - 89$ and $90+$ due to having enough number of observations that would not cause any problems in the calibration process. Some additional auxiliary variables have also been used starting from 2022 to achieve higher accuracy for employment statistics for age group $75 - 89$.

The model used for 2021 is given as follows.

$$\mathrm{LinCal}_{out1574}^{(1)} \quad := \quad region\,(6) + gender + age014 + age75plus$$

where variable descriptions are given by Table A1.2.

The calibration model which has been used since 2022 is given as follows.

$$\begin{aligned}\mathrm{LinCal}^{(2)}_{out1574} \quad &:= \quad region\,(6) + age014 + gender * (age7589 + age90plus) \\ &\quad + regemp\,(2) \times age7589\end{aligned}$$

where '$\times$' refers to the two–way interactions and '$*$' refers to all–way interactions including main effects (see Table A1.2 for variable descriptions).

4.5. Quarterly and yearly individual weights

Let $s^{m_\ell} \subset s$ be a wave sample for a given month m_ℓ corresponding to a given quarter and $\mathcal{S}^{m_\ell} \supset \mathcal{S}^{m_\ell}_w$, given $s^{m_\ell} \supset s^{m_\ell}_w$, be the sample data of month m_ℓ that is obtained as the union of the wave samples of the same month, that is, s^{m_ℓ}. All individuals in $\mathcal{S}^{m_\ell}_w$, for any given month m_ℓ, take positive monthly weights, (4.9) or (4.15), while all those in $\mathcal{S}^{m_\ell} \setminus \mathcal{S}_w$ take zero weights. Monthly weights, (4.9) or (4.15), are multiplied by corresponding coefficients $0 < p_\ell < 1$ which are proportional to the number of weeks in a given month m_ℓ, which is either 4 or 5 to obtain quarterly individual-level weights. Thus, the quarterly weight of individual i for a given quarter q_k, with $k \in \{1, 2, 3, 4\}$, of a calender year is defined by

$$w_i^{q_k} = \begin{cases} p_\ell w_i^{m_\ell} & \text{if } i \in \mathcal{S}^{m_\ell}_w, \\ 0 & \text{if } i \in \mathcal{S}^{m_\ell} \setminus \mathcal{S}^{m_\ell}_w, \end{cases} \tag{4.16}$$

where m_ℓ refers to a month corresponding to quarter q_k.

Since yearly average of the total estimates are calculated as a simple average of the estimated totals of the corresponding quarters of the year of interest, yearly individual-level weight of an individual i for a given calender year is defined as follows.

$$w_i^{year} = \frac{1}{4}\, w_i^{q_k}, \text{ for } i \in \mathcal{S}^{q_k},$$

where $\mathcal{S}^{q_k} = \cup_{\ell=1}^{3} \mathcal{S}^{m_\ell}$, where q_k is a quarter corresponding to the given calender year.

The statistical software R (R Core Team, 2021) is used for individual weighting. The calibration procedure is implemented by using the R package "ReGenesees"(Zardetto, 2023).

5. Yearly household weighting

Household statistics from the NLFS are produced annually by using only wave 2 samples from which data is collected for all household members taken into sample directly or indirectly (i.e. via reference persons). Number of jobless households, number of children aged $0 - 17$ in jobless households and number of adults aged $18 - 59$ by gender in jobless households are the key labour force household statistics followed up annually (e.g. Eurostat metadata on jobless households. [Online; last accessed 6 November 2023]). In labour market statistics, an individual is classified as being "adult" provided any one of the following conditions is fulfilled:

(C.1) being aged above 24 years, or

(C.2) being aged $15 - 24$ and economically active, that is, being in labour force, or

(C.3) being aged $15 - 24$, economically inactive and not living with at least one of the parents in the same household.

An individual not classified as being adult is counted as being “child”. Jobless households refer to households where no members are in employment, that is, either unemployed or economically inactive.

Household weights may simply be calculated by taking an average of individual weights of household members, which is a common way of calculation of household weights in person and household surveys. However, for the NLFS, only those in wave 2 sample have been used for the production of annual household statistics since 2021, and according to the new European Commission implementing regulation of the EU Labour Force Surveys (EC, 2019), household weights need to satisfy a set of conditions given as follows.

