Estimation of the standard error for net changes with the EU Labour Force Survey – How can users independently and appropriately calculate standard errors and confidence intervals?

Bernhard Schimpl-Neimanns, GESIS – Leibniz Institute for the Social Sciences, bernhard.schimpl-neimanns@gesis.org

Stefan Zins, GESIS – Leibniz Institute for the Social Sciences, stefan.zins@gesis.org

Abstract

The EU Labour Force Survey (EU-LFS) is one of the most important official surveys for comparative social research in Europe. As such it is a source for the estimation of indicators that allow the monitoring of economic and social policy. Assessing whether observed changes in indicators are significant requires estimating the variance of estimated changes. This task is challenging as most countries use complex sampling designs and different rotation schemes. Due to their partial overlap between waves, rotational panels allow a more efficient estimation of changes. To account for this, the covariance of cross-sectional estimates has to be estimated. In practice, users can face some difficulties in doing so because longitudinally consistent identifiers are required. However, the data released by Eurostat for scientific purposes currently contain identifiers for the primary sampling units that are randomized per dataset and are only consistent for one year, supporting the erroneous assumption of statistical independence between waves. By taking the example of LFS-data from Austria we show how the available design information can be used to estimate the variance of a change in the cross-sectional estimates. Using the Austrian Microcensus data we circumvent the problem of missing inconsistent longitudinally identifier in the Austrian LFS, thereby showing that proper variance estimation is feasible. We recommend that variables for stratum, clustering, weight, and time consistent unit identifiers should be released if there are no confidentiality concerns.

Keywords: EU Labour Force Survey, Variance Estimation, Linearisation, Net Change

1. Introduction

As part of both the Lisbon Strategy and the “Europe 2020” strategy, annual evaluations of indicators are used to monitor progress towards EU policy goals. Examples of this are changes in labour market participation or in the economic situation of households. When estimating indicators the sampling error is an important measure for the precision of the results and vital to hypothesis testing. For example, if we would like to test if an observed net change over time of estimated indicator values was significant or not.

Almost all EU-LFS surveys are rotational panels that exchange part of the cross-sectional sample every quarter of the year. This represents a compromise between a panel and a repeated cross-sectional survey. The rotational design has the benefit of enabling unbiased estimates for each year, like a cross-sectional survey, while it simultaneously provides the basis for precise estimates of change, like a panel survey. However, the quarterly samples of the rotational panels are not statistically independent, and this temporal dependency must be taken into account when estimating changes over time. Thus, in addition to sufficient information on the cross-sectional sample design, longitudinally consistent identifiers are required. But EU-LFS data released by Eurostat for scientific purposes contains household identifiers which are only consistent for one year (Eurostat 2016). Statistics Austria provides longitudinally consistent identifiers for the Austrian Microcensus (MC), which can also be used for the Austrian EU-LFS. This allows an examination of the potential error in estimation under the erroneous assumption of independent samples between years.

2. The Sampling Design of the Austrian EU-Labour Force Survey

The Austrian LFS is part of the MC. The cross-sectional sampling design of the MC is a stratified cluster sample, with households as clusters and NUTS-2 regions as strata. Each household is kept in the sample for five consecutive quarters, and every quarter, the 20% of households that have been interviewed for five quarters are replaced by a new random sample (for details see Eurostat 2018 and Meraner et al. 2016). This rotational scheme creates an overlap of 80% between consecutive cross-sectional samples. Overall the sample contains approximately 20,000 households with 44,000 persons per quarter.

For each cross-sectional sample survey weights are calculated using register data so that weighted distributions of the sample reproduce the population values. The design weights of the cross-sectional sample (inverse of the inclusion probabilities) are calibrated using the raking method. The register data used consist of the total number of persons in private households in a region separately for categories of gender, age, nationality, household size, and administrative employment status. Furthermore, the total number of private households in a region of a particular household size is included. The calibration specification also requires that each member of the household has the same weight.

