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

Camille VANDERHOEFT, Astrid DEPICKERE, Anja TERMOTE - STATISTICS BELGIUM

Keywords: LFS INDICATORS, PANEL DESIGN, MIXED EFFECTS LOGISTIC REGRESSION, 2-STEP CALIBRATION, BREAKS IN TIME SERIES

1. Quarterly samples: from cross-sectional to quarterly rotating panel survey

- *Cross-sectional survey*: independent quarterly samples[a]
  - …
  - **2016 T1** : S1 (1)
  - **2016 T2** : S2 (1)

- *Panel survey*[b] - Start-up phase (1)
  - **2016 T3** : [RG1 ∪ RG2 ∪ RG3] (1) = S3 (1)
  - **2016 T4** : ~~[RG1 ∪ RG2] (2) ∪~~ [RG4 ∪ RG5 ∪ RG6] (1) = S4 (1)

- *Panel survey*[c] - Start-up phase (2): quasi operational
  - **2017 T1** : RG1(3) ∪ RG2(3) ∪ RG6(2) ∪ RG7(1) = S1'
  - **2017 T2** : RG2(4) ∪ RG3(2) ∪ RG4(2) ∪ RG7(2) ∪ RG8(1) = S2'
  - **2017 T3** : RG3(3) ∪ RG4(3) ∪ RG5(2) ∪ RG8(2) ∪ RG9(1) = S3'
  - **2017 T4** : RG5(3) ∪ RG6(3) ∪ RG9(2) ∪ RG10(1) = S4'

- *Panel survey*[c] - Fully operational
  - **2018 T1** : RG6(4) ∪ RG7(3) ∪ RG10(2) ∪ RG11(1) = S1"
  - **2018 T2** : RG7(4) ∪ RG8(3) ∪ RG11(2) ∪ RG12(1) = S2"
  - **2018 T3** : RG8(4) ∪ RG9(3) ∪ RG12(2) ∪ RG13(1) = S3"
  - **2018 T4** : RG9(4) ∪ RG10(3) ∪ RG13(2) ∪ RG14(1) = S4"
  - **Etc**.

(a) 2-stage sampling: (1) SYS-PPS of PSUs (stratified and sorted frame), (2) SRS of HHs
(b) Each RG (Rotation Group) : by 2-stage sampling
(c) Notation: RGr(w) = respondent sample remaining after wave w from RG no. r; similarly for quarterly respondent samples Sq(w) etc.

2. Quarterly weighting up to *versus* after 2016 T4

Until 2016 T4	Since 2017 T1
< IND; *d*; Strat12 × Sex × Agecat; Lin > i.e. classical post-stratification	**< IND; *d* / $\tilde{\pi}$; Strat12 × Sex × Agecat + RG_c; Lin >** … more general calibration
o Calibrating **IND**ividual respondents o Calibrating to the joint population distribution of **Strat12**, **Sex** and **Agecat** ▪ **Strat12** : NUTS 2 region where **IND** is living (# 12) ▪ **Sex** : sex of **IND** (# 2) ▪ **Agecat** : age class of **IND** (0-4, 5-9, …, 70-74, 75+; # 16) Benchmarks / joint population distribution: from the *National Population Register*	
o Starting from sampling weights ***d***	o Starting from adjusted sampling weights ***d* / $\tilde{\pi}$**, using estimated response probabilities $\tilde{\pi}$
--	o Including contrast constraints **“ RG_c ”** between subsamples
o **Lin**ear method… or any other	o **Lin**ear method… as long as it works
o **Implicit assumption**: sufficient correction for nonresponse through this calibration	o **Explicit correction** for nonresponse and attrition (hopefully better!)
o **Usage**: for quarterly estimation of **CORE** and **STRUCT**ural variables	o **Usage**: for quarterly estimation of **CORE** variables **only**
--	↓↓ 2 INNOVATIONS to be explained ↓↓

2.a. Innovation I : explicit correction for nonresponse / attrition

→ Why ?