(A) Taking account of selection probabilities,

(B) achieving consistency between the number of households in the population and its annual sample estimate based on the household sub-sample (i.e. wave 2 samples corresponding to the year of interest),

(C) achieving consistency between the hoseuhold sub-sample and population distributions of households by household size groups,

(D) achieving consistency between the hosuehold sub-sample and the population with respect to the population distribution by gender and age groups $0-14$ and $15+$,

(E) achieving consistency between estimated number of people in employment, unemployment and outside the labour force by gender and age groups $25-34$, $35-44$ and $45-54$ from the household sub-sample and the yearly averages of estimates of these totals based on the whole sample for the corresponding year, and

(F) achieving consistency for age groups $15-24$, $55-64$ and $65+$ with respect to the statistics mentioned in condition (E) to the fullest extent possible.

In addition to the conditions (A)–(F), it is desirable to have equal weights for all household members in a given household. Therefore, household weights are calculated by taking into account of these issues as well as additional calibration constraints constructed at individual and household levels.

5.1. Initial weights in household weighting

Initial weights in household weighting are determined based on the design weights, that is, the inverse of (4.2) if individuals or (4.1) if households, as well as a set of criteria (i)–(ii) under which households and all corresponding household members take zero initial weights. Let $\Omega_{s;2}^{q_k}$ be a sample of households in wave 2 for a given quarter q_k, with $k \in \{1,2,3,4\}$, corresponding to a given calender year. Let $\Omega_{s;2;w}^{q_k} \subseteq \Omega_{s;2}^{q_k}$ be the sample set consisting of households whose initial weights being equal to the inverse of their selection probabilities (4.1). A household in $\Omega_{s;2}^{q_k} \setminus \Omega_{s;2;w}^{q_k}$ takes zero initial weight if it satisfies any one of the following criteria.

(i) Having a reference person satisfying at least one of the conditions (a)–(c) , or

(ii) Having at least one member who is in the population by the reference date but had no response to the survey.

Criterion (i) ensures that households whose reference persons take zero monthly and quarterly weights are not considered in the household weighting. Criterion (ii) excludes households not fully respondent, that is, not all household members in-scope responded to the survey. Data is needed from all in-scope household members in order to specify the household characteristics collected from the survey, such as, jobless household. Missing observations may be imputed. However, imputation is not allowed for certain variables including the key variable regarding employment status according to the new EU regulation (EC, 2019). Therefore, partially respondent households are excluded in the household weighting by criterion (ii). Some household members may take zero weights although their corresponding households take positive weights. This would be the case for out-of-scope members, that is, not being in the population by the reference data, whose corresponding households take positive initial weights. All household members taking positive weights in the household weighting shall be called *eligible* household members from now on, unless otherwise stated. Let M_κ be the set consisting of eligible household members of household κ.

5.2. Integrative calibration

The EU criteria (A)–(F) involves both individual and household level conditions that need to be satisfied simultaneously in household weighting. This can be achieved by using *integrative calibration* which, in practice, can be implemented in two ways: 1) calibration is applied to household-level data obtained from aggregating individual-level data, or 2) calibration is applied to individual-level data. In the latter case, calibration variables related to household-level variables need to be created appropriately. For example, a calibration variable regarding the number of households can be created at individual level by dividing its value by the number of eligible household members in the corresponding household. This ensures that the sum of such a calibration variable over eligible household members for a given household becomes equal to the value at household-level. For example, the sum over the eligible household members will be 0 or 1 for a household-level dummy variable. For the transformation of household-level variables to individual-level to work for all types of variables in practice, a dummy variable may be created for each category of categorical variables at household-level. The first implementation method of integrative calibration provides equal weights for all eligible household members belonging to the same household. This can be achieved with the second method, for example, by specifying the stage at which the calibration weights should be constant within sampling units in the "calibrate" function of the *survey* package in R (Lumley, 2020). This is applied for household weighting in the NLFS.

The consistency requirements (E)–(F) involves yearly average estimates of population totals based on the whole sample for the year of interest while for conditions (B)–(C), population totals or estimates of these totals may be used in the calibration. Besides, requirements (B)–(F) are related to variables defined at different levels, individual or household. Therefore, we shall divide the auxiliary variables into three types depending on the level they are defined at and whether population totals or estimated population totals of individual-level variables are used in the calibration constraint.

