Notater
Documents
2018/16

Statistics Norway
Statistisk sentralbyrå

Melike Oguz-Alper

New estimation methodology for the Norwegian Labour Force Survey

Documents 2018/16

Melike Oguz–Alper

New estimation methodology for the Norwegian Labour Force Survey

Statistisk sentralbyrå • Statistics Norway
Oslo–Kongsvinger

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

© Statistics Norway
When using material from this publication, Statistics Norway shall be quoted as the source.

Published 26 April 2018

ISBN 978-82-537-9729-8 (electronic)

Symbols in tables	**Symbol**
Category not applicable	.
Data not available	..
Data not yet available	...
Not for publication	:
Nil	-
Less than 0.5 of unit employed	0
Less than 0.05 of unit employed	0.0
Provisional or preliminary figure	*
Break in the homogeneity of a vertical series	—
Break in the homogeneity of a horizontal series	|
Decimal punctuation mark	.

Preface

The aim of this monograph is to document the revision of the previous estimation methodology used for the Norwegian Labour Force Survey (LFS) and the review of the estimators commonly used in Household and Person Surveys in order to find a more efficient estimator which has also a lower non–response bias. Such a revision was brought up to the agenda as a better administrative register, namely *A-ordningen*, in terms of quality and the variety of auxiliary variables for the labour market was 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).

The project was carried out with the collaboration of Division for Labour Market and Wage Statistics and Division for Methods. It was lead by Jørn–Ivar Hamre from Division for Labour Market and Wage Statistics, who made the data available and provided full support regarding variables and the labour market statistics during the project. The monograph has been written by Melike Oguz–Alper from Division for Methods, who implemented the methods presented in this monograph to the Norwegian LFS data.

The author wishes to thank to Anders Holmberg and Magnar Lillegård from Division for Methods and Jørn–Ivar Hamre for their useful comments and suggestions that have significantly improved the first version of this monograph. The author is also grateful to Prof Li–Chun Zhang for his enlightening comments and the clarifications he made on the technical aspects of the work done.

Statistics Norway, 11 April 2018

Jørn Leonhardsen

Abstract

Labour Force Survey (LFS) is an important source of the labour market statistics that provides information about the participation of people aged 15 and over in to the labour market and people outside of the labour market. It is a rotating panel sample survey that is carried out in accordance with the European Union (EU) Council Regulation. Statistics produced are subject to both sampling and non–response errors. Sampling errors are monitored through standard errors, which are provided alongside with the point estimates for the key variables. In that respect, finding an efficient estimator is one of the main goals for the LFS. This requires data sources that includes good auxiliary variables. Thus we aim to find an estimation methodology which better utilises the auxiliary information in the light of a new available data source, namely A–ordningen. In this regard, we compare the regular generalised regression estimator (GREG) and the (multiple) model–calibration estimator, which has been shown to be optimal among a class of calibration estimators, in terms of efficiency by using the Norwegian LFS data. Standard errors are estimated by using the Jackknife linearisation (JL) variance estimator. Overall, for the data used, the (multiple) model–calibration estimators have been more efficient than than the GREG estimators. Thus the former has been chosen to be used in the production of the Norwegian labour force statistics.

Non-response may lead to biased estimates if it is not properly handled in the estimation under a non–uniform response mechanism (i.e. not missing completely at random (MCAR)). We discuss two types of weighting procedures. One of them involves a separate step for non–response adjustment, and the other one handles with non–response as a part of calibration. We have observed, for the data used, that the two–step estimators have provided higher standard errors without reducing non–response bias more. Thus it has been decided to use a one–step (multiple) model–calibration estimator in the production of the Norwegian labour force statistics.

Equal– and unequal–weighted averages of monthly estimates have also been compared in order to investigate the effects of each on quarterly estimates. The former was used by the previous estimation methodology (see Section 4). The latter is proposed being used in the new estimation methodology (see Section 12.4).

The new estimation methodology has been examined with regards to whether or not it satisfies the EU precision requirements. The requirements are met for the data used.

A stratified one–stage cluster sampling is used to select sample units for the Norwegian LFS. We observe that the cluster effect may be ignored in the variance estimation if good auxiliary variables are used in the estimation. This facilitates the computation of variance estimates, especially for changes in statistics over time, for which the variance estimation may be more tedious in rotating panel surveys. The cluster effect is also ignored in the variance estimation procedure previously used.

Contents

Preface . . . 3

Abstract . . . 4

1. Introduction . . . 6
1.1. Utilizing new available data . . . 6
1.2. Estimation methods considered . . . 6
1.3. Methods adjusting for non–response bias . . . 7
1.4. New weighting of monthly estimates . . . 7
1.5. Precision requirements and variance estimation . . . 8
1.6. Sections . . . 8

2. Notation . . . 8

3. The sampling design of the Norwegian LFS . . . 9

4. The previous estimation methodology for the Norwegian LFS . . . 10

5. Parameter of interest . . . 11

6. Point estimators . . . 11
6.1. Reference estimator . . . 11
6.2. One-step GREG estimator . . . 11
6.3. One–step (multiple) model–calibration estimator . . . 12
6.4. Two-step GREG estimator . . . 14
6.5. Two–step (multiple) model–calibration estimator . . . 15

7. Variance estimation . . . 15
7.1. The Jackknife linearisation (JL) variance estimator . . . 16

8. Estimation of a ratio . . . 19

9. Estimation of quarterly totals . . . 20

10. Domain estimation . . . 20

11. Application: Norwegian LFS . . . 21
11.1. One–step estimation . . . 22
11.2. Two–step estimation . . . 23

12. Numerical results . . . 25
12.1. The choice of a new estimation method for Norwegian LFS . . . 25
12.2. Unemployment rate . . . 26
12.3. Estimation over domains . . . 31
12.4. Equal– versus weekly–weighted average for quarterly estimates . . . 31
12.5. EU precision requirements . . . 40
12.6. Effect of clustering . . . 40

13. Conclusion . . . 42

References . . . 44

Appendix A: Derivation of total differentials . . . 47

Appendix B: Tables . . . 48

1. Introduction

1.1. Utilizing new available data

The Norwegian Labour Force Survey (LFS) is an important source of the labour market statistics, which has been conducted by Statistics Norway since 1972 (SSB, 2001, p.19). It is a rotating panel sample survey that is carried out in accordance with the Council Regulation (EC) (1998). The survey provides information about the participation of people aged 15 and over in to the labour market and people outside of the labour market. Statistics produced are subject to both sampling and non–sampling errors. In this working paper, we will focus on sampling and non–response errors. For the latter, we will only consider unit non–response, which today is around 20% in the Norwegian LFS. Both sampling design and non–response errors are taken into account with the previous estimation methodology (see Section 4). Sampling errors are monitored through standard error estimates, which are provided alongside with the point estimates for the key variables. The aim is to document an improved revision of this methodology through comparisons pointing to that the new methodology introduced provides more efficient point estimates as well as lower non–response bias. The revision has been made possible by utilizing better auxiliary information 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) (see `https://www.ssb.no/omssb/om-oss/nyheter-om-ssb/a-ordningen-en-datakilde-for-tre-etater`. [Online; last accessed 07 February 2018]).

Auxiliary information, which may be obtained from administrative registers, censuses or other types of reliable data sources, is commonly used in the LFSs to increase efficiency and ensure consistency with known population quantities. Calibration estimators (Deville and Särndal, 1992) are often used in official statistics for these purposes. Gain in efficiency with a calibration estimator over those that do not involve auxiliary information, for example, Horvitz and Thompson (1952) or Hájek (1964) estimators, is obtained provided that auxiliary information is highly correlated with the outcome variables of interest. Age, gender, region and register based employment status (employed or not) (see Section 4) have been in use for many years for the Norwegian LFS. Although register based employment explains the employment obtained from the LFS quite well, this may not be valid for unemployment and outside of labour force statistics. Therefore, we aim to find better auxiliary variables for the latter and/or an estimation methodology that may incorporate available register variables in to the estimation procedure more efficiently so that we gain in accuracy for all the key variables of interest.

1.2. Estimation methods considered

The generalised regression (GREG) estimator (e.g. Cassel et al., 1976) is a special type of estimator in the class of calibration estimators. A linear relationship between the variable of interest and the auxiliary variable is implicitly assumed with the GREG estimator. Values at the auxiliary variables should be known for all sample units. However, unit–level information is not required for the units outside of the sample, where it is sufficient to know population totals. This is an advantage if unit–level information is not available for all units in the population. Moreover, one set of weights is obtained, which is very practical in the case of huge number of variables involved in the estimation processes. However, when the linear relationship assumption does not hold, the GREG may provide less efficient estimates. This may be the case for the LFS as the response variables are categorical variables. The model–calibration estimator, which can handle such more general cases, is proposed by Wu and Sitter (2001). It is not restricted to a linear working model, unlike the GREG.

The model–calibration estimator is an optimal calibration estimator among a class of calibration estimators in the sense that it minimises the model expectation of the asymptotic design–based variance under the true model and any regular sampling design (Wu, 2003, p.940). Complete auxiliary information is, however, required for the use of it. Because of some possible concerns related to consistency with population totals and obtaining one set of calibration weights (see Section 6.3), a modified version of the model–calibration estimator, namely multiple model–calibration estimator, is proposed by Montanari and Ranalli (2009). They showed that this estimator is also an optimal calibration estimator. In this working paper, we compare the GREG and the multiple model–calibration estimators in terms of efficiency. We are not aware of any national statistical office that uses the model–calibration estimator in the production of labour market statistics. Lehtonen and Veijanen (1998) provides numerical results based on the Finnish LFS by using a logistic generalised difference estimator. However, this is not the same as the model–calibration estimator under a working logistic model. Thus this work will provide important aspects towards the use of this type of estimator in official statistics.

1.3. Methods adjusting for non–response bias

Non-response may lead to biased estimates if it is not properly handled in the estimation under a non–uniform response mechanism (i.e. not missing completely at random (MCAR)). Re–weighting is, in practice, often applied in order to reduce non–response bias. Re–weighting increases design weights of the respondent units in order to compensate those who have not responded. Two re–weighting procedures are common in practice. The first one involves the multiplication of design weights by the inverse response propensities, which can be estimated by a logistic model (e.g. Little, 1986) or a uniform probability model within classes (i.e. re–weighting within response homogeneous classes (RHC)) (e.g. Särndal et al., 1992, p.578). The second one involves the use of calibration (e.g. Lundström and Särndal, 1999). The former may be referred as a two–step weighting approach while the latter may be called a one–step weighting approach (e.g. Haziza and Lesage, 2016; Andersson and Särndal, 2016). If the model is correct, non–response bias is reduced. The estimation methodology may involve one–step or two–step weighting procedure depending on whether or not the non–response adjustment is carried out at a separate step. With a two–step estimation approach, adjusted weights obtained at the first step are used as initial weights in the calibration step. In order to achieve both efficiency and bias reduction, auxiliary information should be highly associated with both the variables of interest and non–response (e.g. Little and Vartivarian, 2005; Nguyen and Zhang, 2016). Besides, if there exists good auxiliary variables that explain the variables of interest, a one–step estimation method may both increase efficiency and reduce non–response bias (Nguyen and Zhang, 2016). Therefore, we aim to investigate if there is any difference between one–step and two–step approaches in terms of point and standard error estimates for the Norwegian LFS data (see Section 12.1).

1.4. New weighting of monthly estimates

Calibration is carried out on monthly data for the Norwegian LFS. Quarterly totals (see Section 9) are obtained by taking an average of the corresponding monthly estimates (e.g. Hamre and Heldal, 2013). All months in a quarter given take equal weights in the previous estimation methodology. However, not all the months include the same number of survey weeks. This may cause under– or over–estimation of some variables if they are more sensible to calender weeks (e.g. Hamre, 2017). Therefore, in Section 12.4, we consider an unequal–weighted average of monthly estimates which are weighted proportional to the number of survey weeks in the corresponding months, and investigate the effect of these two types of weighting methods on quarterly estimates.

1.5. Precision requirements and variance estimation

Norwegian LFS is conducted in alignment with the Council Regulation (EC) (1998). According to the final report of the task force on European Union (EU) LFS (EC, 2014), there are precision requirements for employment and unemployment rates to be fulfilled by the member states. In Section 12.5, we explore if these requirements are hold for the Norwegian LFS with the new estimation methodology.

Estimation of sampling variances alongside with the point estimates is crucial as it gives an idea about the magnitude of the sampling error. In this working paper, we use the Jackknife linearisation (JL) variance estimator to estimate the variances of several estimators, since it has good conditional properties and approximates the customary Jackknife variance estimator very well (e.g. Yung and Rao, 1996). It could be used under stratified multi-stage sampling with unequal probabilities provided that the sampling fractions at the first stage within strata are negligible. It can also be used under item and unit non–response (e.g. Yung and Rao, 2000).

A stratified one–stage cluster sampling is used to select sample units for the Norwegian LFS (see Section 3). Effect of clustering on sampling variance may be ignored if there are very good auxiliary information (e.g. Hagesæther and Zhang, 2009). In Section 12.6, we empirically investigate, in the absence and presence of auxiliary information, the cluster effect on variance estimates.

1.6. Sections

The following Sections are organised as follows. Notations are provided in Section 2. The sampling design of the Norwegian LFS is presented in Section 3. The previous estimation methodology which was in use for the Norwegian LFS for many years is explained in Section 4. Parameter of interest is defined in Section 5. Point estimators are given in Sections 6.1-6.5. The JL variance estimator is provided in Section 7.1. Estimation of ratios and quarterly totals are given in Sections 8 and 9, respectively. Domain estimation is provided by Section 10. Calibration models used in the application are described in Sections 11.1-11.2. Numerical results for the Norwegian LFS are presented in Sections 12.1-12.6. Finally, a general discussion is provided by Section 13.

2. Notation

Let U be a finite population of size N stratified into a finite number of H strata denoted by $U_1, \ldots, U_H$, where $\bigcup_{h\in H} U_h = U$ and $\sum_{h\in H} N_h = N$, where N_h denotes the number of units in U_h. Let each U_h consist of N_h disjoint clusters (i.e. households) U_{hi} of sizes K_{hi}, with $\bigcup_{i\in U_h} U_{hi} = U_h$. The total number of individuals in U shall be denoted by M, where $\sum_{h\in H} M_h = M$, with $\sum_{i\in U_h} K_{hi} = M_h$.

Suppose that we have a stratified one–stage cluster sampling where households are the clustering units. Let s_h denote the sample of households, selected with probabilities π_{hi} from U_h, with $\sum_{i \in U_h} \pi_{hi} = n_h$, where the n_h denote the fixed sample sizes. The whole sample of size n shall be denoted by s, where $s = \cup_{h \in H} s_h$ and $n = \sum_{h \in H} n_h$. Let s_{hi} be the sample of individuals, aged 15-74 years, of size k_{hi}, selected with conditional probabilities equal to one, $\pi_{j|hi} = 1$, within the ith sample household, where $j = 1, \ldots, k_{hi}$.

We may not get a full response in the survey because of various reasons. There is a *unit non–response* when all the items are missing for a given sample unit. In the Norwegian LFS, the unit non–response rate is around 20%. Non–response is expected to occur at the household level as indirect interviews may be carried out if necessary. Nevertheless here, we will define a response indicator at individual level which shall be denoted by r_{hij}. We have $r_{hij} = 1$ if individual unit i in the stratum household unit hj responds, and $r_{hij} = 0$ otherwise.

3. The sampling design of the Norwegian LFS

The Norwegian Labour Force Survey (LFS) is a rotational panel sample survey providing monthly data on labour market status in Norway. A sample of households is quarterly selected from the Central Population Register (CPR) with respect to a stratified one-stage cluster sampling (e.g. Hamre and Heldal, 2013). Each household forms the primary sampling unit in the Norwegian LFS. Thus households are the clustering units. The target population for the Norwegian LFS consists of individuals at 15-74 years old. All individuals falling into the target population in the sample households are included in the survey. The population of households is stratified by the third level regional classification (NUTS III), namely county (19) ('fylke' in Norwegian) (see the county list in Table B.1 (see Appendix B)). Total sample size is disproportionally allocated to counties as such that lesser and more populated counties are, respectively, given larger and smaller sample sizes. The ratio of sampling fractions to the overall sampling fraction in each county is given in Table B.1 (see Appendix B).