EU-LFS micro data from Eurostat is anonymised and household identifiers are randomized for every year (Eurostat 2016). Therefore, tracking households and persons across years is not possible. Fortunately, supplementary files for the Austrian MC with data from the labour market service contain longitudinally consistent household IDs. The MC IDs can be merged one-to-one with the LFS data for each quarter, solving the problem of inconsistent longitudinal household IDs for the LFS. In addition, Statistics Austria provides bootstrap replicate weights, which have been built to reflect the cross-sectional and longitudinal properties of the MC design. The R package mzR (Kowarik & Meraner 2017) use these bootstrap replicate weights to compute estimates and standard errors for the Austrian MC and LFS.

3. Estimation

3.1. Design-based estimates for quarterly cross-sectional samples

All of our statistics of interest are estimated as the ratio of two population totals (see Section 4). We use the following estimator, $\hat{\theta}_{tq}$, to estimate our statistic of interest, ratio θ_{tq} , for the *q*-th quarter of the *t*-th year:

$$\hat{\theta}_{tq} = \frac{\sum_{k \in s_{tq}} w_{tqk} y_{tqk}}{\sum_{k \in s_{tq}} w_{tqk}\, z_{tqk}} \tag{1}$$

Where y_{tqk} and z_{tqk} are the numerator and denominator variables of interest for the *k*-th respondent in s_{tq}, the sample in the *q*-th quarter of the year *t*, and w_{tqk} is the raking weight associated with the *k*-th respondent in s_{tq}. The survey weights are computed by a raking procedure (Meraner et al. 2016). The estimator $\hat{\theta}_{tq}$ is non-linear, so we use a linearisation approach to variance estimation (Särndal et al. 1992, p. 178). We first derive the so-called differential variable d_{tqk} for the ratio $\hat{\theta}_{tq}$: $d_{tqk} = \left(y_{tqk} - \hat{\theta}_{tq} z_{tqk}\right) / \hat{Z}_{tq}$ with $\hat{Z}_{tq} = \sum_{k \in s_{tq}} w_{tqk} z_{tqk}$. Raking weights are based on the sample s_{tq}, and are therefore not independent of y_{tqk} and z_{tqk}. Therefore, we calculate the residuals ϵ_{tqk} of the ordinary least square regression (weighted with the raking weights) of d_{tqk} on the auxiliary variables used for raking (D'Arrigo & Skinner 2010). Finally we use the following variance estimator for $\hat{\theta}_{tq}$:

$$\widehat{Var}(\hat{\theta}_{tq}) = \sum_{h=1}^{H} N_{Itqh}^2 \left(\frac{1}{n_{Itqh}} - \frac{1}{N_{Itqh}}\right) V_{Itqh}^2 \tag{2}$$

Where *H* is the number of strata (see Section 4.1), N_{Itqh}^2 is the number of households in the *h*-th stratum of s_{Itq}, the sample of households in quarter *q* of year *t*, and n_{Itqh} is the number of sampled households in the *h*-th stratum of s_{Itq}. Also, $V_{Itqh}^2 = \sum_{i \in s_{Itqh}} \left(E_{tqhi} - \bar{E}_{tqh}\right)^2 / \left(n_{Itqh} - 1\right)$ with $E_{tqhi} = \sum_{k \in s_{Itqhi}} \epsilon_{tqk}$ and $\bar{E}_{tqh} = E_{tqhi}/n_{Itqh}$, where s_{Itqh} is the set of households in the *h*-th stratum of s_{Itq}, and s_{Itqhi} is the set of respondents in the *i*-th household in the *h*-th stratum of s_{Itq}.

3.2 Estimates for annual averages and net changes

Assuming that the size of the population is (almost) equal in each quarter of the same year, the annual average is the sum of the quarterly estimates divided by 4. The estimator for an annual average of statistic θ in year *t* is given by:

$$\hat{\theta}_t = \frac{1}{4} \cdot \sum_{q=1}^{4} (\hat{\theta}_{tq}) \tag{3}$$

Where $\hat{\theta}_{tq}$ is the estimator for statistic θ in the in the *q*-th quarter of the *t*-th year.