Let $\mathbf{x}$, γ and $\mathbf{z}$ be the vectors auxiliary variables corresponding to the first, second and third type of variables, respectively. The last two types contain individual-level variables while the first type include household-level variables. Suppose that population totals of $\mathbf{x}$ are γ are used in the calibration constraint whereas estimated population totals of $\mathbf{z}$ is considered in the calibration process. Let γ_κ and $\mathbf{z}_\kappa$ be the vectors of household totals for household κ that are defined by $\gamma_\kappa = \sum_{i \in M_\kappa} \gamma_i$ and $\mathbf{z}_\kappa = \sum_{i \in M_\kappa} \mathbf{z}_i$, respectively. Since monthly population register data are used in the monthly calibration, the averages of population totals from monthly register data are used as population totals in the calibration process. Let $\mathbf{X}_t = \sum_{\kappa \in \Omega_t} \mathbf{x}_\kappa$ and $\mathbf{\Gamma}_t = \sum_{\kappa \in \Omega_t} \gamma_\kappa$ be the vectors of population totals of variables, respectively, $\mathbf{x}$ and γ for month t of a given calender year, where Ω_t is the household population for month t. Then, the yearly average of population totals can be defined by $\mathbf{X} = \sum_{t=1}^{12} \mathbf{X}_t/12$ and $\mathbf{\Gamma} = \sum_{t=1}^{12} \mathbf{\Gamma}_t/12$, respectively.

Yearly household calibration weights w_κ are obtained by minimising a chi-square distance function, like (4.6), under the calibration constraint given by

$$\sum_{k=1}^{4} \sum_{\kappa \in \Omega_{s;2}^{q_k}} w_\kappa (\mathbf{x}_\kappa^\top, \boldsymbol{\gamma}_\kappa^\top, \mathbf{z}_\kappa^\top)^\top = (\mathbf{X}^\top, \mathbf{\Gamma}^\top, \hat{\mathbf{Z}}^\top)^\top, \tag{5.1}$$

where $\hat{\mathbf{Z}}$ is a vector of estimated population totals of $\mathbf{z}$ based on the whole sample over the given year that is defined by

$$\hat{\mathbf{Z}} = \frac{1}{4}\sum_{k=1}^{4}\sum_{i\in\mathcal{S}^{q_k}} w_i^{q_k}\mathbf{z}_i,$$

where $w_i^{q_k}$ is given by (4.16) and $\mathcal{S}^{q_k}$ is defined in Section 4.5. All household members take the same weight as the corresponding household, that is, $w_i = w_\kappa$, for any $i \in M_\kappa$ with $\kappa \in \cup_{k=1}^{4}\Omega_{s;2}^{q_k}$. The individual weights w_i also satisfy with the calibration constraint (5.1) as follows.

$$\sum_{k=1}^{4}\sum_{\kappa\in\Omega_{s;2}^{q_k}}\sum_{i\in M_\kappa} w_i\big(\frac{1}{|M_\kappa|}\mathbf{x}_\kappa^\top, \boldsymbol{\gamma}_i^\top, \mathbf{z}_i^\top\big)^\top = (\mathbf{X}^\top, \boldsymbol{\Gamma}^\top, \hat{\mathbf{Z}}^\top)^\top,$$

by assuming $\mathbf{x}$ is originally defined to be a household-level variable while both γ and $\mathbf{z}$ are originally defined as individual-level variables. Lower and upper bounds for the ratio between initial and calibration weights at individual level are set in the calibration process in a similar way as the one described in Section 4.4.

5.3. Numerical results

Linear calibration is used to obtain yearly household weights w_κ. Four models, Model A, Model B, Model C, Model D, with increasing complexities were compared in terms of point and standard error estimates for the key household statistics by using data in 2021. The most parsimonious model, called Model A, is constructed in a way that the EU requirements (A)–(F) are all satisfied. Model B includes variables regarding the key household statistics in addition to the variables in Model A. Models C and D contains all variables in Model B and interactions between some of the existing variables in Model A. The models are defined as follows where the variables used are described in Table A1.3.

$$\begin{aligned}
\text{Model A} \quad &:= \quad 1/hhsize \times hhsizegr\,(5) + gender + age014 \\
&\quad + gender \times age1574 \times (emply + unemply + outLF) \\
&\quad + (age1524 + age2534 + age3544 + age4554) \times (emply + unemply + outLF) \\
&\quad + (age5564 + age6574) \times (emply + notemply),
\end{aligned}$$

$$\begin{aligned}
\text{Model B} \quad &:= \quad \text{Model A variables} \\
&\quad + 1/hhsize \times\ injoblesshh + child017 + child017 \times injoblesshh \\
&\quad + gender \times (adult1859 + adult1859 \times injoblesshh),
\end{aligned}$$

$$\begin{aligned}
\text{Model C} \quad &:= \quad \text{Model B variables} \\
&\quad + gender \times (age1524 + age2534 + age3544 + age4554) \times (emply + notemply) \\
&\quad + gender \times (age5564 + age6574) \times (emply + notemply),
\end{aligned}$$