In each quarter, around $12\,000$ households, or equivalently $24\,000$ individuals, are systematically selected from the stratified population of households. The gross sample of households is randomly distributed over 13 reference weeks in the quarter. The gross sample size of individuals for each month is around $24\,000 * 4/13 = 7\,385$ or $24\,000 * 5/13 = 9\,231$, depending on how many reference weeks, four or five, there are (e.g. Hamre and Heldal, 2013, p.9).

In the rotational panel survey, a sample household stays in the sample for two years, or equivalently, eight consecutive quarters. A panel, which is a sample of households selected into the gross sample at a specific time, is dropped out from the sample and a new panel is introduced into the gross sample to replace it. Thus the $1/8$ of total sample is rotated out each quarter, and the $7/8$ of total sample overlap between two adjacent quarters (e.g. Hamre and Heldal, 2013, p.9).

4. The previous estimation methodology for the Norwegian LFS

The previous estimation methodology which was in use for the Norwegian LFS for many years was first established by Zhang (1998). It can be classified as a two–step GREG estimator. At the first step, design weights are adjusted through a post–stratification procedure, where the post–strata are formed by the cross-classification of five–year age groups from 15 to 74 (12 categories), gender and register based employment, which is further classified into three industry groups if register based employed (4 categories in total: employed in primary, secondary or tertiary industries, or not register based employed). This leads to a total of $2 \times 12 \times 4 = 96$ post–strata. At the second step, a calibration procedure is implemented within each county. This procedure involves calibration against the marginal totals for gender, age and four register based employment groups in each county (i.e. register based employed in primary, secondary or tertiary industries, or not register based employed). Thus the calibration weights are obtained as such that they satisfy $2 + 12 + 4 = 18$ calibration equations within each county, leading to $18 * 19 = 342$ calibration equations overall. A slightly modified version of this estimation procedure (Heldal, 2000) was implemented in the production of Norwegian labour–market statistics, which was programmed in SAS software (SAS Institute Inc., 2013).

The first variation from what was initially suggested by Zhang (1998) reveals itself in the calculation of initial weights, which are not equal to the design weights (4). The former is computed by using the respondent group, but not the original sample selected. Here, in a way that a non–response adjustment is carried out by assuming the MCAR within each county. These adjusted weights are then used in the post–stratification. Apart from this, the initial weights are individual–based, instead of household–based. This is because of practical reasons. The individual–based weighting may not be an issue as long as we have

$$\frac{m_h^r}{n_h^r} \approx \frac{M_h}{N_h}, \tag{1}$$

where M_h is the number of individuals in the population in stratum h, m_h^r is the number of individuals in the respondent sample in stratum h and n_h^r is the number of responding households. Here, (1) means that the the average number of individuals per household in the sample respondent group in stratum h is approximately equal to the average number of individuals per household in population U_h. This is, in fact, is not a starry–eyed assumption.

There are two more variations from the original estimation procedure suggested, one of which is that the biggest age group $70 - 74$ is not further divided into four industry groups in the post–stratification due to the risk of empty cells. Instead, it is divided into two groups: register based employed and not employed. The other variation is that a two–category register based employment status, employed or not employed, instead of four is used in the county level calibration step (Hamre and Heldal, 2013, p.10).

A linearised variance estimator is used for the Norwegian LFS (Hamre and Heldal, 2013). The cluster effect is ignored in the variance estimation. In other words, it is treated as if individuals were selected directly from the population without a household–level clustering. In this way, the sampling variance may be underestimated. However, the extent of the underestimation may be negligible when good auxiliary variables, which could explain the dynamics of the labour force market, are used in the estimation procedure (Hagesæther and Zhang, 2009). Variance estimates with and without taking into account of cluster effect are compared in Tables 16-17.

5. Parameter of interest

Let y_{hij} be the variable of interest associated with the $\{hij\}$th stratum household individual unit. We consider a design-based framework, where the sampling distribution of the sample data $\{y_{hij} : \{hij\} \in s\}$ is only specified by the sampling design. The variable y_{hij} is assumed fixed (non-random) under the design–based framework. Suppose we wish to estimate the population total Y, defined by

$$Y = \sum_{\{hij\} \in U} y_{hij}. \tag{2}$$

We consider several point estimators for (2) that are presented in Sections 6.1-6.5. The estimators are prefixed by *one–step* or *two–step*, except the reference estimator (see Section 6.1), depending on whether or not they involve a separate step for unit non–response adjustment.

6. Point estimators

6.1. Reference estimator

The first estimator of Y is a Hájek (1964) type of estimator defined by

$$\widehat{Y}_H = M \, \frac{\sum_{\{hij\} \in s} d_{hij} r_{hij} y_{hij}}{\sum_{\{hij\} \in s} d_{hij} r_{hij}}, \tag{3}$$

where M is the total number of individuals in the population U, which is assumed to be known, and the d_{hij} are the initial weights given by

$$d_{hij} = (\pi_{hi} \pi_{j|hi})^{-1} = \frac{N_h}{n_h}, \quad \text{with} \quad \{ij : \{ij\} \in U_h\}, \tag{4}$$

as $\pi_{j|hi} = 1$. The estimator (3) does not use any auxiliary information other than the total number of individuals in the population. It is approximately unbiased for Y under full response; that is, $r_{hij} = 1$ for all $\{hij\} \in s$. When there is a unit non–response, the response mechanism has to be *missing completely at random* (MCAR); that is, $\bar{y}^r = \bar{y}$, for the estimator (3) to be an approximately unbiased estimator for Y. Otherwise, it will be biased. Here, $\bar{y}$ is the overall sample mean and $\bar{y}^r$ is the sample mean among the respondent group, which are, respectively, defined by $\bar{y} = \sum_{\{hij\} \in s} d_{hij} y_{hij} / \sum_{\{hij\} \in s} d_{hij}$ and $\sum_{\{hij\} \in s} d_{hij} r_{hij} y_{hij} / \sum_{\{hij\} \in s} d_{hij} r_{hij}$.

6.2. One-step GREG estimator

The generalised regression (GREG) estimator (e.g. Cassel et al., 1976) is a special type of estimator among the class of calibration estimators proposed by Deville and Särndal (1992). The calibration estimator reduces to GREG when a chi–squared distance measure is used. Let $\boldsymbol{x}_{hij}$ be the vector of auxiliary variables associated with the $\{hij\}$th stratum household individual unit, with known population totals $\boldsymbol{X}$. The one–step GREG estimator of Y (Deville and Särndal, 1992) is given by

$$\widehat{Y}_{reg} = \widehat{Y}^r + (\boldsymbol{X} - \widehat{\boldsymbol{X}}^r)^\top \widehat{\boldsymbol{\beta}}^r, \tag{5}$$

where $\widehat{Y}^r$ and $\widehat{\boldsymbol{X}}^r$ are the Horvitz and Thompson (1952) estimators of population totals, respectively, Y and $\boldsymbol{X}$, with $\boldsymbol{X} = \sum_{\{hij\}\in U} \boldsymbol{x}_{hij}$, defined by $\widehat{Y}^r = \sum_{\{hij\}\in s} d_{hij} r_{hij} y_{hij}$ and $\widehat{\boldsymbol{X}}^r = \sum_{\{hij\}\in s} d_{hij} r_{hij} \boldsymbol{x}_{hij}$. The vector of estimated regression coefficient $\widehat{\boldsymbol{\beta}}^r$ is given by

$$\widehat{\boldsymbol{\beta}}^r = (\widehat{\boldsymbol{S}}^r_{xx})^{-1} \widehat{\boldsymbol{S}}^r_{xy}, \tag{6}$$

where

$$\widehat{\boldsymbol{S}}^r_{xx} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \boldsymbol{x}_{hij} \boldsymbol{x}^\top_{hij}, \quad \widehat{\boldsymbol{S}}^r_{xy} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \boldsymbol{x}_{hij} y_{hij}.$$

The regression estimator (5) is equivalent to

$$\widehat{Y}_{reg} = \sum_{\{hij\}\in s} w_{hij} r_{hij} y_{hij},$$

where the w_{hij} are the calibration weights defined by

$$w_{hij} = d_{hij}[1 + \boldsymbol{x}^\top_{hij} (\widehat{\boldsymbol{S}}^r_{xx})^{-1} (\boldsymbol{X} - \widehat{\boldsymbol{X}}^r)], \tag{7}$$

where $\widehat{\boldsymbol{S}}^r_{xx} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \boldsymbol{x}_{hij} \boldsymbol{x}^\top_{hij}$. The calibration weights (7) ensure the consistency with known population total $\boldsymbol{X}$; that is, we have $\sum_{\{hij\}\in s} w_{hij} r_{hij} \boldsymbol{x}_{hij} = \boldsymbol{X}$. A linear regression model is used as an underlying working model for (5). Therefore, it is implicitly assumed that there is a linear relationship between y and $\boldsymbol{x}$.

The estimator (5) does not involve a separate step for adjustment of unit non–response. We aim to achieve three goals at the same time (Särndal and Lundström, 2005):

- reducing non–response bias,
- increasing efficiency,
- ensuring consistency with known population totals.

The estimator (5) is a consistent estimator of Y in the case of full–response. When there is non–response, however, (5) may be assessed under a *model–assisted quasi–randomisation framework* (Nguyen and Zhang, 2016, p.4) . Strictly speaking, one should assume a MCAR response mechanism for design consistency of (5). Otherwise, for example under a *missing–at–random* (MAR) model given $\boldsymbol{x}_{hij}$, the response propensities are assumed to be inversely proportional to the quantity next to the design weights in (7). However, this assumption may not hold in practice (e.g. Haziza and Lesage, 2016; Nguyen and Zhang, 2016).

6.3. One–step (multiple) model–calibration estimator

We gain in efficiency with (5) compared to 3. The GREG estimator (5) is implicitly based on a linear working model. When there is a non–linear relationship between y and $\boldsymbol{x}$, a *model–calibration* estimator proposed by Wu and Sitter (2001) may perform better than the GREG in terms of efficiency if the model is true. A model–calibration estimator uses *complete* auxiliary information unlike the GREG estimator (5), which only uses population totals. The GREG may be favourable when $\boldsymbol{x}$ is not known for all the units in the population. Otherwise, it may worth finding a better incorporation of complete information into the estimation procedure so that we may have even more gain in efficiency.

The one–step model–calibration estimator (Wu and Sitter, 2001) is defined by

$$\widehat{Y}_{mc} = \widehat{Y}^r + \left\{ \sum_{\{hij\}\in U} \mu(\boldsymbol{x}_{hij}, \boldsymbol{\theta}) - \sum_{\{hij\}\in s} d_{hij} r_{hij} \mu(\boldsymbol{x}_{hij}, \boldsymbol{\theta}) \right\} \widehat{\boldsymbol{\beta}}^{r*}, \tag{8}$$

where $\mu(\boldsymbol{x}_{hij}, \boldsymbol{\theta})$ is the conditional expectation of y_{hij} given $\boldsymbol{x}_{hij}$ with respect to the infinite population model defined by

$$E_\xi(y_{hij} \mid \boldsymbol{x}_{hij}) = \mu(\boldsymbol{x}_{hij}, \boldsymbol{\theta}), \quad V_\xi(y_{hij} \mid \boldsymbol{x}_{hij}) = v_{hij}^2 \sigma^2, \tag{9}$$

where $\boldsymbol{\theta}$ and σ^2 are unknown infinite population parameters, v_{hij} is a known function of $\boldsymbol{x}_{hij}$ and $\boldsymbol{\theta}$, and E_ξ and V_ξ are, respectively, the expectation and variance with respect to the infinite population model. Here, $\mu(\boldsymbol{x}_{hij}, \boldsymbol{\theta})$ is a known function of $\boldsymbol{x}_{hij}$ and $\boldsymbol{\theta}$. The vector of estimated regression coefficients $\widehat{\boldsymbol{\beta}}^{r*}$ is given by

$$\widehat{\boldsymbol{\beta}}^{r*} = \left\{ \sum_{\{hij\}\in s} d_{hij} r_{hij} \boldsymbol{\mu}_{hij} \boldsymbol{\mu}_{hij}^\top \right\}^{-1} \sum_{\{hij\}\in s} d_{hij} r_{hij} \boldsymbol{\mu}_{hij} y_{hij},$$

with $\boldsymbol{\mu}_{hij} = \mu(\boldsymbol{x}_{hij}, \boldsymbol{\theta})$.

Linear or non–linear models as well as generalised linear models can be specified by using (9). Model–calibration estimator (8) reduces to the GREG (5) under a linear working model. Model–calibration estimator is design–consistent under full–response. Thus it is robust against model–misspecification (Wu and Sitter, 2001). Moreover, it is an optimum estimator under the model (9) among a class of calibration estimators (Wu, 2003).

There are some drawbacks of using (8) (Montanari and Ranalli, 2009). These are related to consistency and the calibration weights. In the production of official statistics, for example, it may be crucial to ensure consistency with population and sub–population totals. This may not be achieved by a model–calibration estimator if the underlying working–model is not a linear one. Model–calibration estimator (8) requires fitting a separate model for each variable of interest which, in turn, leads to different set of survey weights for each variable. The use of one set of weights is often desirable in the production of official statistics due to the practical reasons, especially when the volume of the statistical production is large. Montanari and Ranalli (2009) proposed a *multiple model–calibration* estimator which tackles these drawbacks. The multiple model–calibration estimator (Montanari and Ranalli, 2009) is given by

$$\widehat{Y}_{mmc} = \widehat{Y}^r + \left(\sum_{\{hij\}\in U} \widehat{\boldsymbol{\eta}}_{hij} - \sum_{\{hij\}\in s} d_{hij} r_{hij} \widehat{\boldsymbol{\eta}}_{hij} \right) \widehat{\boldsymbol{\beta}}^{r**}, \tag{10}$$

where $\widehat{\boldsymbol{\eta}}_{hij} = (\mu(\boldsymbol{x}_{hij}, \widehat{\boldsymbol{\theta}}^r)^\top, \boldsymbol{z}_{hij}^\top)^\top$ and

$$\widehat{\boldsymbol{\beta}}^{r**} = (\widehat{\boldsymbol{S}}^r_{\eta\eta})^{-1} \widehat{\boldsymbol{S}}^r_{\eta y},$$

where

$$\widehat{\boldsymbol{S}}^r_{\eta\eta} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \widehat{\boldsymbol{\eta}}_{hij} \widehat{\boldsymbol{\eta}}_{hij}^\top, \quad \widehat{\boldsymbol{S}}^r_{\eta y} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \widehat{\boldsymbol{\eta}}_{hij} y_{hij}.$$

Here, the working–model parameter $\boldsymbol{\theta}$, which is usually unknown, is replaced by a design–based estimator $\widehat{\boldsymbol{\theta}}^r$, which is defined as the solution of a set of *estimating* *equations* (Wu and Sitter, 2001, p.187). Montanari and Ranalli (2009) showed that (10) is design–consistent under full–response. It is also optimum when the model (9) is true. The vector of variables $\boldsymbol{x}_{hij}$ may be partly or completely included in $\boldsymbol{z}_{hij}$, or they may be completely a different vector of variables. The purpose of having $\boldsymbol{z}_{hij}$ in the calibration model is to fulfil consistency with population totals. One set of weights is obtained like in the case of the GREG (5).

Let $\boldsymbol{\Xi} = \sum_{\{hij\}\in U} \widehat{\boldsymbol{\eta}}_{hij}$ and $\widehat{\boldsymbol{\Xi}}^r = \sum_{\{hij\}\in s} d_{hij} r_{hij} \widehat{\boldsymbol{\eta}}_{hij}$. The calibration weights with the multiple model–calibration estimator (10) is given by

$$w_{hij}^{mmc} = d_{hij} \left\{ 1 + \widehat{\boldsymbol{\eta}}_{hij}^{\top} (\widehat{\boldsymbol{S}}_{\eta\eta}^{r})^{-1} (\boldsymbol{\Xi} - \widehat{\boldsymbol{\Xi}}^r) \right\}. \tag{11}$$

The calibration weights (11) ensure the consistency with known population total $\boldsymbol{Z}$; that is, we have $\sum_{\{hij\}\in s} w_{hij}^{mmc} r_{hij} \boldsymbol{z}_{hij} = \boldsymbol{Z}$.