When estimating annual averages, it must be taken into account that the quarterly estimates are not statistically independent due to the overlap of the rotation design. The variance of estimator $\hat{\theta}_t$ is the given by:

$$Var(\hat{\theta}_t) = \left(\frac{1}{4}\right)^2 \cdot \left[\sum_{q=1}^{4} \sum_{p=1}^{4} Cov(\hat{\theta}_{tq}, \hat{\theta}_{tp})\right] \tag{4}$$

We estimate net change as the difference of the annual average of statistic θ between years *t* and *u* by:

$$\hat{\Delta}_{t,u} = \hat{\theta}_u - \hat{\theta}_t \tag{5}$$

The variance of $\hat{\Delta}_{t,u}$ is given by:

$$Var(\hat{\Delta}_{t,u}) = \left(\frac{1}{4}\right)^2 \cdot \left[\sum_{v \in \{t,u\}} \sum_{w \in \{t,u\}} \sum_{q=1}^{4} \sum_{p=1}^{4} Cov(\hat{\theta}_{vq}, \hat{\theta}_{wp})\right] \tag{6}$$

Where $Cov(\hat{\theta}_{wq}, \hat{\theta}_{vp})$ is the covariance between estimators $\hat{\theta}_{wq}$ and $\hat{\theta}_{vp}$. Note that for *w*=*v* and *q*=*p* we have $Cov(\hat{\theta}_{wq}, \hat{\theta}_{wq}) = Var(\hat{\theta}_{wq})$. An unbiased estimator for $Var(\hat{\theta}_t)$ is given by:

$$\widehat{Var}(\hat{\Delta}_{t,u}) = \left(\frac{1}{4}\right)^2 \cdot \left[\sum_{v\in\{t,u\}} \sum_{w\in\{t,u\}} \sum_{q=1}^{4} \sum_{p=1}^{4} \sum_{k\in s_{vq}} \sum_{l\in s_{wp}} \frac{(\pi_{vwqpkl} - \pi_{vqk}\pi_{wpl})}{\pi_{vwqpkl}} \frac{\varepsilon_{vqk}}{\pi_{vqk}} \frac{\varepsilon_{wpl}}{\pi_{wpl}}\right] \quad (7)$$

Where π_{vqk} is the probability of including element *k* in sample s_{vq}, π_{vwqpkl} is the joint probability of including element *k* into sample s_{vq} and element *l* into sample s_{wp}, and $\varepsilon_{vqk} = \epsilon_{vqk}$ if *v*=*u* and $\varepsilon_{vqk} = -\epsilon_{vqk}$ if *v*=*t*.

Because of its complexity, the above variance estimator is not very practical. That is why approximations to $Var(\hat{\Delta}_{t,u})$ have been sought, which are less complex to estimate (e.g. Berger (2004) and Wood (2008)). We use an estimator proposed by Berger & Priam (2016) that assumes negligible sampling fractions of samples s_{vq} *v*=*t*, *u* and *q*=1,…,4. This estimator uses a multivariate regression of the household totals of the linearised variable $\check{\tau}_{vqhi}$ on the household sample indicators c_{vqhi} and all their first order interactions $c_{vqhi}c_{wphj}$. Where $c_{vqhi} = 1$ if $i \in s_{Ivqh}$ else, $c_{vqhi} = 0$. $\check{\tau}_{vqhi}$ is given by:

$$\check{\tau}_{vqhi} = \frac{\hat{\tau}_{vqhi}}{\pi_{vqhi}} c_{vqhi} \text{ with } \hat{\tau}_{vqhi} = \sum_{k\in s_{Ivqhi}} \varepsilon_{vqk} \quad (8)$$

Where $\pi_{vqhi} = n_{Itqh}/N_{Itqh}$. The ordinary least squares estimate of the correlation matrix of the residuals $\hat{\Upsilon}$ is then used as an estimate for the correlation matrix of the eight cross-sectional estimates $\hat{\theta}_{vq}$ *v*=*t*, *u* and *q*=1,...,4. This gives the following variance estimator:

$$\widehat{Var}^a(\hat{\Delta}_{t,u}) = \left(\varsigma^{1/2}\right)^{\mathrm{T}} \hat{\Upsilon}\, \varsigma^{1/2} \quad (9)$$

Where $\varsigma = \left[\widehat{Var}(\hat{\theta}_{wq})\right]_{\substack{w=\{t,u\} \\ q=1,\dots,4}}$ is the column vector of the variance estimates of the cross-sectional estimates $\hat{\theta}_{vq}$ *v*=*t*, *u* and *q*=1,...,4 and $(x)^{\mathrm{T}}$ is the transpose of a vector *x*.