$$\begin{aligned}
\text{Model D} \quad &:= \quad \text{Model B variables} \\
&\quad + gender \times (age1524 + age2534 + age3544) \times (emply + unemply + outLF) \\
&\quad + gender \times (age4554 + age5564 + age6574) \times (emply + notemply),
\end{aligned}$$

where '$\times$' refers to the two–way interactions. Household-level variables, which are considered via $\mathbf{x}$ in the calibration constraint (5.1), are given by $1/hhsize \times hhsizegr\,(5)$ and $1/hhsize \times injoblesshh$ in Models A–D. Individual-level variables that are denoted by γ in the calibration constraint (5.1) are those that are added by Model B, except from $1/hhsize \times injoblesshh$. The rest of the variables in Models A–D are individual-level variables that are involved in (5.1) via $\mathbf{z}$.

In the numerical work, variables regarding being child or adult in the Models given above were replaced with corresponding proxy variables which were specified by ignoring condition (C.3). This is because of the fact that variables that are required to check this condition were not available in the sample and population data sets when the numerical work was carried out. Ignoring condition (C.3) resulted some of the adults aged $15-24$ being misclassified as child. However, this would not change the general conclusions that could be drawn from the numerical results given by Tables 5.1–5.2.

In Table 5.2, variances are underestimated due to the fact that the estimated totals $\hat{\mathbf{Z}}$ were used in the calibration was ignored in variance estimation. However, assuming the rate of underestimation does not depend on the models given above, our conclusions regarding efficiency would remain same. We also assume that additional auxiliary variables given in Model B are highly associated with the variables of interest corresponding to the parameters in Tables 5.1–5.2, and thus, Models B–D would provide closer values to the true population values. Given these assumptions, the main observations based on the results presented in Tables 5.1–5.2 are given as follows.

- Model A underestimates almost all the parameters compared to Models B–D. Specifically, the differences between the point estimates are remarkable for the number (or proportions) of men and women aged $18-59$ in jobless households.

- Variance estimates are reduced by at least around 50% by Models B–D compared to Model A, except for the number of jobless households, where the variance reduction is around 20%. Hence, more precise estimates are obtained with Models B–D.

- Differences between point and standard error estimates by Models B–D are negligible.

Table 5.1 Point estimates for the key household statistics obtained from data in 2021

Parameter of interest	Model A	Model B	Model C	Model D
Number of jobless households (JlessHHs)	899 231	903 768	902 788	902 617
Number of children* aged $0-17$ in JlessHHs	80 453	83 244	83 502	83 467
Number of adults* aged $18-59$ in JlessHHs	281 533	284 973	284 832	284 625
Number of men* aged $18-59$ in JlessHHs	122 706	143 414	143 340	142 340
Number of women* aged $18-59$ in JlessHHs	158 827	141 559	141 491	142 285
Proportion of children* aged $0-17$ in JlessHHs among all children* aged $0-17$	7.85	7.99	8.01	8.01
Proportion of adults* aged $18-59$ in JlessHHs among all adults* aged $18-59$	9.81	9.97	9.96	9.96
Proportion of men* aged $18-59$ in JlessHHs among all men* aged $18-59$	8.50	9.81	9.81	9.75
Proportion of women* aged $18-59$ in JlessHHs among all women* aged $18-59$	11.14	10.14	10.13	10.18

* Child and adult (man, woman) were specified in an adhoc way by ignoring condition (C.3). This is due to the lack of variables, which are required to check this condition, in the survey data when this numerical work was carried out

Table 5.2 Standard error estimates for the key household statistics obtained from data in 2021

Parameter of interest	Model A	Model B	Model C	Model D
Number of jobless households (JlessHHs)	16 360	14 633	14 580	14 586
Number of children* aged $0-17$ in JlessHHs	7 338	4 538	4 518	4 574
Number of adults* aged $18-59$ in JlessHHs	8 073	5 761	5 780	5 787
Number of men* aged $18-59$ in JlessHHs	5 961	3 967	3 970	3 914
Number of women* aged $18-59$ in JlessHHs	6 729	3 899	3 890	3 860
Proportion of children* aged $0-17$ in JlessHHs among all children* aged $0-17$	0.710	0.431	0.429	0.434
Proportion of adults* aged $18-59$ in JlessHHs among all adults* aged $18-59$	0.277	0.203	0.204	0.204
Proportion of men* aged $18-59$ in JlessHHs among all men* aged $18-59$	0.394	0.270	0.272	0.268
Proportion of women* aged $18-59$ in JlessHHs among all women* aged $18-59$	0.452	0.278	0.276	0.275

* Child and adult (man, woman) were specified in an adhoc way by ignoring condition (C.3). This is due to the lack of variables, which are required to check this condition, in the survey data when this numerical work was carried out.