6.4. Two-step GREG estimator

The two-step GREG estimator involves a separate step for the adjustment of unit non–response in order to reduce the non–response bias. This is carried out in the first step of the estimation procedure. Calibration is performed in the second step. The non–response adjusted weights are used as initial weights in the calibration procedure. Efficiency is achieved when the auxiliary variables are correlated with the variable of interest. Suppose that we have C response homogeneous classes (RHC) (e.g. Särndal et al., 1992, p.578), where a *uniform response mechanism* is hold. Let δ_{hij}^{c} be the RHC indicator with $\delta_{hij}^{c} = 1$ if the $\{hij\}$th stratum household individual unit belongs to the cth RHC, with $c = 1, \ldots, C$, and $\delta_{hij}^{c} = 0$ otherwise. The two-step GREG estimator is given by

$$\widehat{Y}_{reg}^{c} = \widehat{Y}_c + (\boldsymbol{X} - \widehat{\boldsymbol{X}}_c)^{\top} \widehat{\boldsymbol{\beta}}_c, \tag{12}$$

where

$$\widehat{Y}_c = \sum_{c\in C} \sum_{\{hij\}\in s} d_{hij}^{*} r_{hij} \delta_{hij}^{c} y_{hij}, \tag{13}$$

$$\widehat{\boldsymbol{X}}_c = \sum_{c\in C} \sum_{\{hij\}\in s} d_{hij}^{*} r_{hij} \delta_{hij}^{c} \boldsymbol{x}_{hij}, \tag{14}$$

$$\widehat{\boldsymbol{\beta}}_c = (\widehat{\boldsymbol{S}}_{c;xx})^{-1} \widehat{\boldsymbol{S}}_{c;xy}, \tag{15}$$

where

$$\widehat{\boldsymbol{S}}_{c;xx} = \sum_{c\in C} \sum_{\{hij\}\in s} d_{hij}^{*} r_{hij} \delta_{hij}^{c} \boldsymbol{x}_{hij} \boldsymbol{x}_{hij}^{\top}, \quad \widehat{\boldsymbol{S}}_{c;xy} = \sum_{c\in C} \sum_{\{hij\}\in s} d_{hij}^{*} r_{hij} \delta_{hij}^{c} \boldsymbol{x}_{hij} y_{hij},$$

with

$$d_{hij}^{*} = \frac{\widehat{M}_c}{\widehat{M}_c^{r}} d_{hij}, \quad \text{with } \delta_{hij}^{c} = 1, \tag{16}$$

where $\widehat{M}_c = \sum_{\{hij\}\in s} d_{hij} \delta_{hij}^{c}$ and $\widehat{M}_c^{r} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \delta_{hij}^{c}$, are the non–response adjusted weights, which are used as initial weights in calibration instead of (4). The two-step GREG estimator (12) may be re-written as follows.

$$\widehat{Y}_{reg}^{c} = \sum_{c\in C} \sum_{\{hij\}\in s} w_{hij}^{*} r_{hij} \delta_{hij}^{c} y_{hij},$$

where the w^*_{hij} are the calibration weights defined by

$$w^*_{hij} = d^*_{hij}\,[1 + \boldsymbol{x}^\top_{hij}(\widehat{\boldsymbol{S}}_{c;xx})^{-1}(\boldsymbol{X} - \widehat{\boldsymbol{X}}_c)]\cdot \tag{17}$$

The two–step GREG estimator is a consistent estimator of (2) when the assumption of the MAR given the RHCs is true. Otherwise, it will be biased. Yet, the bias may be reduced to a certain extent if the non–response pattern is partially explained by the RHCs and/or the vector of $\boldsymbol{x}$.

6.5. Two–step (multiple) model–calibration estimator

Let $\widehat{\boldsymbol{\Xi}}_c = \sum_{c\in C}\sum_{\{hij\}\in s} d^*_{hij} r_{hij}\delta^c_{hij}\widehat{\boldsymbol{\eta}}_{hij}$. The two–step (multiple) model–calibration estimator involves a separate step for non–response adjustment similar to the two–step GREG estimator (12). Thus it is accordingly defined as follows.

$$\widehat{Y}^c_{mmc} = \widehat{Y}_c + \left(\boldsymbol{\Xi} - \widehat{\boldsymbol{\Xi}}_c\right)\widehat{\boldsymbol{\beta}}^{**}_c, \tag{18}$$

where

$$\widehat{\boldsymbol{\beta}}^{**}_c = \left\{\sum_{c\in C}\sum_{\{hij\}\in s} d^*_{hij} r_{hij}\delta^c_{hij}\widehat{\boldsymbol{\eta}}_{hij}\widehat{\boldsymbol{\eta}}^\top_{hij}\right\}^{-1}\sum_{c\in C}\sum_{\{hij\}\in s} d^*_{hij} r_{hij}\delta^c_{hij}\widehat{\boldsymbol{\eta}}_{hij}y_{hij}.$$

We can re–express (18) by

$$\widehat{Y}^c_{mmc} = \sum_{c\in C}\sum_{\{hij\}\in s} w^{mmc*}_{hij} r_{hij}\delta^c_{hij}y_{hij},$$

where the w^{mmc*}_{hij} are the calibration weights defined by

$$w^{mmc*}_{hij} = d^*_{hij}\left\{1 + \widehat{\boldsymbol{\eta}}^\top_{hij}\,(\widehat{\boldsymbol{S}}_{c;\eta\eta})^{-1}\left(\boldsymbol{\Xi} - \widehat{\boldsymbol{\Xi}}_c\right)\right\},$$

where

$$\widehat{\boldsymbol{S}}_{c;\eta\eta} = \sum_{c\in C}\sum_{\{hij\}\in s} d^*_{hij} r_{hij}\delta^c_{hij}\widehat{\boldsymbol{\eta}}_{hij}\widehat{\boldsymbol{\eta}}^\top_{hij}.$$

The two–step (multiple) model–calibration estimator is design–consistent when there is the MCAR within each RHC. Otherwise, it will be biased although this bias may be reduced to a certain degree depending on how well the RHCs and/or the vector of $\boldsymbol{x}$ explains the non–response mechanism.

7. Variance estimation

Suppose that we wish to estimate, by assuming full response, the variance of the Horvitz and Thompson (1952) estimator of Y, which is unbiased, defined by

$$\widehat{Y}_{HT} = \sum_{\{hij\}\in s} d_{hij}y_{hij}\cdot \tag{19}$$

Assuming that the sampling fractions at the first stage of sample selection, n_h/N_h, are negligible as $n_h \to \infty$ and $N_h \to \infty$, the sample s_h including without replacement set of units is asymptotically equivalent to the sample of with replacement set of units (p.112 Hájek, 1981). This assumption holds for the most household surveys including the LFSs. Thus a variance estimator of (19), by applying an *ultimate cluster approach* (Hansen et al., 1953), is given by

$$v(\widehat{Y}_{HT}) = \sum_{h\in H}\frac{n_h}{n_h - 1}\left\{\sum_{i\in s_h}\widehat{y}^2_{hi} - \frac{1}{n_h}(\sum_{i\in s_h}\widehat{y}_{hi})^2\right\}, \tag{20}$$

where $\widehat{y}_{hi} = \sum_{j \in s_{hi}} d_{hij} y_{hij}$. The variance estimator (20) is called the Hansen and Hurwitz (1943) variance estimator. It can be used under multi-stage sampling if the sampling fractions at the first stage, n_h/N_h are negligible.

In practice, we have often non–response in survey data. Population level information is also used to improve estimates. In Section 7.1, a variance estimator that takes into account the design, non–response and population level information is presented.

7.1. The Jackknife linearisation (JL) variance estimator

We propose using the JL variance estimator to estimate variances of several statistics in the Norwegian LFS. This variance estimator has good conditional properties and approximates the customary Jackknife variance estimator very well (e.g. Yung and Rao, 1996). It is not computer intensive like the customary Jackknife variance estimator. It could be used under stratified multi-stage sampling with unequal probabilities provided that the sampling fractions at the first stage within strata are negligible. It can also be used under item and unit non–response (e.g. Yung and Rao, 2000). It is simple to implement to totals or ratios. However, more analytic derivations are required for application to general smooth statistics.

When the statistics of interest is linear in (19), all units are respondent and no population level information is used, both the customary Jackknife and the JL variance estimators are identical to the customary variance estimator (20).

The linearisation approach may be used to estimate variances of complex statistics (e.g. Deville, 1999). A '*cookbook approach*' is proposed by Binder (1996) for derivation of linearised variables for several complex statistics. It is quite practical to apply. In the following Sections, we use the cookbook approach to derive the linearised variables for the estimators presented in Sections 6.1-6.5.

JL variance estimator for the reference estimator

The variance of (3) may be estimated by using linearised variables, which can be derived by using the cookbook approach (Binder, 1996). The reference estimator $\widehat{Y}_H$ (3) may be re–expressed as a function of estimated totals as follows.

$$\widehat{Y}_H = f(\widehat{Y}^r, \widehat{M}^r, N) = M \frac{\widehat{Y}^r}{\widehat{M}^r}, \tag{21}$$

with $\widehat{M}^r = \sum_{\{hij\} \in s} d_{hij} r_{hij}$. We have a ratio estimator on the right hand side of (21). As N is known, total differentials are only applied to estimated totals $\widehat{Y}^r$ and $\widehat{M}^r$. We obtain, by using the cookbook approach,

$$\{\mathrm{d}\widehat{Y}_H\} = \frac{M}{\widehat{M}^r} \left(\{\mathrm{d}\widehat{Y}^r\} - \frac{\widehat{Y}^r}{\widehat{M}^r} \{\mathrm{d}\widehat{M}^r\} \right) \cdot \tag{22}$$

Binder (1996) proposed replacing the total differential of an estimated total by deviation from its expected value. When we apply this to (22), we obtain

$$\widehat{Y}_H - Y \doteq \frac{M}{\widehat{M}^r} \left(\{\widehat{Y}^r - Y\} - \frac{\widehat{Y}^r}{\widehat{M}^r} \{\widehat{M}^r - M\} \right),$$

or equivalently,

$$\widehat{Y}_H - Y \doteq \frac{M}{\widehat{M}^r} \sum_{j \in s_{hi}} d_{hij} r_{hij} e^y_{hij} + \Omega_0,$$

where Ω_0 denotes the remaining terms not depending explicitly on d_{hij} (Binder, 1996, p.18) and $e^y_{hij} = y_{hij} - \bar{y}^r$, with $\bar{y}^r = \widehat{Y}^r/\widehat{M}^r$. Thus the JL variance estimator of $\widehat{Y}_H$ is given by

$$v_{\mathrm{JL}}(\widehat{Y}_H) = M^2 v(\widehat{e}^y_{hi}) = M^2 \sum_{h \in H} \frac{n_h}{n_h - 1} \left\{ \sum_{i \in s_h} (\widehat{e}^y_{hi})^2 - \frac{1}{n_h} (\sum_{i \in s_h} \widehat{e}^y_{hi})^2 \right\}, \tag{23}$$

where

$$\widehat{e}^y_{hi} = \frac{1}{\widehat{M}^r} \sum_{j \in s_{hi}} d_{hij} r_{hij} e^y_{hij}.$$

JL variance estimator for the one–step GREG estimator

The one–step GREG estimator (5) is a linear function of $\widehat{Y}^r$, $\widehat{\boldsymbol{X}}^r$ and $\widehat{\boldsymbol{\beta}}^r$. When we take the total differentials of both sides of (5), we obtain

$$\{\mathrm{d}\widehat{Y}_{reg}\} = \{\mathrm{d}\widehat{Y}^r\} + (\boldsymbol{X} - \{\mathrm{d}\widehat{\boldsymbol{X}}^r\})^\top \widehat{\boldsymbol{\beta}}^r + (\boldsymbol{X} - \widehat{\boldsymbol{X}}^r)^\top \{\mathrm{d}\widehat{\boldsymbol{\beta}}^r\}. \tag{24}$$

The regression coefficient $\widehat{\boldsymbol{\beta}}^r$ can further be written as a function of estimated totals. We obtain, by applying total differentials to (6),

$$\{\mathrm{d}\widehat{\boldsymbol{\beta}}^r\} = (\widehat{\boldsymbol{S}}^r_{xx})^{-1} \left(\{\mathrm{d}\widehat{\boldsymbol{S}}^r_{xy}\} - \widehat{\boldsymbol{\beta}}^r \{\mathrm{d}\widehat{\boldsymbol{S}}^r_{xx}\} \right). \tag{25}$$

Total differentials $\{\mathrm{d}\widehat{Y}^r\}$, $\{\mathrm{d}\widehat{\boldsymbol{X}}^r\}$, $\{\mathrm{d}\widehat{\boldsymbol{S}}^r_{xy}\}$ and $\{\mathrm{d}\widehat{\boldsymbol{S}}^r_{xx}\}$ can be, respectively, replaced by $\widehat{Y}^r - Y$, $\widehat{\boldsymbol{X}}^r - \boldsymbol{X}$, $\widehat{\boldsymbol{S}}^r_{xy} - \boldsymbol{S}_{xy}$ and $\widehat{\boldsymbol{S}}^r_{xx} - \boldsymbol{S}_{xx}$, where $\boldsymbol{S}_{xy} = \sum_{\{hij\} \in U} \boldsymbol{x}_{hij} \boldsymbol{x}^\top_{hij}$ and $\boldsymbol{S}_{xx} = \sum_{\{hij\} \in U} \boldsymbol{x}_{hij} y_{hij}$, in (24)-(25) (Binder, 1996). After some algebra, we obtain

$$\{\mathrm{d}\widehat{Y}_{reg}\} = \widehat{Y}_{reg} - Y \doteq \sum_{\{hij\} \in s} w_{hij} r_{hij} \epsilon^*_{hij} + \Omega^*,$$

where

$$\epsilon^*_{hij} = y_{hij} - \widehat{\boldsymbol{\beta}}^r \boldsymbol{x}_{hij},$$

w_{hij} is defined by (7) and Ω^* contains the terms not depending explicitly on d_{hij}. Thus the JL variance estimator of the one-step regression estimator is given by

$$v_{\mathrm{JL}}(\widehat{Y}_{reg}) = v(\widehat{\epsilon}^*_{hi}) = \sum_{h \in H} \frac{n_h}{n_h - 1} \left\{ \sum_{i \in s_h} (\widehat{\epsilon}^*_{hi})^2 - \frac{1}{n_h} (\sum_{i \in s_h} \widehat{\epsilon}^*_{hi})^2 \right\}, \tag{26}$$

where

$$\widehat{\epsilon}^*_{hi} = \sum_{j \in s_{hi}} w_{hij} r_{hij} \epsilon^*_{hij}. \tag{27}$$

The variance estimator (26) is similar to the model-assisted variance estimator suggested by Deville and Särndal (1992, p.380). They are, in fact, asymptotically equivalent (Yung and Rao, 1996). It is not only design-consistent, but also nearly model-unbiased. When we use d_{hij} in (27), (26) becomes equivalent to the *standard linearisation variance estimator*, which treats $\widehat{\boldsymbol{\beta}}^r$ known. In practice, the fact that $\widehat{\boldsymbol{\beta}}^r$ is estimated is often ignored.

JL variance estimator for the one–step model–calibration estimator

Linearised variables for the one–step multiple model–calibration estimator (10) can be obtained in a similar way to the one–step GREG estimator. Applying total differentials to both sides of (10) leads to

$$\{\mathrm{d}\widehat{Y}_{mmc}\} = \{\mathrm{d}\widehat{Y}^r\} + (\boldsymbol{\Xi} - \{\mathrm{d}\widehat{\boldsymbol{\Xi}}^r\})^\top \widehat{\boldsymbol{\beta}}^{r**} + (\boldsymbol{\Xi} - \widehat{\boldsymbol{\Xi}}^r)^\top \{\mathrm{d}\widehat{\boldsymbol{\beta}}^{r**}\}, \tag{28}$$

where

$$\{\mathrm{d}\widehat{\boldsymbol{\beta}}^{r**}\} = (\widehat{\boldsymbol{S}}^r_{\eta\eta})^{-1}\left(\{\mathrm{d}\widehat{\boldsymbol{S}}^r_{\eta y}\} - \widehat{\boldsymbol{\beta}}^{r**}\{\mathrm{d}\widehat{\boldsymbol{S}}^r_{\eta\eta}\}\right)\cdot \tag{29}$$

We obtain, after some algebra by using (28) and (29),

$$\{\mathrm{d}\widehat{Y}_{mmc}\} = \widehat{Y}_{mmc} - Y \doteq \sum_{\{hij\}\in s} w^{mmc}_{hij} r_{hij} \epsilon^{**}_{hij} + \Omega^{**},$$

where

$$\epsilon^{**}_{hij} = y_{hij} - \widehat{\boldsymbol{\beta}}^{r**}\widehat{\boldsymbol{\eta}}_{hij},$$

w^{mmc}_{hij} is defined by (11) and Ω^{**} contains the terms not depending explicitly on d_{hij}. Thus the JL variance estimator of the one-step multiple model–calibration estimator can be obtained by replacing $\widehat{\epsilon}^{*}_{hi}$ with $\widehat{\epsilon}^{**}_{hi}$ in (26) and (27), where $\widehat{\epsilon}^{**}_{hi} = \sum_{j\in s_{hi}} w^{mmc}_{hij} r_{hij}\epsilon^{**}_{hij}$. Here, the fact that $\boldsymbol{\theta}$ is estimated is ignored in the variance estimation.