4. Results

4.1. Data preparation

We use the most recent annual data available from the Austrian LFS for 2016 (2017 release) and 2015 (2016 release). Variables for the identification of strata and households, as well as for the reporting quarter, are created according to the sampling design. Due to coarsening, the stratification variable distinguishes 3 strata (groups of NUTS-2 regions) instead of the actual 9. Indicators are defined for the numerator and denominator variables of LFS-based EU 2020 headline indicators and the associated subpopulations. We are interested in estimating the indicators *Employment rate, Tertiary educational attainment*, and the *Employment rate of older workers*. The last indicator is not an EU 2020 headline indicator, but it is used in monitoring (European Commission 2018: p. 50). The third EU 2020 indicator *Early leavers from education and training* (of the population aged 18-24) cannot be replicated because the age information in the data does not correspond to the age group of the headline indicator.

4.2. Estimates

We would like to investigate how our own estimates, based on the linearisation approach, compare to official results based on the original data and the bootstrap replicated weights for variance estimation. We have used publications from Statistics Austria for this purpose (see e.g. Statistik Austria 2017a, 2017b), but due to the higher accuracy (without rounding), we report results achieved with the R program mzR. For reasons of space, only summary results for annual comparisons are shown below. The auxiliary variables used for weighting are constructed according to the specifications in Meraner et al. (2016), pp. 7-8. There are some restrictions. Instead of 9 federal states, only 3 are available in the LFS. The data does not allow us to identify Turkish citizenship. Hence, only 5 different nationalities can be used. Because administrative employment status is unknown, we use administrative job-seeker status as a proxy, which originates from the Public Employment Service (AMS) that is available in a supplementary file for the Austrian MC. This supplementary file also contains the longitudinally consistent household ID. These changes to the calibration specifications can be expected to cause differences between our own estimates and those of Statistics Austria.

Table 1 shows our results for the employment rate. We estimate an increase from 74.3% in 2015 to 74.8% in 2016, close to the EU-28 target of 75%, but still below the national target of 77%. Our estimator for change using linearization, ${}^{1}_{lin}\Delta_{2015,2016}$, gives a standard error of 0.32%. The corresponding 95% confidence interval ranges from 0.01% to 1.09%. As it does not include the null, the estimated change is statistically significant. This corresponds to the result for ${}^{1}_{\mathrm{mzR}}\widehat{\Delta}_{2015,2016}$, ${}_{\mathrm{mzR}}\Delta_{2015,2016}$, obtained with the bootstrap method using the mzR program. ${}^{1}_{\mathrm{w/o}}\widehat{\Delta}_{2015,2016}$ is our estimator if we had no access to longitudinally consistent identifiers for the households and assume statistical independence between the years. Consequently, ${}^{1}_{\mathrm{w/o}}\widehat{\Delta}_{2015,2016}$ has a larger standard error and the corresponding confidence interval does include the null, thus the estimated change would not be regarded as significant.

Table 1. Employment Rate: Annual Averages, Net Changes, and Standard Error Estimates.

Estimator	Ratio (%)	S.E. (%)	CI l 2.5%	CI u 97.5%
${}^{1}\hat{\theta}_{2016}$	74.80	0.2269	74.36	75.25
${}^{1}\hat{\theta}_{2015}$	74.26	0.2281	73.81	74.70
${}^{1}_{lin}\hat{\Delta}_{2015,2016}$	0.55	0.2773	0.01	1.09
${}^{1}_{w/o}\hat{\Delta}_{2015,2016}$	*0.55*	*0.3218*	*-0.08*	*1.18*
${}^{1}_{mzR}\Delta_{2015,2016}$	0.55	0.1536	0.24	0.85

Source: EU-LFS 2015 (2016 release) and 2016 (2017 release), annual data, weighted results, own calculation. The indicator is defined as the percentage of the population aged 20-64 in employment by the total population of the same age group.

