New weighting methods, including explicit correction of sampling weights for non-response and attrition, in the reformed Belgian Labour Force Survey

C. Vanderhoeft, A. Depickere and A. Termote

Statbel (Statistics Belgium)

Keywords: LFS indicators, panel design, mixed effects logistic regression, 2-step calibration, breaks in time series.

1. INTRODUCTION

In 2017, Statbel, the Belgian statistical office, introduced a major reform for the Labour Force Survey (LFS): after 18 years of working with a continuous survey, the switch was made to a panel survey. The most important aspects of this reform are:

- The transfer to an **infra-annual** (**"quarterly"** for the Belgian case***)* rotating panel** design. A sample (called *rotation group* (RG)) of private households is drawn each quarter, independently of previously drawn samples. The sample is *rotating* in the sense that any specific rotation group stays in the survey during 18 months (6 quarters), after which it is replaced by a new rotation group. In the panel survey, each member (of at least 15 years old) in a selected household is asked to complete a questionnaire four times, i.e. during four *waves*, according to a 2(2)2 scenario: a selected household/individual is asked to complete a questionnaire during two consecutive quarters (wave 1 and wave 2), is then not in the survey during the next two quarters, and is again asked to complete a questionnaire during the next two quarters (wave 3 and wave 4).

- The introduction of **mixed mode data collection** techniques. In the first wave and after an introductory letter, the selected households are contacted by an interviewer and CAPI is used to collect the data. In the three follow-up waves, data can be delivered through CAWI or CATI, according to the household's preference.

- Application of the **wave approach**. Information on structural variables is gathered in the first wave only; information on core variables is collected in all four waves.

- A revision of the **weighting methods**. More attention is paid to the correction of effects of non-response (in the first wave) and panel attrition (in the follow-up waves). This resulted in a two-step weighting approach: in step 1, response probabilities are estimated through a mixed effects logistic regression model, and aggregates of the estimated probabilities are used to correct the sampling weights; in step 2, the corrected weights from step 1 are calibrated to the population of interest.

The present text is focussing on the latter aspect of the reform, i.e. the weighting methods. We show the effect of changing the weighting method by comparing estimates for various LFS indicators based on the new 2-step approach, with estimates based on the

old 1-step approach which was used for the continuous LFS from 1999 to 2016. We argue that the new method is better correcting for non-response and attrition bias. Furthermore, it will be shown that different changes in the new methodology are, to some extent, cancelling out, causing only moderate breaks in time series for some major indicators.

For more information on the various aspects of the new panel design, we refer to **Fout! Verwijzingsbron niet gevonden.** or [2]. We discuss the methodology, as outlined in the next paragraphs, at length in [3].

2. METHODS

2.1. Sampling

The sampling design is basically the same as before the reform. Each quarter, a stratified PPS-SYS sample of PSUs is drawn in the first stage, and for each selection of a PSU, a fixed number of private households is selected randomly in stage 2. Before the reform, each such sample served, after simple post-stratification, to estimate all quarterly LFS indicators for structural and core variables. After the reform, each such sample, called *rotation group*, primarily serves to replace a similar rotation group that was drawn and launched six quarters earlier.

2.2. Calibration until 2016 T4 for quarterly estimation of all variables

The "traditional" calibration model used in each quarter for the continuous LFS can be formulated briefly as follows:

$$\mathbf{< IND;\ } \boldsymbol{d}\mathbf{;\ Strat12 \times Sex \times Agecat;\ Lin >}$$

This means that the sample of individual respondents is calibrated to the population by a complete crossing of variables Strat12 (NUTS 2 level, i.e. the sampling strata), Sex and Agecat (5-year age classes, with open ended last class 75+), starting from the sampling weights *d*. The choice of the linear method, indicated by "Lin" in the above formal expression, is irrelevant, because of the complete crossing: this model is an ordinary post-stratification model. The (quarterly) reference population consisted of all Belgian citizens, including those living in collective households.

2.3. Calibration since 2017 T1 for quarterly estimation of core variables

A consequence of the new panel design is that quarterly figures are based on a quarterly sample which consists of households/individuals from four different, but independent, rotation groups which are in four different waves. The weighting procedure therefore should take into account that non-response and successive panel attrition cause some imbalance between the four parts of the quarterly sample. Also, the potential effect of the mixed mode approach should not be overlooked.

The traditional model is adapted to the following:

$$\mathbf{< IND;\ } \boldsymbol{d/\tilde{p}}\mathbf{;\ Strat12 \times Sex \times Agecat + RG\_c;\ Lin >}$$

Thus, the sampling weights *d* are corrected using estimated response probabilities $\tilde{p}$, and the linear structure of the calibration model is extended with contrast constraints, i.e. balancing constraints, between the sub-samples of the quarterly sample. In the presentation we will deal extensively with the *random intercept logistic regression model*

that is used to estimate the response probabilities; the random effect is introduced through a variable identifying the (sampled) PSUs. The approach is inspired by [4]. Also, the introduction of contrast constraints between the RGs, indicated by "RG_c" in the above formal expression, will be motivated carefully.

2.4. Estimating yearly indicators for core variables

Yearly estimates of indicators for core variables are merely (unweighted) averages of the corresponding four quarterly estimates. No new weighting is needed.