Recall, e.g. for **2017 T2**, complexity of the quarterly sample:

$$RG2(4) \cup RG3(2) \cup RG4(2) \cup RG7(2) \cup RG8(1) = S2'$$

- RG(s) in wave 1 : effect of “initial” nonresponse (NR) only
- RG(s) in wave 2 : effect of “initial” NR + attrition from wave 1
- RG(s) in wave 4 : effect of “initial” NR + attrition from wave 1 + attrition from wave 2 + attrition from wave 3

To be expected:

- **Different** response **levels** after different numbers of waves / for different RGs
  - Expected and observed: ~70% after wave 1, ~60% after wave 2, ~54% after wave 3, ~48% after wave 4
- **Different** response **mechanisms** at different stages
  - Initial NR : non-contact, refusal, disability, lack of interviewer, …
  - Attrition from wave *w* to *w*+1 : interviewer drop-out, respondent getting bored, …
- Availability of **predictors**
  - For initial sample: background characteristics from *National Population Register* (sex, age, HH composition, …)
  - For respondents: ILO status (StatBIT), level of education, … (→ *interesting, but not yet used*)

→ How ?

The approach:

- Modelling **per RG** *r*
- Modelling response **at household (HH) level**
  - o Working with interviewers…
  - o … contacting the HH, or its *reference person*
- Modelling **cumulative response**, i.e. response from initial sample to wave *w*, for RG*r**(*w*)
  - o A simplification…
  - o … limiting availability of predictors
- Modelling **differences between PSUs**
  - o Randomly selected PSUs (stage 1 in sampling)
  - o Between PSU differences not fully captured through predictor(s)
  - o Many-to-1 relationship between PSUs and INTerviewers

Notation:

- RG*r**(*w*) = entire initial sample or rotation group no. r, with indication of response or nonresponse after wave w

The model for any RG*r**(*w*): a ***random intercept logistic regression model***

- **Dependent** variable = cumulative response indicator for HH *j* in PSU *k* with distribution $y_{jk}|u_k \sim Bin(1, \pi_{jk})$
- **Linear predictor** $\eta_{jk} = \beta_0 + HHtype_j^T \beta_1 + Origin_j^T \beta_2 + Province_j^T \beta_3 + Urbanisation_j^T \beta_4 + u_k$
  - o ***Fixed effect*** variables
    - HH-type (# 5), Origin (# 3), Province (# 11), Degree of urbanization (# 3)
    - Available for all initially selected HHs
  - o ***Random effect*** u_k of PSU on intercept, with $(u_k) \sim N(0, \sigma_u^2 I)$
    - Measuring geographic variation in (non)response
    - Measuring variation between interviewers
- **Logit link** function: $\eta_{jk} = g(\pi_{jk}) = logit(\pi_{jk}) = \ln\left(\frac{\pi_{jk}}{1-\pi_{jk}}\right)$
- **…**
- **Smoothing** the estimates $\hat{\pi}_{jk}$