Table 2 presents our results for the employment rate of older people using the same types of estimators as in Table 1. The indicator rose by 2.9 percentage points to 49.2% in 2016 compared to the previous year. The confidence intervals for the difference in all three estimators ${}^{2}_{\mathrm{lin}}\widehat{\Delta}_{2015,2016}$, ${}^{2}_{\mathrm{w/o}}\widehat{\Delta}_{2015,2016}$, and ${}^{2}_{\mathrm{mzR}}\widehat{\Delta}_{2015,2016}$ do not include the null , i.e. with all three estimators, the estimated change test as statistically significant. But ${}^{2}_{\mathrm{w/o}}\widehat{\Delta}_{2015,2016}$ has a 17 percentage points higher standard error than ${}^{2}_{\mathrm{lin}}\widehat{\Delta}_{2015,2016}$.

Table 2. Employment Rate of Older People: Annual Averages, Net Changes, and Standard Error Estimates.

Year	Ratio (%)	S.E. (%)	CI l 2.5%	CI u 97.5%
${}^{2}\hat{\theta}_{2016}$	49.17	0.5497	48.09	50.24
${}^{2}\hat{\theta}_{2015}$	46.27	0.5460	45.20	47.34
${}_{lin}^{2}\widehat{\Delta}_{2015,2016}$	2.89	0.6615	1.60	4.19
${}_{w/o}^{2}\widehat{\Delta}_{2015,2016}$	*2.89*	*0.7748*	*1.37*	*4.41*
${}_{mzR}^{2}\widehat{\Delta}_{2015,2016}$	2.89	0.6137	1.73	4.06

Source: See Table 1. The indicator is calculated by dividing the number of persons aged 55 to 64 in employment by the total population of the same age group.

Table 3. Tertiary educational attainment: Annual Averages, Net Changes, and Standard Error Estimates.

Year	Ratio (%)	S.E. (%)	CI l 2.5%	CI u 97.5%
${}^{3}\hat{\theta}_{2016}$	40.11	0.8724	38.40	41.82
${}^{3}\hat{\theta}_{2015}$	38.74	0.8848	37.00	40.47
${}_{lin}^{3}\Delta_{2015,2016}$	1.38	1.0640	-0.71	3.46
${}_{w/o}^{3}\Delta_{2015,2016}$	*1.38*	*1.2425*	*-1.06*	*3.81*
${}_{mzR}^{3}\Delta_{2015,2016}$	1.38	1.0722	-0.80	3.35

Source: See Table 1. The indicator is defined as the percentage of the population aged 30-34 who have successfully completed tertiary studies.

Table 3 displays the results for tertiary educational attainment. It should be at least 40% by 2020 in the EU. Both this goal and the national target of 38% were already achieved by 2016. The increase of around 1.4 percentage points between 2015 and 2016 is not statistically significant for any of the three estimates. And, as before, when looking at the employment rate of older people, ${}_{\mathrm{w/o}}^{3}\widehat{\Delta}_{2015,2016}$ has a 17 percentage points higher standard error than ${}_{\mathrm{lin}}^{3}\Delta_{2015,2016}$.

As expected, the standard errors estimated with the linearisation approach, taking into account the covariance between years, differ from the results obtained with the bootstrap method. While the standard error estimated using the linearisation approach differ by a factor of round 1.8 (0.2773 / 0.1536) for the employment rate indicator, the difference in the employment rate of older workers indicator (0.6615 / 0.6137 ≈ 1.1) and the tertiary educational attainment rate indicator (1.0640 / 1.0722 ≈ 1.0) is substantially smaller. Apart from the fact that the methods differ, it can be assumed that one reason for the differences is the missing or coarsened information on stratification and calibration specifications available in the data. In particular, the administrative employment status is unknown. Nevertheless, decisions on the statistical significance of annual changes lead to the same conclusions. The confidence intervals estimated using the linearisation approach are - with the exception of the employment rate indicator - close to those estimated using the bootstrap method.