5.4. Final model in household weighting

The numerical results given by Tables 5.1–5.2 clearly show that Models B–D are much better than Model A in terms of accuracy for the key household-level statistics considered in the numerical work. Since Models C and D do not substantially provide better results than Model B, and also, they are more complex than Model B which may lead to less robust estimates over time especially when the calibration cells become small due to missing observations, it was decided to use Model B in the calculation of yearly household weights for the NLFS.

Some changes were applied in Model B starting from 2022 due to the same reasons for changes made in the calibration model for individual weighting for those outside of $15-74$ (see Section 4.4). The age group $75+$ was further divided into two groups as $75-89$ and $90+$. Besides, interactions with gender as well as employment status were added to Model B. Hence, the final model which has been used since 2022 is given as follows where changes are highlighted with bold font.

$$\begin{aligned}
\text{Model B}^{(final)} \quad := \quad & 1/hhsize \times hhsizegr\,(5) + gender + age014 \\
& + \mathbf{gender} \times (\mathbf{age7589} + \mathbf{age90plus}) \\
& + gender \times age1574 \times (emply + unemply + outLF) \\
& + (age1524 + age2534 + age3544 + age4554) \times (emply + unemply + outLF) \\
& + (age5564 + age6574) \times (emply + notemply) + \mathbf{age7589} \times \mathbf{emply} \\
& + 1/hhsize \times injoblesshh + child017 + child017 \times injoblesshh \\
& + gender \times (adult1859 + adult1859 \times injoblesshh).
\end{aligned}$$

Weighting program was written in statistical software R (R Core Team, 2021). The calibration procedure for household weighting is implemented by using the R package “survey”(Lumley, 2020).

6. Conclusion

The revision made in the weighting methodology of the NLFS in 2018 was remarkable in the sense that a model-calibration approach was introduced to be implemented for the production of labour market statistics for the first time in the history of the NLFS. To the best of our knowledge, there is no any other statistical office that has used model-calibration approach for the production of its LFS statistics. Although the theory of model-calibration approach is first developed by Wu and Sitter (2001), we see an earlier application of logistic generalised difference estimator on the Finnish labour force survey (Lehtonen and Veijanen, 1998) which is similar to model-calibration in the sense that the relationship between the outcome variable y and model covariates is determined by a logistic model rather than being limited to a linear model. This revision was made as a result of an empirical investigation (Oguz-Alper, 2018) following the availability of many and good quality of register based variables that could be used as auxiliary variables for the NLFS.

The revision made in 2021 described in this document is the second revision in the weighting methodology of the NLFS in the last decade. The main weighting methodology was not changed in these sense that the model-calibration approach with the same auxiliary variables (see Section 4.4) is still used for the age group $15-74$. However, significant changes occurred in the calculation of selection probabilities and initial weights due to the new sampling design that was gradually phased in quarter by quarter, changes in the data structure and the new weighting requirements by the new EU legislation (EC, 2019). A linear calibration method was introduced for those within age groups $0-14$ and $75+$. An integrative calibration approach has been used for the calculation of yearly household weights for the first time in the history of the NLFS. Previously, the naïve method was used for the calculation of family weights by taking the average of individual weights within families. Besides, some auxiliary variables, being highly associated with the key annual household statistics, were taken into account in the construction of the calibration model for household weighting (see Section 5.3) in order to increase the accuracy of these statistics.

A future revision in the current weighting methodology might be considered by transforming the current calibration procedure to wave-specific calibration which will allow to estimate wave-specific sampling errors that could be used in the production of seasonally adjusted monthly series (e.g. Hamre et al., 2022). Besides, there is a project ongoing to implement a mixed-mode data collection method for the NLFS. In this regard, wave-specific calibration might be especially useful when dealing with possible mode effects that might be caused when the data collection method varies among waves. However, it is currently infeasible to apply a wave-specific calibration with the calibration model (4.13) used for the age group $15 - 74$ due to empty or very small calibration cells for monthly wave-based data. To deal with this problem, one may implement a wave-specific calibration by using quarterly data instead of monthly data, so that the number of sample observations used in the calibration will increase although this may not guarantee to have sufficient number of observations for all waves, especially for later waves such as waves $5 - 8$. In case this is a feasible solution, a separate calibration procedure with a simplified model, possibly with less number of variables, needs to be applied to monthly data in order to produce monthly weights. With such a procedure, quarterly weights will not be a simple function (4.16) of monthly weights, unlike the current methodology, which allows a linear aggregation of monthly total estimates to obtain quarterly total estimates. Thus, if consistency between monthly and quarterly estimates, not seasonally adjusted, is desired for specific statistics over specific domains, this needs to be taken into account in the construction of calibration model for the production of quarterly weights. As a conclusion, an empirical study may be considered as a future investigation to evaluate the advantages and disadvantages of transition to a wave-specific calibration as well as the feasibility of such a procedure given a set of requirements.