JL for the two-step GREG estimator

The variance of (12) may be estimated by using the cookbook approach (Binder, 1996). At first, we apply total differentials to both sides of (12). This leads to

$$\{\mathrm{d}\widehat{Y}^c_{reg}\} = \{\mathrm{d}\widehat{Y}_c\} + (\boldsymbol{X} - \{\mathrm{d}\widehat{\boldsymbol{X}}_c\})^\top \widehat{\boldsymbol{\beta}}_c + (\boldsymbol{X} - \widehat{\boldsymbol{X}}_c)^\top \{\mathrm{d}\widehat{\boldsymbol{\beta}}_c\}\cdot \tag{30}$$

The total differentials on the right hand side of equation (30) are derived by rewriting $\widehat{Y}_c$, $\widehat{\boldsymbol{X}}_c$ and $\widehat{\boldsymbol{\beta}}_c$ as functions of estimated totals in Appendix A.

Using (A.3), (A.4), (A.5), (A.8) and (A.9) (see Appendix A), and after some algebra, we obtain

$$\{\mathrm{d}\widehat{Y}^c_{reg}\} = \widehat{Y}^c_{reg} - Y \doteq \sum_{c\in C}\sum_{\{hij\}\in s} d_{hij}\delta^c_{hij} e_{hij} + \Omega, \tag{31}$$

where

$$e_{hij} = \widehat{a}_c g_{hij} r_{hij}\epsilon_{hij} + \bar{\epsilon}^r_c(1 - \widehat{a}_c r_{hij}), \tag{32}$$

where g_{hij} are the g-weights defined by

$$g_{hij} = [1 + \boldsymbol{x}^\top_{hij}(\widehat{\boldsymbol{S}}_{c;xx})^{-1}(\boldsymbol{X} - \widehat{\boldsymbol{X}}_c)],$$

where $\widehat{\boldsymbol{S}}_{c;xx}$ is given by (A.7), and

$$\bar{\epsilon}^r_c = \frac{1}{\widehat{M}^r_c}\sum_{\{hij\}\in s} d_{hij} g_{hij}\delta^c_{hij} r_{hij}\epsilon_{hij},$$

with

$$\epsilon_{hij} = y_{hij} - \widehat{\boldsymbol{\beta}}_c \boldsymbol{x}_{hij}\cdot$$

Therefore, the JL variance estimator of (12) is given by (e.g Yung and Rao, 2000)

$$v_{\mathrm{JL}}(\widehat{Y}^c_{reg}) = v(\widehat{e}_{hi}) = \sum_{h\in H} \frac{n_h}{n_h - 1}\left\{\sum_{i\in s_h} \widehat{e}^2_{hi} - \frac{1}{n_h}(\sum_{i\in s_h} \widehat{e}_{hi})^2\right\}, \tag{33}$$

where

$$\widehat{e}_{hi} = \sum_{c\in C}\sum_{j\in s_{hi}} d_{hij}\delta^c_{hij}e_{hij}.$$

The variance estimator (33) takes into account the sampling design, unit non–response and population level information. It can be noticed that the JL variance estimators of (3) and (5); that is, (23) and (26), are the special cases of (33).

JL for the two-step multiple model–calibration estimator

The variance of (18) can be obtained in a similar way as in Section 1. We can easily show that

$$\{\mathrm{d}\widehat{Y}^c_{mmc}\} = \widehat{Y}^c_{mmc} - Y \doteq \sum_{c\in C}\sum_{\{hij\}\in s} d_{hij}\delta^c_{hij}e^{mmc}_{hij} + \Omega_{mmc}, \tag{34}$$

where

$$e^{mmc}_{hij} = \widehat{a}_c\, g^{mmc}_{hij} r_{hij}\, \epsilon^{mmc}_{hij} + \bar{\epsilon}^{r*}_c(1 - \widehat{a}_c\, r_{hij}), \tag{35}$$

where g^{mmc}_{hij} are the g-weights defined by

$$g^{mmc}_{hij} = [1 + \widehat{\boldsymbol{\eta}}^{\top}_{hij}(\widehat{\boldsymbol{S}}_{c;\eta\eta})^{-1}(\boldsymbol{\Xi} - \widehat{\boldsymbol{\Xi}}_c)],$$

and

$$\bar{\epsilon}^r_c = \frac{1}{\widehat{M}^r_c}\sum_{\{hij\}\in s} d_{hij}g^{mmc}_{hij}\delta^c_{hij}r_{hij}\epsilon^{mmc}_{hij},$$

with

$$\epsilon^{mmc}_{hij} = y_{hij} - \widehat{\boldsymbol{\beta}}^{**}_c\widehat{\boldsymbol{\eta}}_{hij}.$$

Thus the JL variance estimator of (18) can be obtained by replacing $\widehat{e}_{hi}$ with $\widehat{e}^{mmc}_{hi}$ in (33), where $\widehat{e}^{mmc}_{hi} = \sum_{c\in C}\sum_{j\in s_{hi}} d_{hij}\delta^c_{hij}e^{mmc}_{hij}$.

8. Estimation of a ratio

Suppose that the parameter of interest is a ratio of two population totals defined by

$$R = \frac{Y}{W}, \tag{36}$$

where $Y = \sum_{\{hij\}\in U} y_{hij}$ and $W = \sum_{\{hij\}\in U} w_{hij}$. For example, we may be interested in the unemployment rate, which is defined as the ratio of people unemployed, Y, among all 15-74-year-old people in the labour force, W. Let $\widehat{Y}$ and $\widehat{W}$ be any of the design–based estimators, which are defined in Sections 6.1-6.5, of Y and W, respectively. Thus a design–based estimator of R is given by

$$\widehat{R} = \frac{\widehat{Y}}{\widehat{W}}. \tag{37}$$

Taking total differentials of both sides of (37) leads to

$$\{\mathrm{d}\widehat{R}\} = \frac{1}{\widehat{W}}\left(\{\mathrm{d}\widehat{Y}\} - \widehat{R}\{\mathrm{d}\widehat{W}\}\right)\cdot$$

The total differentials of $\widehat{Y}$ and $\widehat{W}$ can be accordingly obtained depending on which design–based estimator is used. Let e^{y}_{hij} and e^{w}_{hij} be the linearised variables associated with $\widehat{Y}$ and $\widehat{W}$, respectively, after applying the cookbook approach (Binder, 1996). Thus the linearised variable for (37) can be defined by

$$\tau_{hij} = \frac{1}{\widehat{W}}(e^{y}_{hij} - \widehat{R}\, e^{w}_{hij}),$$

Therefore, the JL variance estimator of $\widehat{R}$ can be obtained by replacing $\widehat{y}_{hi}$ with $\widehat{\tau}_{hi}$ in (20), where $\widehat{\tau}_{hi}$ is a design–based estimator of the cluster total of τ_{hij}.

9. Estimation of quarterly totals

Point estimation for the Norwegian LFS is carried out monthly. Equal– or unequal–weighted averages of the monthly estimates may be used to estimate quarterly totals. Let $\widehat{Y}_{m_t}$ be a design–based estimator of Y for the t-th month in a given quarter, where m stands for 'month' and $t \in \{1, 2, 3\}$. An estimator of a quarterly total is defined by

$$\widehat{Y}_q = \sum_{t\in\{1,2,3\}} f_{m_t}\widehat{Y}_{m_t}, \tag{38}$$

where the f_{m_t} are the weights given to each month in the quarter of interest. An equal–weighted average of monthly estimates, where $f_{m_t} = 1/3$, was used in the previous estimation methodology (see Section 4). We suggest using an unequal–weighted average, where the f_{m_t} are proportional to the number of survey weeks in the relevant months, in the new estimation methodology (see Section 12.4). In this case, the f_{m_t} are given by $4/13$ and $5/13$ for a month with four and five weeks, respectively. We shall call the unequal–weighted averaging method the *weekly–weighted* averaging method henceforth.

The variance estimator of (38) is given by

$$v_{\mathrm{JL}}(\widehat{Y}_q) = \sum_{t\in\{1,2,3\}} f^2_{m_t}\, v_{\mathrm{JL}}(\widehat{Y}_{m_t})$$

as monthly samples are independent from each other due to the random allocation of quarterly sample to the weeks of a quarter. The expression for the variance estimator $v_{\mathrm{JL}}(\widehat{Y}_{m_t})$ depends on the estimator used for monthly totals.

10. Domain estimation

Estimation over domains are important in LFSs. For example, age, sex and regional distribution of labour market may be an particular interest of researchers and policy makers. Therefore, in this Section, point and variance estimation over domains shall be presented. Let Φ be the domain of interest and ϕ_{hij} be a domain indicator for stratum household individual unit $\{hij\}$ defined by

$$\phi_{hij} = \begin{cases} 1 & \text{if } \{hij\} \in \Phi, \\ 0 & \text{if } \{hij\} \notin \Phi\cdot \end{cases}$$

A population total over a domain is defined by

$$Y_\Phi = \sum_{\{hij\}\in U} \phi_{hij} y_{hij}. \tag{39}$$

We consider an estimator of a domain total that involves the use of ϕ_{hij} wherever the variable of interest, y_{hij}, appears in Sections 6.1-6.5. This applies also to variance estimators of domain point estimators. Thus y_{hij} is replaced with $\phi_{hij} y_{hij}$ in Section 7.1.

Domains may not be necessarily given by design strata. They may cross-over strata. Thus domain sizes are random if domain of interest is not used in the design. Here, we assume that we have enough number of sample observations in domains of interest such that we have reliable estimates. This is a reasonable assumption for the Norwegian LFS as publication domains are usually large.

A ratio over a domain is given by

$$R_\Phi = \frac{Y_\Phi}{W_\Phi},$$

where $Y_\Phi = \sum_{\{hij\}\in U} \phi_{hij} y_{hij}$ and $W_\Phi = \sum_{\{hij\}\in U} \phi_{hij} w_{hij}$. In order to obtain point and variance estimates for a ratio, we may replace y_{hij} and w_{hij} with $\phi_{hij} y_{hij}$ and $\phi_{hij} w_{hij}$, respectively, in the expressions of the relevant estimator (see Sections 6.1-8).

11. Application: Norwegian LFS

In the application, monthly Norwegian LFS 2015, 2016 and 2017 data are used. Several calibration models are considered for one–step and two–step GREG and multiple model–calibration estimators. Here, we aim to find the best estimator among others in terms of efficiency and unbiasedness for the Norwegian LFS as well as empirically respond several research questions provided below.

- Do the two–step weighting approaches reduce the non–response bias more than those with one–step weighting?

- Which estimator is better for ratios, where both enumerator and dominator are estimated: GREG or multiple model–calibration?

- Is the multiple model–calibration estimator more efficient than the proxy method to the previous estimation method used by SSB over important publication domains?

- What type of averaging method should be used to estimate quarterly totals: the equal– versus weekly–weighted average of monthly estimates?

- Does the multiple model–calibration estimator provide estimates satisfying the precision requirement of EU for national employment and unemployment rates and regional level unemployment rates?

- When may clustering have significant effect on the sampling variance?

The models used for the estimators are described in Sections 11.1-11.2. A description of variables used in these models are provided by Table B.2 (see Appendix B).

11.1. One–step estimation

Two different calibration models are considered for the one–step GREG estimator (5).

- **One–step GREG model 1** is given by:

$$\mathbf{PS} := \sim regemp\,(2) \times age\,(11) \times gender\,(2),$$

where the variables in the model are all categorical and the numbers within the parentheses show the number of categories for each (see Table B.2 for details). Here, '$\times$' refers to cross–classification; that is, three–way interaction between variables. The biggest age group $70 - 74$ is merged with the nest biggest $65 - 69$ to avoid empty and/or small cells. The model PS is one of models used by Nguyen and Zhang (2016) as an analysis model for employment and unemployment. They call it SSB model although it should not be seen as the previous estimation method used by SSB. The PS stands for the *post–stratified* estimator.

The post–stratified estimator is a special case of GREG estimator. The weights obtained with the model PS, under the model–assisted framework (e.g. Deville and Särndal, 1992), is independent of the distance function chosen in the calibration procedure (e.g. Haziza and Lesage, 2016, p.143). Therefore, whatever the distance function is, we obtain the same set of weights (7) , and so the same GREG estimator. Thus the PS estimator is unbiased for Y provided that we have the MCAR within each post–strata, or the MAR mechanism given post–stratum cells. Even the MAR assumption does not hold, The PS estimator may still decrease the non–response bias to a certain extent as well as it increases the efficiency provided that the variables used in the model are highly associated with the variables of interest and the response mechanism (e.g. Little and Vartivarian, 2005, p.7).

- **One–step GREG model 2** is given by:

$$\mathbf{GREG}_{\mathbf{1\text{–}stp}} := \sim age\,(13) + gender\,(2) \times [\{regemp\,(7) + age\,(6) + income\,(5)\} + \{regemp\,(2) \times marstat\,(2) \times age\,(2)\} + \{region\,(7) \times age\,(3) \times regemp\,(2)\}] + education\,(4) + tiltak\,(3) + country\,(3),$$

where '$\times$' refers to interactions between variables. Variables in the model $\mathrm{GREG}_{1\text{–stp}}$ were decided based on their relationships with the employment status (a multinomial logistic regression model was fitted to see which variables significantly affect the employment status), the expert views provided by the Division for Labour Market and Wage Statistics and national users' needs (see Table B.2 for variable descriptions). The one–step GREG estimator with the model $\mathrm{GREG}_{1\text{–stp}}$ is unbiased provided that the MCAR characterise the response mechanism. Otherwise, it may exhibit some bias, which may be reduced to a certain degree depending on how good the variables in the model explain the variables of interest.

The following calibration model is considered for the one–step multiple model–calibration estimator (10).

- **One–step multiple model–calibration model** is given by:

$$\mathbf{MC}_{\mathbf{1\text{–}stp}} := \sim age\,(12) \times gender\,(2) + age\,(8) + [age\,(3) \times gender\,(2) \times \{\widehat{p}_e + \widehat{p}_u + \widehat{p}_o\}] + [region\,(7) \times \{\{\widehat{p}_e + \widehat{p}_u + \widehat{p}_o + age\,(3) + gender\,(2)\}] + regemp\,(4) \times gender\,(2),$$

where $\widehat{p}_e$, $\widehat{p}_u$ and $\widehat{p}_o$ are the predicted probabilities for an individual to be, respectively, employed, unemployed and outside of the labour force, and '$\times$' refers to interactions between variables (see Table B.2 for variable descriptions). These probabilities are predicted by fitting a *multinomial logistic model*. We have $\widehat{p}_e+\widehat{p}_u+\widehat{p}_o=1$ for each individual. Variables in the model are given by

$$\begin{aligned}\text{Mnomlog} \sim\ & gender\,(2) * [age\,(13) + regemp\,(7) + \{age\,(11) * regemp\,(2)\} + \\ & \{regemp\,(2) * marstat\,(2) * age\,(2)\}] + education\,(4) + county\,(19) + \\ & familysize\,(3) + tiltak\,(3) + country\,(3),\end{aligned}$$

where '$*$' refers to all way interactions between variables up to the order equal to the number of variables where it appears (see Table B.2 for variable descriptions).

11.2. Two–step estimation

Two–step GREG and multiple model–calibration estimators, as mentioned in Sections 6.4-6.5, involves a separate phase for non–response adjustment. We assume the MCAR mechanism within RHCs, which are constructed in two different ways in the application here. Non–response adjustment is performed within these classes (see (16)). Second phase involves calibration. Two calibration models are considered for the two–step GREG estimator.

First phase: formation of RHC

Two ways are followed to form RHCs. The first one is based on the cross–classification of registered employment status (2 groups), age (11 groups) and gender (2 groups) (see Table B.2 for variable descriptions). Unbiased estimation requires that the non–response mechanism within each of these cells is MCAR. The RHCs are formed by

$$\mathbf{RHC}_1 \coloneqq\ \sim regemp\,(2) \times age\,(11) \times gender\,(2),$$

where '$\times$' refers to cross–classification between variables.