5. Conclusions

The LFS data are used to monitor change between indicators such as the employment rate. Estimating the variance of change for this type of rotating sampling surveys is challenging because temporal correlations between cross-sectional estimates must be calculated. However, the data currently released by Eurostat contains household identifiers which are only consistent within the same year. By taking LFS-data from Austria and supplementary data for the Austrian Microcensus with longitudinally consistent household identifiers, an empirical examination of the potential error in estimations under the erroneous assumption of statistically independent samples between consecutive years was possible. Our design-based estimation with the linearisation approach to variance estimation has proven to be consistent with the proposed estimation strategy by Statistics Austria using the R mzR package. But assuming erroneous statistical independence between the years leads to standard errors with an upward bias and confidence intervals that are too wide. In the case of the employment rate indicator, assuming statistical independence results a negative value for the lower limit of the confidence interval, leading to not rejecting the null hypotheses that the indicator did not change between 2015 and 2016. This example shows that assuming statistical independence can sometimes lead to a different conclusion, as compared to more accurate variance estimators. Similar problems can also arise in the analysis of gross changes and other longitudinal analyses.

Both to avoid potential statistical fallacies and to enable researchers to estimate more accurate standard errors and confidence intervals on their own, we recommend that variables for stratum, clustering, weighting, and completely time consistent unit identifiers should be released if there are no confidentiality concerns. This would considerably improve variance estimations based on anonymised microdata.

6. References

Berger, Y. G. (2004), Variance estimation for measures of change in probability sampling, Canadian Journal of Statistics, 32, pp. 451-467.

Berger, Y. G. and Priam, R. (2016), A simple variance estimator of change for rotating repeated surveys: an application to the European Union Statistics on Income and Living Conditions household surveys, Journal of the Royal Statistical Society: Series A, 179, pp. 251-272.

D'Arrigo, J. and Skinner, C.J. (2010), Linearization variance estimation for generalized raking estimators in the presence of nonresponse, Survey methodology, 36, pp. 181-192.

European Commission (2018), Commission Staff Working Document, Country Report Austria 2018, accompanying the document “Communication from the Commission to the European Parliament, the Council, the European Central Bank and the Euro-group. SWD(2018) 218 final. Brussels. Available at: https://ec.europa.eu/info/sites/info/files/2018-european-semester-country-report-austria-en.pdf (Accessed: 03 April 2018).

Eurostat (2018), Labour force surveys in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2016, Publications Office of the European Union, Luxembourg.

Eurostat (2016), Criteria for the anonymization of LFS Microdata. 2016 release. Available at: http://ec.europa.eu/eurostat/documents/1978984/6037342/Criteria-for-the-anonymisation.pdf (Accessed: 03 April 2017).

Kowarik, A. and Meraner A. (2017), mzR: mzR – To compute estimates and their errors for Austrian microcensus data. R package version 1.1.0.

Meraner, A, Gumprecht, D, and Kowarik A. (2016), Weighting Procedure of the Austrian Microcensus using Administrative Data. Austrian Journal of Statistics, 45, pp. 3-14.

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

Statistik Austria (2017a), Arbeitsmarktstatistik Mikrozensus-Arbeitskräfteerhebung 2014, 4. Quartal, Schnellbericht 5.8. Wien. Available at: http://www.statistik.at/wcm/idc/idcplg?IdcService=GET_PDF_FILE&RevisionSelectionMethod=LatestReleased&dDocName=111797 (Accessed: 04 April 2018).

Statistik Austria (2017b), Arbeitsmarktstatistiken. Ergebnisse der Mikrozensus-Arbeitskräfteerhebung und der Offenen-Stellen-Erhebung. Wien. Available at: http://www.statistik.at/wcm/idc/idcplg?IdcService=GET_NATIVE_FILE&RevisionSelectionMethod=LatestReleased&dDocName=112779 (Accessed: 04 April 2018).

Wood, J. (2008), On the Covariance Between Related Horvitz-Thompson Estimators. Journal of Official Statistics, 24, pp. 53–78.