References

Andersson, P. and Särndal, C. E. (2016). Calibration for nonresponse treatment: In one or two steps? *Statistical Journal of the IAOS*, 32:375–381.

Bethel, J. (1989). Sample allocation in multivariate surveys. Survey Methodology.

Cassel, C. M., Särndal, C. E., and Wretman, J. H. (1976). Some results on generalized difference estimation and generalized regression estimation for finite populations. *Biometrika*, 63:615–620.

Deville, J. C. and Särndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87(418):376–382.

EC (2016). Tf3 – consolidated implementing act in the labour market domain: structure and content. Commission Regulation (EC), Doc.: Eurostat/F3/LAMAS/28/16. [Online; last accessed 18 March 2021]. https://circabc.europa.eu/sd/a/71c42ee4-702d-4b7e-a8ab-d925cf399dad/Doc%2028%20Item%204.8%20TF3%20-%20Consolidated%20implementing%20act%20in%20the%20labour%20market%20domain%20-%20structure%20and%20content.pdf.

EC (2019). Commision implementing regulation (eu) no 2019/2240 of 16 december 2019. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R2240&from=EN. [Online; last accessed 18 March 2021].

Hamre, J. I. and Heldal, J. (2013). Improved calculation and dissemination of coefficients of variation in the norwegian lfs. Documents 46/2013, Statistics Norway. https://www.ssb.no/en/arbeid-og-lonn/artikler-og-publikasjoner/_attachment/148090?_ts=142476a8ad0. [Online; last accessed 21 September 2017].

Hamre, J. I., Hungnes, H., Jansen, X., Pham, D., Sandvik, O., and Skjerpen, T. (2022). Break estimation in the norwegian lfs due to the 2021 redesign. Documents 2022/3, Statistics Norway. https://www.ssb.no/arbeid-og-lonn/sysselsetting/artikler/break-estimation-in-the-norwegian-lfs-due-to-the-2021-redesign.documentation-of-the-methods-and-some-results/_/attachment/inline/2d706750-a335-4505-b313-f31f8db3c14e:371261092c490557f38ba811a10d4a601125b3df/NOT2022-03.pdf. [Online; last accessed 1 October 2023].

Haziza, D. and Lesage, É. (2016). A discussion of weighting procedures for unit nonresponse. *Journal of Official Statistics*, 32(1):129–145.

Heldal, J. (2000). Kalibrering av aku. Notater 2000/07, Statistisk sentralbyrå. https://www.ssb.no/a/publikasjoner/pdf/notat_200007/notat_200007.pdf. [Online; last accessed 21 September 2017].

Horgen, E. H., Lien, H. H., Sandvik, O., and Sundt, C. S. (2023). Dokumentasjon av arbeidskraftundersøkelsen etter omleggingen i 2021. Notater 2023/37, Statistisk sentralbyrå. https://www.ssb.no/arbeid-og-lonn/sysselsetting/artikler/dokumentasjon-av-arbeidskraftundersokelsen-etter-omleggingen-i-2021. [Online; last accessed 04 November 2023].

Jentoft, S. (2022). The norwegian labour force survey sampling design. Documents 2022/34, Statistics Norway. https://www.ssb.no/arbeid-og-lonn/sysselsetting/artikler/the-norwegian-labour-force-survey-sampling-design/_/attachment/inline/59132a2d-6774-4b2e-9116-a7fb50432f45:cf0acc6440df9b3ded84c2aeea020b47f8b30eda/NOT2022-34.pdf. [Online; last accessed 17 April 2023].

Lehtonen, R. and Veijanen, A. (1998). Logistic generalized regression estimators. *Survey Methodology*, 24(1):51–55.

Little, R. J. and Vartivarian, S. (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodolology*, 31(2):161–168.

Lumley, T. (2020). survey: Analysis of complex survey samples. R package version 4.0.