The second method for the construction of the RHCs involves allocating sample units into homogeneous classes in terms of their response propensities, p, which are predicted by using a logistic model, which is given by

$$\begin{aligned}\text{logit}\,\widehat{p} \sim\ & age\,(12) + regemp\,(7) + education\,(4) + marstat\,(2) + country\,(3) + \\ & income\,(5) + famsize\,(3)\cdot\end{aligned}$$

Here, p is defined as the probability of response given a set of explanatory variables denoted by $\boldsymbol{X}$ such that $p = \mathrm{P}(r = 1 \mid \boldsymbol{X} = \boldsymbol{x})$, where r is the response indicator (see Section 2). After predicting p, sample units are assigned into five homogeneous classes by using the *K–means clustering* method. The use of these RHCs instead of the direct use of the predicted response propensities, $\widehat{p}$, may provide some robustness against a model misspecification for the response propensities (e.g Haziza and Lesage, 2016, p.134). Thus the second type of RHCs are denoted by

$$\mathbf{RHC}_2 \coloneqq\ \mathcal{C}_5(p),$$

where $\mathcal{C}_5(\widehat{p})$ refers to the five RHCs formed by the K–means clustering of the predicted response propensities. Five classes are usually enough to reduce non–response bias to a certain extent. More classes may just increase the variance without providing a significant improvement in the bias, which may compensate the price paid (e.g. Eltinge and Yansaneh, 1997, p.35-36).

Two phases combined: non–response plus calibration

The weights obtained in the first phase (non–response adjustment within the RHC) are used as initial weights in the calibration procedure, which shall be called the *second phase* in the estimation procedure. The first three estimation methods presented below are examples of the two–step GREG estimator given by (12).

- **Two–step GREG estimation method 1**, which shall be denoted by $\text{GREG}^{\text{ssb}}_{\text{2–stp}}$, is obtained by slightly modifying the previous estimation method used by SSB (see Section 4).

  First phase $:=$ RHC_1,

  Second phase $:=$ $\sim region\,(7) : \{regemp\,(2) + age\,(11) + gender\,(2)\}\cdot$

Since the two–step method $\text{GREG}^{\text{ssb}}_{\text{2–stp}}$ is a slightly modified version of the method used by SSB, we shall call $\text{GREG}^{\text{ssb}}_{\text{2–stp}}$ a *proxy* SSB *method*. The reason of not using exactly the same calibration model (in the second phase) as the previous one is to avoid having very small cells which may lead to convergence problems in calibration constraints, and eventually large survey weights and sampling variance. Because of these reasons, one may anticipate that the variance estimate of the proxy estimator $\text{GREG}^{\text{ssb}}_{\text{2–stp}}$ is equal to or lower than the original previous estimator used by SSB (see Section 4). Therefore, any method superior than the former, $\text{GREG}^{\text{ssb}}_{\text{2–stp}}$, will also be superior than the latter in terms of efficiency.

- **Two–step GREG estimation method 2**, which shall be denoted by $\text{GREG}^{\text{rhc1}}_{\text{2–stp}}$, involves combination of model $\text{GREG}_{\text{1–stp}}$ and the first type of non–response adjustment cells, RHC_1, such that

  First phase $:=$ RHC_1,

  Second phase $:=$ model $\text{GREG}_{\text{1–stp}}$.

- **Two–step GREG estimation method 3**, which shall be denoted by $\text{GREG}^{\text{rhc2}}_{\text{2–stp}}$, involves combination of model $\text{GREG}_{\text{1–stp}}$ and the second type of non–response adjustment cells, RHC_2, such that

  First phase $:=$ RHC_2,

  Second phase $:=$ model $\text{GREG}_{\text{1–stp}}$.

We consider two estimation methods for the two–step multiple model–calibration estimator given by (18). These are as follows.

- **Two–step multiple model–calibration estimation method 1**, which shall be denoted by $\text{MC}^{\text{rhc1}}_{\text{2–stp}}$, involves combination of model $\text{MC}_{\text{1–stp}}$ and the first type of non–response adjustment cells , RHC_1, such that

  First phase $:=$ RHC_1,

Second phase $:=$ model $\text{MC}_{1\text{–stp}}$.

- **Two–step multiple model–calibration estimation method 2**, which shall be denoted by $\text{MC}^{\text{rhc2}}_{2\text{–stp}}$, involves combination of model $\text{MC}_{1\text{–stp}}$ and the second type of non–response adjustment cells, RHC_2, such that

  First phase $:=$ RHC_2,

  Second phase $:=$ model $\text{MC}_{1\text{–stp}}$.

12. Numerical results

In this Section, we compare several point estimators (see (3), (5), (10), (12) and (18)) by using Norwegian LFS 2015, 2016 and 2017 data. Several models (see Sections 11.1-11.2) are considered for each estimator. Results are presented by Tables 1-17. The reference estimator (3) shall be denoted by $\widehat{Y}_H$. The PS and $\text{GREG}_{1\text{–stp}}$ are examples of the one–step GREG estimator (5). The $\text{MC}_{1\text{–stp}}$ is a one–step multiple model–calibration estimator defined by (10). The $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$, $\text{GREG}^{\text{rhc1}}_{2\text{–stp}}$ and $\text{GREG}^{\text{rhc2}}_{2\text{–stp}}$ are types of the two–step GREG estimator given by (12). The $\text{MC}^{\text{rhc1}}_{2\text{–stp}}$ and $\text{MC}^{\text{rhc2}}_{2\text{–stp}}$ are examples of the two-step multiple model–calibration estimator defined by (18). Standard errors of the point estimators are computed by using the JL variance estimator (see Section 7.1).

Calibration and non–response adjustment procedures are applied to monthly LFS data sets. Quarterly estimates are obtained by using either the equal– or weekly–weighted average of monthly estimates (see Section 9). In Tables 1-10, the equal–weighted averaging method is used. In Tables 15-17, the weekly–weighted averaging method is used. The equal– and weekly–weighted averaging methods are compared in Table 14. In Tables 1-17, the Q1, Q2, Q3 and Q4, which appear after each year as suffix, stand for first, second, third and fourth quarters of the associated year, respectively. The Q1 covers the first three months in a calender year: January–March. The other Q–s should be referred accordingly. The statistical software R Development Core Team (2014) is used for implementation.

12.1. The choice of a new estimation method for Norwegian LFS

In this Section, we compare several estimation methods in terms of point and standard error estimates for unemployment, employment and outside of labour force. Results are provided by Tables 1-3.

The reference estimator significantly underestimates unemployment and number of people outside of labour force, while it significantly overestimates employment in comparison to other estimators. This is due to the fact that the non–response mechanism is not the MCAR. This probably leads to downward bias for unemployment and outside of labour force, and upward bias in the estimates for employment. Downward bias indicates also that people who are unemployed or outside of labour force tend to respond less than those who are employed. Based on this interpretation, the proxy method to the previous estimation method used by SSB (see Section 4) does not reduce the non–response bias as much as the other methods do. This can be seen that overall, the $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$ provides the highest employment and lowest number of people outside of labour force. The PS estimator seems to provide more unstable estimates for employment and outside of labour force. We do not observe a significant difference for point estimates of total unemployment between different methods excluding the reference estimator.

We obtain a significant variance reduction for employment and outside of labour force by using auxiliary information in the estimation. This can be seen from Tables 2-3 as such that the standard errors of $\widehat{Y}_H$ for these statistics is almost double of those that are obtained with the other methods. However, the other methods provide higher standard errors for unemployment than the reference estimator provides (see Table 1). This may be explained by the fact that the relationship between y_{hij} (i.e. unemployed or not here) and $\boldsymbol{x}_{hij}$ or $\widehat{\boldsymbol{\eta}}_{hij}$, depending on which estimator we refer to, is not so strong as it should be (e.g. Wu and Sitter, 2001, p.189). Nevertheless, the reference estimator provides more biased estimates than the others which may eventually lead to higher mean square errors. Thus there is a question of bias–variance trade–off here.

The one–step multiple model–calibration estimator, $\text{MC}_{\text{1–stp}}$, provides smallest standard errors for all parameters among other methods, except the reference estimator, where it provides smallest standard errors for unemployment. Standard errors of total unemployment are reduced by at least 9% and at most 15% with the one–step multiple model–calibration estimator in comparison to the $\text{GREG}^{\text{ssb}}_{\text{2–stp}}$ (see Table 1). Reductions in standard errors of total employment and out of labour force are around 7-9$\%$ and 9-11$\%$, respectively (see Tables 2-3). Here, the comparison of one– and two–step estimators may not be quite fair. However, even the two–step multiple model–calibration estimator where the first phase is the same as the $\text{GREG}^{\text{ssb}}_{\text{2–stp}}$ (see Section 11.2) provides much lower standard errors for all cases.

We observe that the relationship between y_{hij} and $\widehat{\boldsymbol{\eta}}_{hij}$ is much stronger than the one between y_{hij} and $\boldsymbol{x}_{hij}$. This may be the reason of that the $\text{MC}_{\text{1–stp}}$ is more efficient than the $\text{GREG}_{\text{1–stp}}$ for all cases (see Tables 1-3). One may claim that this comparison may not be fair as the former includes many variables in the calibration model (see Section 11.1) that may increase the variance eventually. However, this has not been the case for the data used. For example, the post–stratified estimator, PS, includes only three variables leading to 44 calibration cells, which is much less than the one with the $\text{MC}_{\text{1–stp}}$. However, it provides much higher standard errors, even for total employment which is, in fact, highly correlated with the register based employment that is used in the PS model (see Section 11.1). Moreover, several models with less or more variables, have been investigated for one–step GREG estimator before deciding the last model. The model presented here has provided better results in terms of efficiency and unbiasedness among all others.

When it comes to the discussion of one–step or two–step estimation, we do not observe a particular superiority of two–step methods over those with those with one–step weighting. Two–step methods provide higher standard errors without providing further adjustment for non–response bias significantly. This may be explained by the fact that we have good explanatory variables in the calibration models (e.g. Nguyen and Zhang, 2016). Therefore, the one–step multiple model–calibration estimator seems to be the best among others for the data used.

12.2. Unemployment rate

In this Section, we investigate the performance of the $\text{MC}_{\text{1–stp}}$ for national level unemployment rate, which is computed as a ratio of total number of unemployed people (15-74) to the total number of people in labour force. Both numerator and denominator are estimated from the survey. Thus it is important to evaluate if the

Table 1: Point and standard error estimates of the number of people (15–74) unemployed in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Estimator	2015Q1	2015Q2	2015Q3	2015Q4	2016Q1	2016Q2	2016Q3	2016Q4	2017Q1
$\widehat{Y}_H$	96 (4.56)	102 (4.70)	105 (4.71)	93 (4.52)	110 (4.80)	109 (4.81)	112 (4.80)	96 (4.51)	102 (4.59)
$\mathrm{MC}_{1\text{–stp}}$	119 (4.61)	122 (4.79)	126 (4.84)	120 (4.96)	136 (5.01)	134 (5.00)	134 (4.84)	120 (4.67)	120 (4.55)
$\mathrm{MC}^{\mathrm{rhc1}}_{2\text{–stp}}$	119 (4.63)	122 (4.82)	126 (4.89)	120 (5.02)	136 (5.02)	134 (5.04)	134 (4.88)	120 (4.71)	120 (4.59)
$\mathrm{MC}^{\mathrm{rhc2}}_{2\text{–stp}}$	119 (4.68)	122 (4.85)	127 (4.97)	120 (5.11)	137 (5.20)	134 (5.10)	135 (4.99)	121 (4.86)	120 (4.65)
$\mathrm{GREG}_{1\text{–stp}}$	118 (4.71)	120 (4.94)	125 (5.09)	118 (5.20)	137 (5.34)	129 (5.21)	133 (5.16)	117 (4.92)	118 (4.76)
$\mathrm{GREG}^{\mathrm{rhc1}}_{2\text{–stp}}$	118 (4.77)	120 (4.99)	126 (5.17)	120 (5.32)	139 (5.50)	130 (5.30)	134 (5.23)	118 (5.05)	119 (4.82)
$\mathrm{GREG}^{\mathrm{rhc2}}_{2\text{–stp}}$	118 (4.74)	120 (4.98)	127 (5.20)	119 (5.28)	137 (5.40)	130 (5.30)	133 (5.21)	118 (5.00)	118 (4.79)
PS	118 (5.30)	121 (5.36)	129 (5.44)	116 (5.39)	137 (5.70)	133 (5.66)	138 (5.63)	121 (5.42)	123 (5.34)
$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	120 (5.44)	125 (5.51)	131 (5.55)	117 (5.45)	135 (5.55)	131 (5.51)	137 (5.57)	120 (5.30)	121 (5.19)

Table 2: Point and standard error estimates of the number of people (15–74) employed in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Estimator	2015Q1	2015Q2	2015Q3	2015Q4	2016Q1	2016Q2	2016Q3	2016Q4	2017Q1
$\widehat{Y}_H$	2 724 (13.77)	2 750 (13.69)	2 733 (13.91)	2 725 (14.07)	2 716 (13.96)	2 722 (14.02)	2 734 (13.99)	2 704 (14.18)	2 692 (14.08)
$\text{MC}_{1\text{–stp}}$	2 611 (7.14)	2 649 (7.25)	2 647 (7.63)	2 631 (7.26)	2 624 (7.03)	2 640 (7.15)	2 654 (7.33)	2 628 (6.97)	2 621 (6.79)
$\text{MC}^{\text{rhc1}}_{2\text{–stp}}$	2 611 (7.22)	2 649 (7.34)	2 647 (7.72)	2 631 (7.35)	2 624 (7.11)	2 640 (7.26)	2 654 (7.41)	2 628 (7.10)	2 622 (6.89)
$\text{MC}^{\text{rhc2}}_{2\text{–stp}}$	2 611 (7.20)	2 650 (7.35)	2 647 (7.72)	2 631 (7.42)	2 625 (7.19)	2 640 (7.29)	2 654 (7.45)	2 628 (7.15)	2 622 (6.91)
$\text{GREG}_{1\text{–stp}}$	2 615 (7.39)	2 654 (7.59)	2 647 (7.96)	2 636 (7.67)	2 623 (7.37)	2 640 (7.51)	2 657 (7.80)	2 626 (7.35)	2 621 (7.19)
$\text{GREG}^{\text{rhc1}}_{2\text{–stp}}$	2 615 (7.44)	2 654 (7.66)	2 647 (7.99)	2 635 (7.72)	2 623 (7.47)	2 638 (7.59)	2 656 (7.85)	2 625 (7.40)	2 620 (7.24)
$\text{GREG}^{\text{rhc2}}_{2\text{–stp}}$	2 615 (7.45)	2 655 (7.65)	2 646 (8.00)	2 635 (7.71)	2 622 (7.43)	2 639 (7.57)	2 657 (7.84)	2 627 (7.41)	2 621 (7.24)
PS	2 619 (7.57)	2 660 (7.81)	2 652 (8.21)	2 644 (7.86)	2 623 (7.63)	2 638 (7.77)	2 650 (8.04)	2 623 (7.68)	2 617 (7.49)
$\text{GREG}^{\text{ssb}}_{2\text{–stp}}$	2 621 (7.64)	2 661 (7.89)	2 652 (8.26)	2 645 (7.86)	2 632 (7.48)	2 648 (7.70)	2 662 (8.04)	2 633 (7.54)	2 623 (7.33)

Table 3: Point and standard error estimates of the number of people (15–74) outside of labour force in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Estimator	2015Q1	2015Q2	2015Q3	2015Q4	2016Q1	2016Q2	2016Q3	2016Q4	2017Q1
$\widehat{Y}_H$	1 061 (13.40)	1 038 (13.36)	1 063 (13.53)	1 095 (13.76)	1 095 (13.65)	1 098 (13.73)	1 093 (13.69)	1 149 (13.91)	1 161 (13.84)
$\text{MC}_{1\text{–stp}}$	1 150 (7.59)	1 119 (7.65)	1 127 (7.82)	1 163 (7.58)	1 161 (7.54)	1 156 (7.54)	1 151 (7.73)	1 201 (7.48)	1 214 (7.25)
$\text{MC}^{\text{rhc1}}_{2\text{–stp}}$	1 150 (7.67)	1 119 (7.73)	1 127 (7.90)	1 163 (7.67)	1 161 (7.61)	1 156 (7.63)	1 151 (7.82)	1 201 (7.59)	1 213 (7.36)
$\text{MC}^{\text{rhc2}}_{2\text{–stp}}$	1 150 (7.65)	1 118 (7.71)	1 126 (7.91)	1 162 (7.76)	1 160 (7.68)	1 155 (7.65)	1 151 (7.86)	1 200 (7.69)	1 213 (7.38)
$\text{GREG}_{1\text{–stp}}$	1 148 (7.79)	1 116 (7.96)	1 128 (8.15)	1 160 (8.00)	1 163 (7.88)	1 160 (7.92)	1 150 (8.20)	1 206 (7.89)	1 216 (7.65)
$\text{GREG}^{\text{rhc1}}_{2\text{–stp}}$	1 147 (7.86)	1 116 (8.03)	1 127 (8.19)	1 159 (8.11)	1 161 (7.98)	1 161 (8.01)	1 150 (8.27)	1 206 (7.99)	1 217 (7.74)
$\text{GREG}^{\text{rhc2}}_{2\text{–stp}}$	1 147 (7.87)	1 115 (8.02)	1 127 (8.22)	1 160 (8.08)	1 163 (7.94)	1 160 (7.99)	1 149 (8.27)	1 204 (7.97)	1 216 (7.72)
PS	1 143 (8.25)	1 109 (8.32)	1 119 (8.56)	1 154 (8.39)	1 162 (8.47)	1 158 (8.40)	1 152 (8.70)	1 205 (8.36)	1 215 (8.17)
$\text{GREG}^{\text{ssb}}_{2\text{–stp}}$	1 140 (8.36)	1 105 (8.47)	1 118 (8.69)	1 152 (8.41)	1 155 (8.36)	1 151 (8.31)	1 139 (8.66)	1 196 (8.24)	1 211 (8.04)

$\text{MC}_{1\text{–stp}}$ provides better efficiency for such a parameter than the $\text{GREG}_{1\text{–stp}}$ and the $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$. Point and standard error estimates are presented in Table 4. Coefficient of variation estimates are given by Table 5. We eliminate the two–step multiple model–calibration and the GREG estimators in this comparison since they provide higher variance without reducing non–response bias more (see Tables 1-3). The PS estimator is not presented in Tables 4- 5 either as it provides the highest variances overall (see Tables 1-3).