2.5. Calibration since 2017 for yearly estimation of structural variables

Data on structural variables are only collected in wave 1. Therefore, the sample used to produce yearly estimates for structural variables consists of four different, and independent, rotations groups, which are all in wave 1, though observed in four different quarters. Eurostat asks to assure consistency between some quarterly and yearly figures.

The model used to calibrate the wave 1 sub-sample for estimation of structural variables, can be formulated as follows:

$$< \mathbf{IND};\ \boldsymbol{d/\tilde{p}};\ \mathbf{Strat12 \times Sex \times Agecat + Region \times Sex \times Agecat^* \times StatBIT;\ Lin} >$$

The new term in the linear structure of the model, where Agecat* is a grouped version of Agecat, Region is NUTS 1 level and StatBIT is ILO status, allows to satisfy the consistency requirements of Eurostat. Benchmarks are full-sample annual averages from LFS.

3. RESULTS

Net breaks in time series for various indicators are composed of effects of several aspects of the switch from a continuous LFS to the panel survey. Each of them is expected to have a considerable impact, but it will be shown that the joint effect is rather modest.

3.1. Effect of correcting sampling weights for non-response and panel attrition

By applying old and new weighting methods to both old and new data, it can be shown that explicit correction for non-response and attrition has a systematic positive effect on estimates of indicators such as unemployment rates, and a systematic negative effect on estimates of indicators such as employment rates.

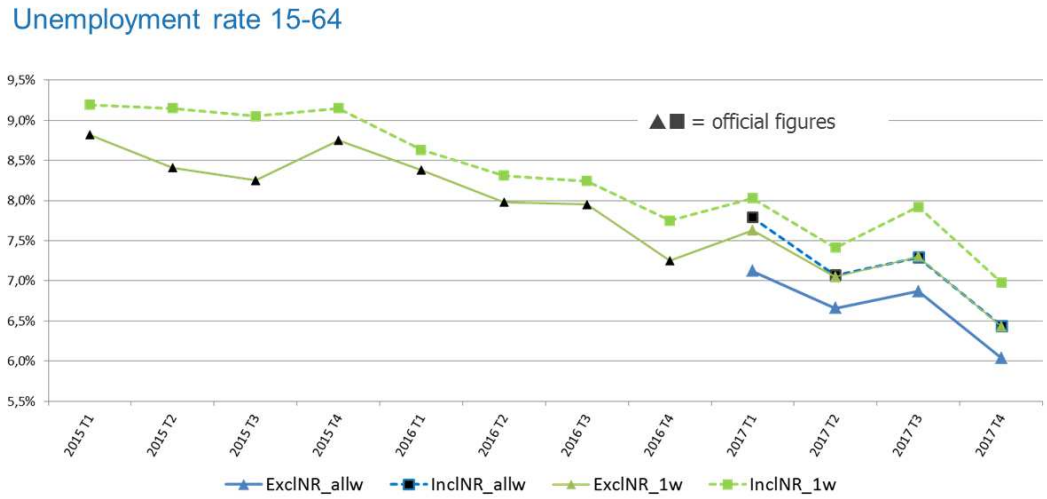
3.2. Effect of introducing the panel approach, and mixed mode data collection

The combined effect of working each quarter with data from 4 waves instead of 1, and collecting data with CAPI in wave 1 and CAWI/CATI in waves 2, 3 and 4, is shown to be negative on, for instance, unemployment rates, and positive on, for instance, employment rates.

3.3. Overall effect of the new methodology

The picture below shows that, for instance for the unemployment rate in age class 15-64, the above mentioned effects are cancelling out: the shift from the solid green to the dashed green line, and from the solid blue to the dashed blue line, indicate the positive effect of explicit correction for non-response and, if data from all waves are included, attrition; the shift from the solid green to the solid blue line, and from the dashed green to

dashed blue line, indicate the negative effect of using data from all waves instead of one (first) wave. Similar results will be presented for other indicators.



4. CONCLUSIONS

We have developed a mixed effects logistic regression model to estimate response probabilities for LFS, and demonstrated the (quality improving) impact of explicit correction for non-response (and attrition). It has also been shown that the impact of introducing the panel design and the mixed mode data collection approach, is cancelled out (or at least reduced) by an explicit correction of sampling weights, before proper calibration. Published time series therefore do not suffer a lot from the introduction of the new panel design in the first quarter of 2017.

REFERENCES

[1] A. Termote en A. Depickere, De hervormde enquête naar de arbeidskrachten in 2017, Statbel Analyses (2018). Downloadable from: https://statbel.fgov.be/nl/over-statbel/methodologie/analyses

[2] A. Termote et A. Depickere, Réforme de l'enquête sur les forces de travail en 2017, Analyses Statbel (2018). Downloadable from: https://statbel.fgov.be/fr/propos-de-statbel/methodologie/analyses

[3] C. Vanderhoeft, A. Depickere and A. Termote, New weighting methods, including explicit correction of sampling weights for non-response and attrition, in the reformed Belgian Labour Force Survey, Analyses Statbel (2019). (To be published)

[4] J. Pickery and A. Carton, Oversampling in relation to differential response rates, Survey Research Methods Vol.2 Nr. 2 (2008), 83-92.