Lundström, S. and Särndal, C. E. (1999). Calibration as a standard method for treatment of nonresponse. *Journal of Official Statistics*, 15(2):305–327.

Montanari, G. and Ranalli, G. (2009). Multiple and ridge model calibration for sample surveys. Workshop on Calibration and Estimation in Surveys , Statistics Canada. http://www.stat.unipg.it/~giovanna/papers/ranalliWCES.pdf. [Online; last accessed 16 November 2017].

Nguyen, N. D. and Zhang, L. (2020). An appraisal of common reweighting methods for nonresponse in household surveys based on the norwegian labour force survey and the statistics on income and living conditions survey. *Journal of Official Statistics*, 36:151–172.

Oguz-Alper, M. (2018). New estimation methodology for the norwegian labour force survey. Documents 2018/16, Statistics Norway. https://www.ssb.no/en/arbeid-og-lonn/artikler-og-publikasjoner/_attachment/346996?_ts=162d85820c8. [Online; last accessed 16 March 2021].

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. http://www.R-project.org, Vienna, Austria.

Vedø, A. and Rafat, D. (2003). Sammenligning av utvalgsplaner i aku. Notater 2003/56, Statistisk sentralbyrå. http://www.ssb.no/emner/08/02/20/notat_200460/notat_200460.pdf. [Online; last accessed 20 September 2017].

Villund, O. (2009). Measuring working hours in the norwegian labour force survey. Reports 2009/3, Statistisk sentralbyrå. https://www.ssb.no/a/english/publikasjoner/pdf/rapp_200903_en/rapp_200903_en.pdf. [Online; last accessed 06 November 2023].

Wu, C. and Sitter, R. (2001). A model-calibration approach to using complete auxiliary information from survey data. *Journal of the American Statistical Association*, 96(453):185–193.

Zardetto, D. (2023). Regenesees: R evolved generalized software for sampling estimates and errors in surveys. R package version 2.3.

Zhang, L. C. (1998). Dokumentasjonsrapport: Den nye estimeringsmetoden for arbeidskraftundersøkelsen (aku). Notater 1998/01, Statistisk sentralbyrå. https://www.ssb.no/a/histstat/not/not_9801.pdf. [Online; last accessed 21 September 2017].

Zhang, L. C., Thomsen, I., and Kleven, O. (2013). On the use of auxiliary and paradata for dealing with non-sampling errors in household surveys. *International Statistical Review*, 81:270–288.

Appendix: Tables and Figures

Table A1.1 The ratios between the within-stratum sampling fractions and the overall sampling fraction for family-sampling. Used from the second quarter of 2020 to the fourth quarter of 2020

Region code	Region name	County code	County name	Ratio
01	Oslo and Viken	03	Oslo	0.871
		30	Viken	0.934
02	Innlandet	34	Innlandet	1.187
03	Agder and Sør–Østlandet	38	Vestfold and Telemark	1.000
		42	Agder	1.187
04	Vestlandet	11	Rogaland	1.000
		15	Møre and Romsdal	1.000
		46	Vestland	0.963
05	Trøndelag	50	Trøndelag	1.000
06	Nord-Norge	18	Nordland	1.000
		54	Troms and Finnmark	1.265

Source: An internal unpublished report (in Norwegian) titled "Justering av utvalgsplan for AKU i 2020" that was written by Jørn-Ivar Hamre, Senior Adviser, Division of Methodology, Statistics Norway.

Figure A1.1 An example of the panel structure of the NLFS from 2020 to 2022

2020				2021				2022			
Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
W8											
W7	W8										
W6	W7	W8									
W5	W6	W7	W8								
W4	W5	W6	W7	W8							
W3	W4	W5	W6	W7	W8						
W2	W3	W4	W5	W6	W7	W8					
W1	W2	W3	W4	W5	W6	W7	W8				
	W1	W2	W3	W4	W5	W6	W7	W8			
		W1	W2	W3	W4	W5	W6	W7	W8		
			W1	W2 + HH	W3	W4	W5	W6	W7	W8	
				NEW1	NEW2 + HH	NEW3	NEW4	NEW5	NEW6	NEW7	NEW8
					NEW1	NEW2 + HH	NEW3	NEW4	NEW5	NEW6	NEW7
						NEW1	NEW2 + HH	NEW3	NEW4	NEW5	NEW6
							NEW1	NEW2 + HH	NEW3	NEW4	NEW5
								NEW1	NEW2 + HH	NEW3	NEW4
									NEW1	NEW2 + HH	NEW3
										NEW1	NEW2 + HH
											NEW1

Source: Jørn-Ivar Hamre, Senior Advisor, Division of Methodology, Statistics Norway.
Q: quarter; W: wave; HH: collection of household data from all household members; Yellow colour: family sampling and regional classification used between 2018 and the first quarter of 2020; Blue colour: family sampling and regional classification used between the second and the fourth quarters of 2020; Green colour (or NEW): person sampling and regional classification used since 2021.