The $\text{MC}_{1\text{–stp}}$ and the GREG estimators reduce non–response bias significantly when we compare them with the reference estimator. There is no significant difference between the methods using auxiliary information in terms of point estimates. The multiple model–calibration estimator provides smallest standard errors and CVs among others.

Table 4: Point and standard error estimates of unemployment rates in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in percentage–unit points (%). Standard error estimates are presented within parentheses.

Period	Estimator			
	$\widehat{Y}_H$	$\text{MC}_{1\text{–stp}}$	$\text{GREG}_{1\text{–stp}}$	$\text{GREG}^{\text{ssb}}_{2\text{–stp}}$
2015Q1	3.4 (0.161)	4.4 (0.164)	4.3 (0.168)	4.4 (0.193)
2015Q2	3.6 (0.164)	4.4 (0.168)	4.3 (0.174)	4.5 (0.192)
2015Q3	3.7 (0.165)	4.6 (0.170)	4.5 (0.179)	4.7 (0.194)
2015Q4	3.3 (0.160)	4.4 (0.176)	4.3 (0.184)	4.2 (0.192)
2016Q1	3.9 (0.169)	4.9 (0.176)	4.9 (0.188)	4.9 (0.194)
2016Q2	3.9 (0.169)	4.8 (0.175)	4.7 (0.183)	4.7 (0.193)
2016Q3	3.9 (0.168)	4.8 (0.169)	4.8 (0.180)	4.9 (0.193)
2016Q4	3.4 (0.160)	4.4 (0.165)	4.3 (0.174)	4.3 (0.187)
2017Q1	3.6 (0.163)	4.4 (0.162)	4.3 (0.169)	4.4 (0.184)

Table 5: Coefficient of variation (CV) estimates unemployment rates in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Estimates are given in percentage–unit points (%).

Period	CV (%)			
	$\widehat{Y}_H$	$\text{MC}_{1\text{–stp}}$	$\text{GREG}_{1\text{–stp}}$	$\text{GREG}^{\text{ssb}}_{2\text{–stp}}$
2015Q1	4.73	3.75	3.90	4.40
2015Q2	4.58	3.82	4.02	4.29
2015Q3	4.46	3.74	3.96	4.14
2015Q4	4.82	4.04	4.28	4.54
2016Q1	4.33	3.56	3.80	3.96
2016Q2	4.38	3.64	3.92	4.09
2016Q3	4.25	3.50	3.78	3.93
2016Q4	4.66	3.79	4.10	4.30
2017Q1	4.49	3.70	3.94	4.17

12.3. Estimation over domains

In this Section, the performance of the one–step multiple model–calibration estimator $\text{MC}_{1\text{–stp}}$ is investigated for domain estimation by comparing with the $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$. The one–step GREG estimator $\text{GREG}_{1\text{–stp}}$ is not presented here as the multiple model–calibration estimator provides better estimates than the $\text{GREG}_{1\text{–stp}}$ in terms of efficiency for totals as well as ratios (see Tables 1-5). Tables 6-11 show that the $\text{MC}_{1\text{–stp}}$ provides more efficient estimates for all domains than the $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$ does. We have significant variance reduction for employment over domains when auxiliary information is used in the estimation process. Although register based employment status (i.e. employed or not), which is highly correlated with employment from LFS, is used with the $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$, standard errors for employment over domains with the $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$ are not so small as those obtained from the $\text{MC}_{1\text{–stp}}$. This may be explained by the fact that interaction terms between domains and register based employment status are not used in the calibration model for $\text{GREG}^{\text{ssb}}_{2\text{–stp}}$. In other words, calibration is not carried out within these domains. This may lead to higher variances over domains especially when domain–level characteristics are different from overall in terms of the statistics of interest.

12.4. Equal– versus weekly–weighted average for quarterly estimates

In Table 14, quarterly estimates obtained from the one–step multiple model–calibration estimator under the equal– and weekly–weighted averaging methods are presented. There are no significant differences in point and standard error estimates for employment, unemployment and outside of labour force between the two weighting methods. However, it may be more sensible to use the weekly–weighted averaging method as it may provide smoother quarterly weights by reducing the effect of a five–week month while increasing the effects of four-week months on quarterly estimates. Apart from this, other type of variables or estimates over certain domains may be more sensible to calender weeks that the use of an equal–weighted average may lead to over– or under–estimation of variables of interest (e.g. Hamre, 2017).

Table 6: Point and standard error estimates of unemployment by gender in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	Female			Male		
	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$
2015Q1	43 (2.98)	51 (3.06)	53 (3.64)	53 (3.34)	68 (3.42)	67 (3.99)
2015Q2	45 (3.13)	53 (3.19)	52 (3.55)	57 (3.44)	69 (3.46)	72 (4.18)
2015Q3	46 (3.13)	53 (3.13)	55 (3.61)	59 (3.48)	74 (3.69)	76 (4.25)
2015Q4	45 (3.07)	55 (3.28)	54 (3.60)	49 (3.17)	64 (3.58)	63 (3.98)
2016Q1	45 (3.04)	53 (3.14)	53 (3.47)	65 (3.66)	83 (3.81)	82 (4.36)
2016Q2	42 (2.93)	48 (2.94)	47 (3.25)	68 (3.79)	86 (3.98)	83 (4.47)
2016Q3	51 (3.25)	58 (3.27)	60 (3.68)	61 (3.55)	76 (3.52)	77 (4.25)
2016Q4	41 (2.91)	48 (3.01)	48 (3.31)	56 (3.45)	72 (3.58)	72 (4.20)
2017Q1	43 (2.96)	48 (2.95)	49 (3.29)	59 (3.43)	72 (3.39)	72 (3.99)

Table 7: Point and standard error estimates of employment by gender in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	Female			Male		
	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$
2015Q1	1 275 (11.58)	1 229 (4.96)	1 239 (7.00)	1 450 (11.92)	1 382 (5.02)	1 382 (6.97)
2015Q2	1 288 (11.61)	1 254 (5.02)	1 261 (7.04)	1 462 (11.84)	1396 (5.21)	1 400 (7.06)
2015Q3	1 290 (11.68)	1 248 (5.32)	1 254 (7.23)	1 443 (11.94)	1 399 (5.42)	1 398 (7.20)
2015Q4	1 289 (11.76)	1 236 (4.96)	1 250 (7.05)	1 436 (12.00)	1 395 (5.18)	1 394 (7.07)
2016Q1	1 287 (11.64)	1 243 (4.88)	1 253 (6.98)	1 430 (11.88)	1 381 (5.02)	1 379 (7.00)
2016Q2	1 292 (11.75)	1 249 (4.94)	1 257 (7.04)	1 429 (11.93)	1 391 (5.10)	1 390 (7.07)
2016Q3	1 310 (11.75)	1 255 (5.11)	1 268 (7.14)	1 424 (11.89)	1 398 (5.15)	1 394 (7.19)
2016Q4	1 297 (11.72)	1 247 (4.76)	1 255 (7.00)	1 407 (12.04)	1 381 (5.04)	1 378 (7.10)
2017Q1	1 293 (11.54)	1 240 (4.67)	1 246 (6.88)	1 400 (11.90)	1 381 (4.94)	1 377 (6.99)

Table 8: Point and standard error estimates of unemployment by age groups in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	15–24			25–54			55-74		
	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$
2015Q1	33 (2.60)	37 (2.63)	35 (2.70)	57 (3.51)	73 (3.52)	78 (4.54)	7 (1.23)	9 (1.37)	7 (1.32)
2015Q2	42 (3.01)	46 (2.91)	45 (3.07)	53 (3.37)	69 (3.55)	73 (4.38)	6 (1.15)	7 (1.15)	7 (1.21)
2015Q3	38 (2.87)	43 (2.90)	41 (2.98)	58 (3.54)	75 (3.69)	81 (4.57)	9 (1.38)	9 (1.17)	9 (1.39)
2015Q4	29 (2.48)	35 (2.59)	32 (2.63)	55 (3.47)	76 (3.98)	76 (4.52)	9 (1.38)	9 (1.29)	9 (1.49)
2016Q1	38 (2.79)	43 (2.86)	41 (2.91)	61 (3.56)	82 (3.83)	82 (4.49)	11 (1.52)	11 (1.29)	12 (1.54)
2016Q2	43 (3.08)	48 (3.05)	46 (3.12)	58 (3.49)	77 (3.75)	77 (4.37)	8 (1.28)	9 (1.20)	9 (1.35)
2016Q3	36 (2.76)	41 (2.76)	40 (2.91)	67 (3.73)	85 (3.77)	88 (4.60)	9 (1.34)	9 (1.21)	9 (1.43)
2016Q4	30 (2.55)	35 (2.68)	33 (2.68)	56 (3.44)	74 (3.59)	76 (4.36)	10 (1.43)	11 (1.23)	11 (1.52)
2017Q1	34 (2.61)	39 (2.64)	37 (2.70)	58 (3.48)	72 (3.48)	74 (4.19)	10 (1.46)	10 (1.24)	10 (1.53)

Table 9: Point and standard error estimates of employment by age groups in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	15–24			25–54			55-74		
	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$
2015Q1	331 (8.19)	330 (3.92)	331 (5.23)	1 790 (14.88)	1 753 (4.83)	1 755 (6.87)	604 (11.76)	528 (3.52)	535 (5.60)
2015Q2	346 (8.29)	346 (3.99)	344 (5.17)	1 791 (14.84)	1 767 (5.01)	1 775 (6.96)	614 (11.76)	536 (3.43)	542 (5.58)
2015Q3	349 (8.49)	346 (4.57)	346 (5.43)	1 770 (14.99)	1 768 (5.18)	1 767 (7.03)	614 (11.80)	533 (3.27)	538 (5.51)
2015Q4	339 (8.34)	328 (3.91)	334 (5.17)	1 780 (15.09)	1 772 (5.22)	1 781 (7.11)	605 (11.77)	531 (3.20)	531 (5.51)
2016Q1	322 (8.06)	325 (3.81)	324 (5.23)	1 796 (15.00)	1 771 (4.90)	1 777 (6.76)	598 (11.53)	528 (3.25)	531 (5.46)
2016Q2	320 (7.99)	325 (3.97)	327 (5.16)	1 792 (15.05)	1 778 (4.98)	1 782 (6.83)	611 (11.65)	536 (3.18)	538 (5.39)
2016Q3	335 (8.15)	339 (4.29)	341 (5.35)	1 795 (14.91)	1776 (4.94)	1 781 (6.92)	604 (11.47)	538 (3.25)	541 (5.49)
2016Q4	316 (7.94)	322 (3.94)	322 (5.21)	1 791 (15.07)	1 775 (4.74)	1 778 (6.78)	596 (11.50)	531 (3.2)	533 (5.50)
2017Q1	305 (7.68)	310 (3.65)	311 (5.03)	1 801 (14.99)	1 778 (4.76)	1 781 (6.75)	587 (11.34)	533 (3.22)	532 (5.54)

Table 10: Point and standard error estimates of employment among females by age groups in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	female, 15-24			female, 25-74		
	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\mathrm{MC}_{1\text{–stp}}$	$\mathrm{GREG}^{\mathrm{ssb}}_{2\text{–stp}}$
2015Q1	164 (5.76)	164 (2.79)	169 (4.50)	1 111 (10.57)	1 065 (4.12)	1 070 (7.06)
2015Q2	171 (5.89)	176 (2.84)	177 (4.61)	1 117 (10.58)	1 078 (4.15)	1 084 (7.09)
2015Q3	169 (5.89)	172 (3.27)	173 (4.67)	1 121 (10.76)	1 077 (4.22)	1 081 (7.17)
2015Q4	160 (5.76)	1 60 (2.73)	163 (4.46)	1 129 (10.83)	1 076 (4.17)	1 088 (7.15)
2016Q1	160 (5.73)	163 (2.67)	165 (4.54)	1 127 (10.68)	1 080 (4.05)	1 088 (7.06)
2016Q2	159 (5.66)	162 (2.74)	164 (4.51)	1 134 (10.77)	1 087 (4.11)	1 093 (7.07)
2016Q3	164 (5.71)	169 (2.99)	169 (4.59)	1 146 (10.73)	1 087 (4.12)	1 099 (7.12)
2016Q4	152 (5.58)	159 (2.78)	158 (4.48)	1 145 (10.74)	1 088 (3.83)	1 097 (7.04)
2017Q1	147 (5.37)	151 (2.60)	151 (4.34)	1 146 (10.68)	1 089 (3.88)	1 096 (7.01)

Table 11: Point and standard error estimates of employment among males by age groups in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	male, 15-24			male, 25-74		
	$\widehat{Y}_H$	$\text{MC}_{1\text{–stp}}$	$\text{GREG}^{\text{ssb}}_{2\text{–stp}}$	$\widehat{Y}_H$	$\text{MC}_{1\text{–stp}}$	$\text{GREG}^{\text{ssb}}_{2\text{–stp}}$
2015Q1	168 (5.93)	166 (2.66)	162 (4.39)	1 282 (11.41)	1 216 (4.22)	1 220 (7.04)
2015Q2	175 (5.95)	170 (2.79)	167 (4.37)	1 288 (11.39)	1 226 (4.39)	1 233 (7.18)
2015Q3	180 (6.17)	174 (3.16)	174 (4.61)	1 263 (11.37)	1 224 (4.39)	1 224 (7.19)
2015Q4	179 (6.13)	168 (2.76)	171 (4.49)	1 257 (11.42)	1 227 (4.39)	1 223 (7.19)
2016Q1	162 (5.76)	162 (2.69)	159 (4.43)	1 267 (11.36)	1 219 (4.22)	1 219 (7.12)
2016Q2	161 (5.73)	163 (2.84)	163 (4.42)	1 269 (11.45)	1 227 (4.19)	1 227 (7.16)
2016Q3	171 (5.84)	171 (2.99)	172 (4.49)	1 253 (11.32)	1 228 (4.16)	1 223 (7.20)
2016Q4	164 (5.80)	163 (2.77)	165 (4.50)	1 243 (11.49)	1 219 (4.17)	1 213 (7.13)
2017Q1	158 (5.63)	160 (2.59)	160 (4.38)	1 241 (11.39)	1 222 (4.19)	1 217 (7.09)