Table A1.2 Descriptions of the variables used in individual weighting

Variable	Label	Categories
Gender	gender	male, female
Age	age (2)	15-59, 60-74
	age (3)	15-24, 25-54, 55-74
	age (8)	15-17, 18-19, 20-24, 25-39, 40-54, 55-61, 62-66, 67-74
	age (11)	five–year age groups from 15 to 64, and 65-74
	age (12)	five–year age groups from 15 to 74
	age (13)	15-17, 18-19 and five–year age groups from 20 to 74
	age014	1: if $age \leq 14$, and 0: otherwise
	age75plus	1: if $age \geq 75$, and 0: otherwise
	age7589	1: if $75 \leq age \leq 89$, and 0: otherwise
	age90plus	1: if $age \geq 90$, and 0: otherwise
Register based employment status	regemp (2)	employed, not employed
	regemp (4)	full–time employed, part–time employed, self-employed, others
	regemp (7)	full–time employed, part–time employed, self-employed, unemployed for 90 days or less, unemployed for more than 90 days, permanently disabled, outside of labour force
NUTS II	region (6)	Oslo and Viken, Innlandet, Agder and Sør-Østlandet, Vestlandet, Trøndelag, Nord-Norge
NUTS III	county (11)	Oslo, Rogaland, Møre and Romsdal, Nordland, Viken, Innlandet, Vestfold and Telemark, Agder, Vestland, Trøndelag, Troms and Finnmark
Education	education (3)	1: upper secondary school: 2: higher education; 3: primary school or others
Marital status	marstat (2)	2: married or registered partner; 1: others
Household size	hhsize (3)	1: one person; 2: two persons ; 3: three or more persons
Country of origin[1]	country (3)	0: not immigrants; 1: immigrants coming from the EEA (The European Economic Area), USA, New Zealand, Canada and Australia; 2: immigrants coming from other countries or stateless
Scheme[2]	scheme (3)	2: unemployed, ordinary scheme participant, salary subsidies, skill–training scheme, temporary employment scheme and other ordinary schemes; 3: occupationally handicapped or reduced working capacity in scheme and not in scheme; 1: others

[1] Derived based on a register variable providing country background back to three generations and a register variable indicating immigration status.
[2] More information about scheme can be found here: About scheme participants. [Online; last accessed 1 October 2023].

Table A1.3 Descriptions of the variables used in household weighting

Variable	Label	Description/Categories
Gender	gender	male, female
Age	age014	1: if $age \leq 14$, and 0: otherwise
	age1574	1: if $15 \leq age \leq 74$, and 0: otherwise
	age1524	1: if $15 \leq age \leq 24$, and 0: otherwise
	age2534	1: if $25 \leq age \leq 34$, and 0: otherwise
	age3544	1: if $35 \leq age \leq 44$, and 0: otherwise
	age4554	1: if $45 \leq age \leq 54$, and 0: otherwise
	age5564	1: if $55 \leq age \leq 64$, and 0: otherwise
	age6574	1: if $65 \leq age \leq 74$, and 0: otherwise
	age7589	1: if $75 \leq age \leq 89$, and 0: otherwise
	age75plus	1: if $age \geq 75$, and 0: otherwise
	age90plus	1: if $age \geq 90$, and 0: otherwise
Register based employment status	emply	1: if employed, and 0: otherwise
	unemply	1: if unemployed, and 0: otherwise
	outLF	1: if out of labour force, and 0: otherwise
	notemply	1: if not employed, and 0: otherwise
Maturity[1] & age	child017	1: if child aged $0-17$, and 0: otherwise
	adult1859	1: if adult aged $18-59$, and 0: otherwise
Household size	hhsize	Number of eligible household members
	hhsizegr (5)	Five household size groups: 1, 2, 3 , 4 or $5+$
Jobless household	injoblesshh	1: if member of a jobless household, and 0: otherwise

[1] Maturity is defined according to the conditions given in Section 5. For the numerical work presented in Section 5.3, condition (C.3) was ignored in the determination of maturity due to the lack of variables in survey and register data at the time when the numerical work was carried out.