Table 12: Point and standard error estimates of unemployment among females by age groups in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	female, 15-24			female, 25-74		
	$\widehat{Y}_H$	MC_{1-stp}	$GREG^{ssb}_{2-stp}$	$\widehat{Y}_H$	MC_{1-stp}	$GREG^{ssb}_{2-stp}$
2015Q1	13 (1.68)	16 (1.78)	15 (1.85)	29 (2.48)	36 (2.49)	38 (3.16)
2015Q2	19 (2.02)	20 (2.08)	20 (2.14)	26 (2.35)	32 (2.38)	32 (2.81)
2015Q3	16 (1.86)	18 (1.90)	17 (1.99)	30 (2.52)	35 (2.48)	37 (3.03)
2015Q4	14 (1.69)	17 (1.84)	15 (1.86)	31 (2.57)	39 (2.73)	39 (3.12)
2016Q1	16 (1.80)	18 (1.95)	18 (1.93)	28 (2.42)	35 (2.44)	35 (2.90)
2016Q2	17 (1.87)	19 (1.83)	18 (1.97)	25 (2.26)	29 (2.31)	29 (2.61)
2016Q3	17 (1.87)	19 (1.86)	19 (2.00)	34 (2.64)	39 (2.66)	41 (3.08)
2016Q4	13 (1.62)	14 (1.73)	14 (1.74)	28 (2.41)	34 (2.44)	34 (2.81)
2017Q1	14 (1.69)	16 (1.71)	15 (1.77)	29 (2.40)	32 (2.37)	34 (2.74)

Table 13: Point and standard error estimates of unemployment among males by age groups in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). The equal–weighted averaging method is used for quarterly estimates (see Section 9). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	male, 15-24			male, 25-74		
	$\widehat{Y}_H$	MC_{1-stp}	$GREG^{ssb}_{2-stp}$	$\widehat{Y}_H$	MC_{1-stp}	$GREG^{ssb}_{2-stp}$
2015Q1	19 (1.98)	22 (1.91)	20 (1.97)	34 (2.71)	46 (2.85)	47 (3.50)
2015Q2	23 (2.18)	25 (1.99)	25 (2.21)	34 (2.65)	44 (2.82)	48 (3.56)
2015Q3	22 (2.15)	25 (2.18)	24 (2.23)	37 (2.77)	49 (2.96)	52 (3.67)
2015Q4	16 (1.77)	18 (1.78)	16 (1.84)	33 (2.64)	46 (3.12)	46 (3.56)
2016Q1	22 (2.08)	25 (2.05)	24 (2.19)	43 (3.02)	59 (3.23)	59 (3.82)
2016Q2	27 (2.39)	29 (2.38)	28 (2.41)	41 (2.92)	56 (3.16)	56 (3.77)
2016Q3	19 (2.04)	22 (2.02)	21 (2.15)	42 (2.90)	55 (2.90)	56 (3.70)
2016Q4	18 (1.97)	21 (2.03)	19 (2.06)	38 (2.80)	50 (2.90)	53 (3.67)
2017Q1	20 (1.97)	23 (1.99)	22 (2.06)	39 (2.82)	49 (2.75)	51 (3.46)

Table 14: Point and standard error estimates of labour market status in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1) under the equal– and weekly–weighted averaging methods (see Section 9). Multiple model–calibration estimator is used (see (10) and model MC_{1-stp} in Section 11.1). Point estimates are given in 1 000 units. Standard error estimates are presented within parentheses.

Period	unemployed		employed		outside of labour force	
	equal	weekly	equal	weekly	equal	weekly
2015Q1	119 (4.61)	119 (4.57)	2 611 (7.14)	2 610 (7.10)	1 150 (7.59)	1 152 (7.53)
2015Q2	122 (4.79)	122 (4.78)	2 649 (7.25)	2 647 (7.21)	1 119 (7.65)	1 122 (7.62)
2015Q3	126 (4.84)	127 (4.84)	2 647 (7.63)	2 646 (7.59)	1 127 (7.82)	1 128 (7.78)
2015Q4	120 (4.96)	120 (4.94)	2 631 (7.26)	2 632 (7.31)	1 163 (7.58)	1 162 (7.59)
2016Q1	136 (5.01)	136 (4.97)	2 624 (7.03)	2 625 (7.07)	1 161 (7.54)	1 161 (7.54)
2016Q2	134 (5.00)	134 (4.97)	2 640 (7.15)	2 642 (7.20)	1 156 (7.54)	1 154 (7.55)
2016Q3	134 (4.84)	134 (4.81)	2 654 (7.33)	2 652 (7.33)	1 151 (7.73)	1 154 (7.73)
2016Q4	120 (4.67)	119 (4.64)	2 628 (6.97)	2 626 (6.99)	1 201 (7.48)	1 204 (7.49)
2017Q1	120 (4.55)	121 (4.56)	2 621 (6.79)	2 621 (6.81)	1 214 (7.25)	1 213 (7.26)

12.5. EU precision requirements

In this Section, we investigate if the precision requirements for national level employment and unemployment rates, and regional level (NUTS II) unemployment rates for EU LFSs (EC, 2014) are fulfilled for the Norwegian LFS where the one–step multiple model–calibration estimator accompanied with the weighted average method is used for quarterly estimates. The condition for national level employment and unemployment rates is defined by

$$\widehat{se}(\,\widehat{p}\,) \leq \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{7800\sqrt{M_{15-74}}-4500}}, \tag{40}$$

where $\widehat{p}$ denotes estimated quarterly ratio of either employment or unemployment to the total population aged 15-74 and $\widehat{se}(\,\widehat{p}\,)$ is an standard error estimate of $\widehat{p}$. Here, M_{15-74} is the total population aged 15-74 expressed in millions (EC, 2014, p.13, 23). For regional level unemployment rates, we have

$$\widehat{se}(\,\widehat{p}_{\Phi;u}\,) \leq \sqrt{\frac{\widehat{p}_{\Phi;u}\,(1-\widehat{p}_{\Phi;u})}{A}}, \tag{41}$$

with

$$A = \begin{cases} 1300 & \text{if } M_{\Phi;15-74} \geq 0.300 \text{ million} \\ 1300(0.300)^{-1} & \text{if } M_{\Phi;15-74} < 0.300 \text{ million} \end{cases},$$

where $\widehat{p}_{\Phi;u}$ is estimated ratio of unemployment to the total population aged 15-74 in region Φ, which is denoted by $M_{\Phi;15-74}$, and $\widehat{se}(\,\widehat{p}_{\Phi;u}\,)$ is an standard error estimate of $\widehat{p}_{\Phi;u}$ (EC, 2014, p.18). Here, $M_{\Phi;15-74}$ is expressed in millions. It should be noticed that the criteria (40)-(41) involves the use of total population aged 15-74 at national or regional level, but not estimated population in the labour force, for computation of unemployment rates.

Precision requirements are fulfilled for employment and unemployment rates for all quarters (i.e. from first quarter of 2015 to first quarter of 2017). However, in Table 15, only results for the first quarter of 2015 (2015Q1) are presented as the results for other quarters were quite similar to those regarding 2015Q1.

12.6. Effect of clustering

Effect of clustering on variance estimates are compared for the reference estimator and the one–step multiple model–calibration estimator in Tables 16-17. Variance estimator without clustering does not involve aggregation of unit–level linearised variables to cluster–level. Here, we ignore the household cluster sampling of the individuals and treat them as if they were directly selected. This may underestimate the variance. However, downward bias may be negligible if good auxiliary variables are used in the estimation (e.g. Hagesæther and Zhang, 2009). Therefore, we compare two cases: i). reference estimator which do not use auxiliary information, except M, ii). one–step multiple model–calibration estimator that uses auxiliary information.

Variances regarding the reference estimator for unemployment are under–estimated mostly by 4-5$\%$ with the variance estimator without clustering. Downward bias is a little bit lower than those values, that is around 2-3$\%$ for most of the cases, with the one–step multiple model–calibration estimator (see Table 16). We have significant differences between the cluster effects obtained from the two estimators for employment (see Table 17). Variances for the reference estimator are underestimated

Table 15: Check for the precision requirements of EU for national level employment and unemployment rates, and NUTS II level unemployment rates in Norway in the first quarter of 2015. Multiple model–calibration estimator is used (see (10) and model $\text{MC}_{\text{1–stp}}$ in Section 11.1). Weighted average method is used for quarterly estimates (see Section 9).

Domain	M^1_{15-74}	$\widehat{p}^2$	$\widehat{se}(\widehat{p})$	criteria[3]	satisfy[4]
Overall	3.881	67.3	0.18	0.45	yes
Overall	3.881	3.1	0.12	0.17	yes
NO01	0.939	3.3	0.26	0.50	yes
NO02	0.288	3.3	0.42	0.51	yes
NO03	0.733	3.7	0.30	0.52	yes
NO04	0.566	2.9	0.29	0.46	yes
NO05	0.657	2.7	0.26	0.45	yes
NO06	0.335	2.6	0.36	0.44	yes
NO07	0.363	2.4	0.34	0.43	yes

[1] Population aged 15-74 expressed in millions.

[2] Employment rate for the first row and unemployment rates for the others (see Section 12.5).

[3] Given by the right hand sides of (40) and (41) for the first two rows and the others, respectively.

[4] Satisfy is *yes* if $\widehat{se}(\widehat{p}) \leq$ criteria.

Table 16: Variance estimates with and without clustering (i.e. $v_{\text{cl}}(\cdot)$ and $v(\cdot)$, respectively) for unemployment in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). Weighted average is used for quarterly totals (see Section 9). Point estimates (second and sixth columns) are given in 1 000 units.

Period	$\widehat{Y}_H$				$\text{MC}_{\text{1–stp}}$			
	$\widehat{Y}_H$	$v_{\text{cl}}(\widehat{Y}_H)$	$v(\widehat{Y}_H)$	R^1_H	$\widehat{Y}_{mmc}$	$v_{\text{cl}}(\widehat{Y}_{mmc})$	$v(\widehat{Y}_{mmc})$	R^2_{mmc}
2015Q1	95	20.3	19.3	0.95	119	20.9	20.5	0.98
2015Q2	101	21.8	20.4	0.94	122	22.8	21.7	0.95
2015Q3	105	22.2	21.3	0.96	127	23.4	23.1	0.99
2015Q4	94	20.4	19.3	0.95	120	24.4	23.6	0.97
2016Q1	110	23.0	21.9	0.95	136	24.7	23.9	0.97
2016Q2	110	22.9	21.9	0.96	134	24.7	24.0	0.97
2016Q3	112	22.6	21.9	0.97	134	23.1	22.7	0.98
2016Q4	96	20.0	19.2	0.96	119	21.5	21.1	0.98
2017Q1	102	21.1	20.2	0.96	121	20.8	20.4	0.98

[1] Given by $v(\widehat{Y}_H)/v_{\text{cl}}(\widehat{Y}_H)$.

[2] Given by $v(\widehat{Y}_{mmc})/v_{\text{cl}}(\widehat{Y}_{mmc})$.

by around 12% when clustering is taken into account. On the other hand, under–estimation of variances is mostly around 2-3$\%$ for the multiple model–calibration estimator. Therefore, effect of clustering on variance estimates may be negligible when we have explanatory variables that strongly explain the variables of interest. Otherwise, one should be cautious when applying a unit–level variance estimator under a cluster sampling design as clustering may have a significant effect on variance estimates.

Table 17: Variance estimates with and without clustering (i.e. $v_{\text{cl}}(\cdot)$ and $v(\cdot)$, respectively) for employment in Norway over 2015 Quarter 1 (2015Q1) – 2017 Quarter 1 (2017Q1). Weighted average is used for quarterly totals (see Section 9). Point estimates (second and sixth columns) are given in 1 000 units.

Period	$\widehat{Y}_H$				$\text{MC}_{1\text{–stp}}$			
	$\widehat{Y}_H$	$v_{\text{cl}}(\widehat{Y}_H)$	$v(\widehat{Y}_H)$	R_H^1	$\widehat{Y}_{mmc}$	$v_{\text{cl}}(\widehat{Y}_{mmc})$	$v(\widehat{Y}_{mmc})$	R_{mmc}^2
2015Q1	2 720	188.4	167.2	0.89	2 610	50.4	49.1	0.97
2015Q2	2 745	186.7	165.9	0.89	2 647	52.0	51.8	1.00
2015Q3	2 730	192.8	169.3	0.88	2 646	57.6	56.3	0.98
2015Q4	2 726	195.9	173.1	0.88	2 632	53.4	52.1	0.98
2016Q1	2 717	192.7	170.2	0.88	2 625	50.0	48.8	0.98
2016Q2	2 724	194.0	170.8	0.88	2 642	51.8	50.3	0.97
2016Q3	2 733	194.5	168.1	0.86	2 652	53.8	52.0	0.97
2016Q4	2 701	199.9	173.5	0.87	2 626	48.9	47.2	0.97
2017Q1	2 691	196.6	171.7	0.87	2 621	46.4	46.2	1.00

[1] Given by $v(\widehat{Y}_H)/v_{\text{cl}}(\widehat{Y}_H)$.
[2] Given by $v(\widehat{Y}_{mmc})/v_{\text{cl}}(\widehat{Y}_{mmc})$.

13. Conclusion

We have compared several estimators in terms of point and standard error estimates by using the Norwegian LFS data of nine quarters from 2015 to 2017. We have observed that all calibration estimators reduces the non–response bias to a certain extent in comparison to the reference estimator (e.g. see Tables 1-3). They provide more efficient estimates for employment and outside of labour force than those obtained from the reference estimator.

We have observed that the two–step estimators have provided higher standard errors without improvement of non–response bias. Therefore, a one–step calibration estimator is recommended over a two–step estimator as long as the calibration model includes variables which are strongly correlated with the variables of interest. Overall, the one–step multiple model–calibration estimator (see Section 11.1) has provided lowest standard errors for unemployment, employment and outside of labour force compared to the other calibration estimators used in the application (see Section 12). Moreover, it satisfies the EU precision requirements (see Section 12.5). Therefore, we propose using this estimator, where the weekly–weighted averaging method is utilised for quarterly estimates (see Section 9), in the new estimation methodology for the Norwegian LFS.

The calibration model which shall be used in the production is slightly different from the one given in Section 11.1. In order to obtain a better precision for labour market statistics over immigrant groups, a cross–classification of register based employment status with country background has been added into the model. Therefore, the final model used in the new estimation methodology is given by

$$\mathbf{MC}_{\mathbf{1\text{–}stp}}^{\mathbf{final}} := \sim age\,(12) \times gender\,(2) + age\,(8) + [age\,(3) \times gender\,(2) \times \{\widehat{p}_e + \widehat{p}_u + \widehat{p}_o\}] + [region\,(7) \times \{\{\widehat{p}_e + \widehat{p}_u + \widehat{p}_o + age\,(3) + gender\,(2)\}\} + regemp\,(4) \times gender\,(2) + \mathbf{regemp\,(2)} \times \mathbf{country\,(3)}\cdot$$

There has also been a slight change in the multinomial logistic model presented in Section 11.1. The four-factor education variable shall be replaced by a three–factor education variable by grouping those with primary school education and the others (see Table B.2). Therefore, the final multinomial logistic model, which is used to predict p_e, p_u and p_o, is given by

$$\text{Mnomlog} \sim gender\,(2) * [age\,(13) + regemp\,(7) + \{age\,(11) * regemp\,(2)\} + \{regemp\,(2) * marstat\,(2) * age\,(2)\}] + \mathbf{education}\,(\mathbf{3}) + county\,(19) + familysize\,(3) + tiltak\,(3) + country\,(3)\cdot$$

Changes in labour market statistics over time are also the main interest of the LFSs. Since these statistics are estimated, variances of change estimates should also be provided alongside with point estimates in order to judge whether or not observed changes are statistically significant. Variance estimation of changes in rotating panel surveys have been discussed by several others (e.g. Holmes and Skinner, 2000; Berger, 2004; Qualité and Tillé, 2008; Qualité, 2009; Oguz-Alper and Berger, 2015; Berger and Priam, 2016). In the Norwegian LFS, the variance estimator provided by Hamre and Heldal (2013, p.11-12) is used for the estimation of variances of changes. This variance estimator involves covariance estimation that is based on the overlapping sample between the periods of interest. In this way, the covariance between two estimated totals at different time periods may be significantly overestimated that may lead to a large negative bias in the variance estimator of change if there is a strong correlation. In that case, variance estimators that always provide positive variance estimates may be preferred (e.g. Qualité, 2009; Berger, 2004). In this respect, we suggest investigating a variance estimator for change estimates for the Norwegian LFS as a future research work.

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.

Berger, Y. G. (2004), "Variance estimation for measures of change in probability sampling," *Canadian Journal of Statistics*, 32(4), 451–467.

Berger, Y. G., and Priam, R. (2016), "A Simple Variance Estimator of Change for Rotating Repeated Surveys: an Application to the EU-SILC Household Surveys," *Journal of the Royal Statistical Society, Series A*, p. 22pp. Available at: DOI: `http://dx.doi.org/10.1111/rssa.12116`. [Online; accessed 09 June 2015].

Binder, D. A. (1996), "Linearization Methods for Single Phase and Two-phase Samples: A Cookbook Approach," *Survey Methodology*, 22(1), 17–22.

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. (1999), "Variance estimation for complex statistics and estimators: linearization and residual techniques," *Survey Methodology*, 25, 193–203.

Deville, J. C., and Särndal, C. E. (1992), "Calibration Estimators in Survey Sampling," *Journal of the American Statistical Association*, 87(418), 376–382.

EC (1998), "Council Regulation (EC) No 577/98 of 9 March 1998,", `http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31998R0577&from=EN`. [Online; accessed 30 Januar 2018].

EC (2014), "Final report of the task force on EU LFS precision requirements,", Commission Regulation (EC), Doc.: Eurostat/F3/LAMAS/06/14. [Online; accessed 04 January 2017]. `https://circabc.europa.eu/webdav/CircaBC/ESTAT/labmarstatwg/Library/statistics_working/2014/1.%2024-25%20June%20(LFS)/2.%20Documents/Doc%2006%20Item%202.5%20Final%20Report%20TF%20Precision%20Requirements.pdf`.

Eltinge, J., and Yansaneh, I. (1997), "Diagnostics for Formation of Nonresponse Adjustment Cells, with an Application to Income Nonresponse in the U.S. Consumer Expenditure Survey," *Survey Methodology*, 23, 33–40.

Hagesæther, N., and Zhang, L. C. (2009), "A Note on the Effect of Auxiliary Information on the Variance of Cluster Sampling," *Journal of Offical Statistics*, 25(3), 397–404.

Hájek, J. (1964), "Asymptotic Theory of Rejective Sampling with Varying Probabilities from a Finite Population," *The Annals of Mathematical Statistics*, 35(4), 1491–1523.

Hájek, J. (1981), *Sampling from a Finite Population*, New York: Marcel Dekker.

Hamre, J. I. (2017), "Kvartalsgjennomsnitt med ukeproporsjonal vekting av månedene,". Internt notat. Statistics Norway, Oslo. Available from the author upon request.

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; accessed 21 September 2017].

Hansen, M. H., and Hurwitz, W. N. (1943), “On the Theory of Sampling from Finite Populations,” *The Annals of Mathematical Statistics*, 14(4), pp. 333–362.

Hansen, M., Hurwitz, W., and Madow, W. (1953), *Sample Survey Methods and Theory, volume I*, New York: John Wiley and Sons.

Haziza, D., and Lesage, É. (2016), “A Discussion of Weighting Procedures for Unit Nonresponse,” *Journal of Offical 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; accessed 21 September 2017].

Holmes, D. J., and Skinner, C. J. (2000), “Variance estimation for Labour Force Survey estimates of level and change,” *Government Statistical Service Methodology Series*, . The Office for National Statistics, London, United Kingdom, 21, 40pp.

Horvitz, D. G., and Thompson, D. J. (1952), “A Generalization of Sampling Without Replacement From a Finite Universe,” *Journal of the American Statistical Association*, 47(260), 663–685.

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

Little, R. J. (1986), “Survey Nonresponse Adjustments for Estimates of Means,” *International Statistical Review*, 54, 139–157.

Little, R. J., and Vartivarian, S. (2005), “Does Weighting for Nonresponse Increase the Variance of Survey Means?,” *Survey Methodolology*, 31(2), 161–168.

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., and Zhang, L. C. (2016), “Reweighting for Nonresponse in Household Surveys at Statistics Norway,”, Unpublished working paper.

Oguz-Alper, M., and Berger, Y. G. (2015), “Variance estimation of change of poverty based upon the Turkish EU-SILC survey,” *Journal of Official Statistics*, 31, 155–175.

Qualité, L. (2009), Unequal probability sampling and repeated surveys,. PhD Thesis, University of Neuchatel, Switzerland.

Qualité, L., and Tillé, Y. (2008), “Variance estimation of changes in repeated surveys and its application to the Swiss survey of value added,” *Survey Methodology*, 34, 173–181.

R Development Core Team (2014), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing. http://www.R-project.org, Vienna, Austria.

Särndal, C. E., and Lundström, S. (2005), *Estimation in Surveys with Nonresponse*, Chichester: Wiley.

Särndal, C.-E., Swensson, B., and Wretman, J. (1992), *Model Assisted Survey Sampling*, New York: Springer-Verlag.

SSB (2001), "Labour Force Survey 2001,", Official Statistics of Norway, C 748, Statistics Norway. `http://www.ssb.no/a/nos/`. [Online; last accessed 30 January 2001].

SAS Institute Inc. (2013), *What's New in Base SAS® 9.4: Details* Cary, NC: SAS Institute Inc.

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; accessed 20 September 2017].

Wu, C. (2003), "Optimal calibration estimators in survey sampling," *Biometrika*, 90(4), 937–951.

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

Yung, W., and Rao, J. (1996), "Jackknife linearization variance estimators under stratified multi-stage sampling," *Survey Methodology*, 22(1), 23–31.

Yung, W., and Rao, J. (2000), "Jackknife variance estimation under imputation for estimators using poststratification information," *Journal of the American Statistical Association*, 95(451), 903–915.

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; accessed 21 September 2017].

Appendix A: Derivation of total differentials

The total differentials $\{\mathrm{d}\widehat{Y}_c\}$, $\{\mathrm{d}\widehat{\boldsymbol{X}}_c\}$ and $\{\mathrm{d}\widehat{\boldsymbol{\beta}}_c\}$ in (30) can be derived by re–writing $\widehat{Y}_c$, $\widehat{\boldsymbol{X}}_c$ and $\widehat{\boldsymbol{\beta}}_c$ as functions of estimated totals. We have, by using (13) and (16),

$$\widehat{Y}_c = f(\widehat{Y}_c^r, \widehat{M}_c^r, \widehat{M}_c) = \sum_{c\in C} \frac{\widehat{M}_c}{\widehat{M}_c^r}\widehat{Y}_c^r, \tag{A.1}$$

where

$$\widehat{Y}_c^r \;=\; \sum_{\{hij\}\in s} d_{hij} r_{hij} \delta_{hij}^c y_{hij},$$

and the estimates $\widehat{M}_c^r$ and $\widehat{M}_c$ are defined in Section 6.4. By applying total differentials to (A.1), we obtain

$$\{\mathrm{d}\widehat{Y}_c\} = \sum_{c\in C}\left\{\frac{\widehat{M}_c}{\widehat{M}_c^r}\{\mathrm{d}\widehat{Y}_c^r\} - \frac{\widehat{M}_c}{(\widehat{M}_c^r)^2}\widehat{Y}_c^r\{\mathrm{d}\widehat{M}_c^r\} + \frac{1}{\widehat{M}_c^r}\widehat{Y}_c^r\{\mathrm{d}\widehat{M}_c\}\right\}\cdot \tag{A.2}$$

When we replace the total differentials of estimated totals in (A.2) by deviations from their expected values (Binder, 1996), we obtain

$$\{\mathrm{d}\widehat{Y}_c\} = \widehat{Y}_c - Y \doteq \sum_{c\in C}\sum_{\{hij\}\in s} d_{hij}\delta_{hij}^c \varrho_{hij}^y + \Omega_1, \tag{A.3}$$

where

$$\varrho_{hij}^y = \widehat{a}_c r_{hij} y_{hij} + \bar{y}_c^r(1 - \widehat{a}_c r_{hij}),$$

where $\widehat{a}_c = \widehat{M}_c/\widehat{M}_c^r$ is the nonresponse adjustment factor for the cth RHC, $\bar{y}_c^r = \widehat{Y}_c^r/\widehat{M}_c^r$ is the estimated mean within the cth RHC, and Ω_1 denotes the remaining terms not depending explicitly on d_{hij} (Binder, 1996, p.18).

The estimate $\widehat{\boldsymbol{X}}_c$ may be re–written as a function of totals in a similar manner as $\widehat{Y}_c$ by using (14) and (16). Thus applying a similar procedure, (A.1) through (A.3), to $\widehat{\boldsymbol{X}}_c$, we obtain

$$\{\mathrm{d}\widehat{\boldsymbol{X}}_c\} = \widehat{\boldsymbol{X}}_c - \boldsymbol{X} \doteq \sum_{c\in C}\sum_{\{hij\}\in s} d_{hij}\delta_{hij}^c \boldsymbol{\varrho}_{hij}^x + \boldsymbol{\Omega}_2, \tag{A.4}$$

where

$$\boldsymbol{\varrho}_{hij}^x = \widehat{a}_c r_{hij}\boldsymbol{x}_{hij} + \bar{\boldsymbol{x}}_c^r(1 - \widehat{a}_c r_{hij}),$$

with $\bar{\boldsymbol{x}}_c^r = \widehat{\boldsymbol{X}}_c^r/\widehat{M}_c^r$, where $\widehat{\boldsymbol{X}}_c^r = \sum_{\{hij\}\in s} d_{hij} r_{hij}\delta_{hij}^c \boldsymbol{x}_{hij}$.

Applying total differentials to $\widehat{\boldsymbol{\beta}}_c$, we obtain

$$\{\mathrm{d}\widehat{\boldsymbol{\beta}}_c\} = (\widehat{\boldsymbol{S}}_{c;xx})^{-1}\left(\{\mathrm{d}\widehat{\boldsymbol{S}}_{c;xy}\} - \widehat{\boldsymbol{\beta}}_c\{\mathrm{d}\widehat{\boldsymbol{S}}_{c;xx}\}\right)\cdot \tag{A.5}$$

Total differentials $\{\mathrm{d}\widehat{\boldsymbol{S}}_{c;xy}\}$ and $\{\mathrm{d}\widehat{\boldsymbol{S}}_{c;xx}\}$ are obtained by re-writing, respectively, $\widehat{\boldsymbol{S}}_{c;xy}$ and $\widehat{\boldsymbol{S}}_{c;xx}$ in terms of estimated totals. Thus we have

$$\widehat{\boldsymbol{S}}_{c;xy} \;=\; f(\widehat{\boldsymbol{S}}_{c;xy}^r, \widehat{M}_c^r, \widehat{M}_c) = \sum_{c\in C}\frac{\widehat{M}_c}{\widehat{M}_c^r}\widehat{\boldsymbol{S}}_{c;xy}^r, \tag{A.6}$$

$$\widehat{\boldsymbol{S}}_{c;xx} \;=\; f(\widehat{\boldsymbol{S}}_{c;xx}^r, \widehat{M}_c^r, \widehat{M}_c) = \sum_{c\in C}\frac{\widehat{M}_c}{\widehat{M}_c^r}\widehat{\boldsymbol{S}}_{c;xx}^r, \tag{A.7}$$

with

$$\widehat{\boldsymbol{S}}^r_{c;xy} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \delta^c_{hij} \boldsymbol{x}_{hij} y_{hij}$$
$$\widehat{\boldsymbol{S}}^r_{c;xx} = \sum_{\{hij\}\in s} d_{hij} r_{hij} \delta^c_{hij} \boldsymbol{x}_{hij} \boldsymbol{x}^\top_{hij}.$$

Taking total differentials of (A.6) and (A.7) and following a similar procedure provided above, (A.2) through (A.3), we obtain

$$\{\mathrm{d}\widehat{\boldsymbol{S}}_{c;xy}\} = \widehat{\boldsymbol{S}}_{c;xy} - \boldsymbol{S}_{xy} \doteq \sum_{c\in C} \sum_{\{hij\}\in s} d_{hij} \delta^c_{hij} \boldsymbol{\omega}^{xy}_{hij} + \boldsymbol{\Omega}_4, \qquad \text{(A.8)}$$
$$\{\mathrm{d}\widehat{\boldsymbol{S}}_{c;xx}\} = \widehat{\boldsymbol{S}}_{c;xx} - \boldsymbol{S}_{xx} \doteq \sum_{c\in C} \sum_{\{hij\}\in s} d_{hij} \delta^c_{hij} \boldsymbol{\omega}^{xx}_{hij} + \boldsymbol{\Omega}_3, \qquad \text{(A.9)}$$

where

$$\boldsymbol{\omega}^{xy}_{hij} = \widehat{a}_c r_{hij} \boldsymbol{x}_{hij} y_{hij} + (\widehat{M}^r_c)^{-1} \widehat{\boldsymbol{S}}_{c;xy}(1 - \widehat{a}_c r_{hij}),$$
$$\boldsymbol{\omega}^{xx}_{hij} = \widehat{a}_c r_{hij} \boldsymbol{x}_{hij} \boldsymbol{x}^\top_{hij} + (\widehat{M}^r_c)^{-1} \widehat{\boldsymbol{S}}_{c;xx}(1 - \widehat{a}_c r_{hij}).$$

Appendix B: Tables

Table B.1: The ratio of sampling fractions to the overall sampling fraction in Norwegian counties

Region code	Region name	County code	County name	Ratio between sampling fractions
01	Akershus og Oslo	02	Akershus	0.871
		03	Oslo	0.871
02	Hedmark og Oppland	04	Hedmark	1.000
		05	Oppland	1.000
03	Sør–Østlandet	01	Østfold	1.000
		06	Buskerud	1.000
		07	Vestfold	1.000
		08	Telemark	1.000
04	Agder og Rogaland	09	Aust–Agder	1.484
		10	Vest–Agder	1.000
		11	Rogaland	1.000
05	Vestlandet	12	Hordaland	0.871
		14	Sogn og Fjordane	1.398
		15	Møre og Romsdal	1.000
06	Trøndelag	16	Sør–Trøndelag	1.000
		17	Nord–Trøndelag	1.151
07	Nord–Norge	18	Nordland	1.000
		19	Troms	1.000
		20	Finnmark	1.851

Source: Vedø and Rafat (2003, p.7).

Table B.2: Descriptions of variables used in the estimation (see Sections 11.1-1).

Variable[1]	Label	Categories
Gender	gender (2)	male, female
Age	age (2)	15-59, 60-74
	age (3)	15-24, 25-54, 55-74
	age (6)	15-24, 25-39, 40-54, 55-61, 62-66, 67-74
	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
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 (7)	Akershus and Oslo, Hedmark and Oppland, Sør–Østlandet Agder and Rogaland, Vestlandet, Trøndelag, Nord–Norge
NUTS III	region (19)	Østfold, Akershus, Oslo, Hedmark, Oppland, Buskerud, Vestfold Telemark, Aust–Agder, Vest–Agder, Rogaland, Hordaland, Sogn and Fjordane, Møre and Romsdal, Sør–Trøndelag, Nord–Trøndelag, Nordland, Troms Romsa, Finnmark

[1] Variables used in the estimation come from up–to–date administrative registers.

Table B.2 continued: Descriptions of variables used in the estimation (see Sections 11.1-1).

Variable[1]	Label	Categories
Education	education (4)	primary school, high school, higher education, others
	education (3)	high school, higher education, primary school and others
Marital status	marstat (2)	married or registered partner, others
Family size	familysize (3)	1 person, 2 persons and 3 or more persons
Country of origin[2]	country (3)	1: not immigrants; 2: immigrants coming from EEA (European Economic Area), USA, New Zealand, Canada and Australia; 3: immigrants coming from other countries, stateless or others
Income	income (5)	four categories formed by quartiles of income and one category for not wage–earners
Scheme[3]	tiltak (3)	1: unemployed, ordinary scheme participant, salary subsidies, skill–training scheme, temporary employment scheme and other ordinary schemes; 2: occupationally handicapped or reduced working capacity in scheme and not in scheme; 3: others

[1] Variables used in the estimation come from up–to–date administrative registers.

[2] Based on a register variable providing country background back to three generations and a register variable indicating immigration status.

[3] More information about scheme can be found here: `https://www.nav.no/no/NAV+og+samfunn/Statistikk/Arbeidssokere+og+stillinger+-+statistikk/Tiltaksdeltakere`

Statistics Norway

Postal address:
PO Box 8131 Dept
NO-0033 Oslo

Office address:
Akersveien 26, Oslo
Oterveien 23, Kongsvinger

E-mail: ssb@ssb.no
Internet: www.ssb.no
Telephone: + 47 62 88 50 00

ISBN 978-82-537-9729-8 (electronic)

Statistisk sentralbyrå
Statistics Norway

Design: Siri Boquist