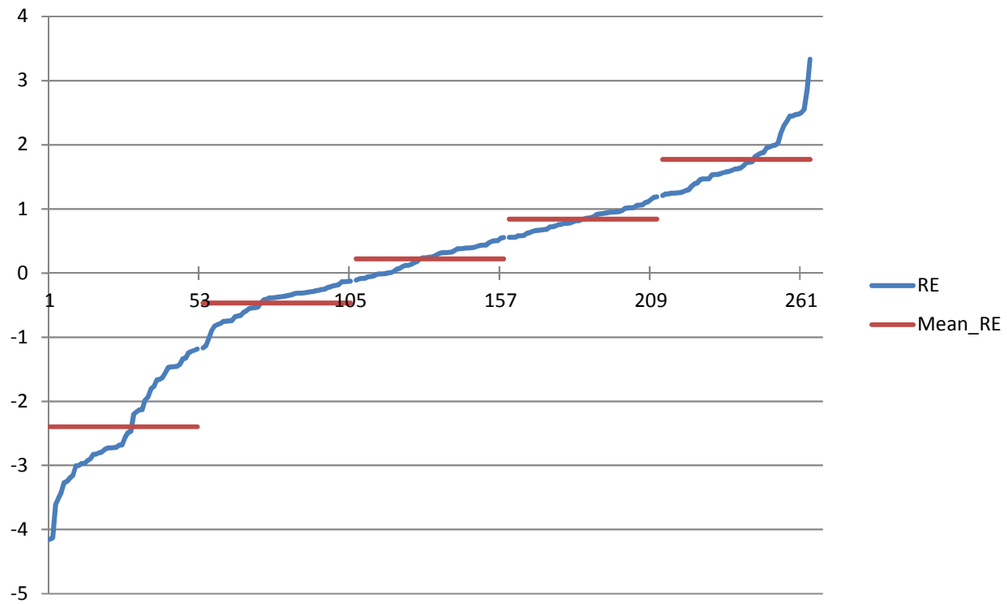
→ Quality? Prediction power, e.g. for RG*r**(*w*) in 2017 T4, before and after smoothing

RG *r*	Wave *w*	**AUC** (×) ~ predictive power of the model, *using*…		
		Original estimated probs. $\hat{\pi}_{jk} = g^{-1}(\hat{\eta}_{jk})$ $= g^{-1}(x_{jk}^T\hat{\beta} + \hat{\boldsymbol{u}}_{\boldsymbol{k}})$	Marginal probs. $\hat{\pi}_{jk}^0 = g^{-1}(x_{jk}^T\hat{\beta} + \mathbf{0})$	Smoothed probs. $\tilde{\pi}_{jk} = g^{-1}(x_{jk}^T\hat{\beta} + \hat{\boldsymbol{u}}_{\boldsymbol{Q(k)}})$
5	4	0.8445	0.6261 (*)	0.8368
6	3	0.8469	0.6279	0.8374
9	2	0.8158	0.6474	0.8059
10	1	0.8303	0.6228	0.8181

(×) AUC = area under the ROC curve

(*) 0.6291 for corresponding *fixed effects* model (RE $u_k \equiv 0$, or $\sigma_u^2 = 0$)

Smoothing the RE part of the linear predictor → **smoothed** $\tilde{\pi}_{jk}$



Horizontal axis: numbering of PSUs after sorting by $\hat{\boldsymbol{u}}_{\boldsymbol{k}}$

Vertical axis:

RE = $\hat{\boldsymbol{u}}_{\boldsymbol{k}}$ for $k = 1, \ldots, 286$

- Very variable $\hat{\boldsymbol{u}}_{\boldsymbol{k}}$
- Extreme $\hat{\boldsymbol{u}}_{\boldsymbol{k}}$ → extreme $\hat{\pi}_{jk}$ → extreme $d_{jk}/\hat{\pi}_{jk}$
- Effects on estimates in certain domains

Mean_RE = $\hat{\boldsymbol{u}}_{\boldsymbol{Q(k)}}$ for $k = 1, \ldots, 286$

- $\hat{\boldsymbol{u}}_{\boldsymbol{Q(k)}}$ = average of REs in $Q(k)$-th quintile
- Less extreme $\hat{\boldsymbol{u}}_{\boldsymbol{Q(k)}}$ → less extreme $d_{jk}/\tilde{\pi}_{jk}$
- Moderate effect on estimates

2.b. Innovation II : calibration with contrast constraints

→ Why and how ?

Recall, e.g. for **2017 T2**, complexity of the quarterly sample:

$$RG2(4) \cup RG3(2) \cup RG4(2) \cup RG7(2) \cup RG8(1) = S2'$$

- 3 “normal” RGs + 2 small RGs (result of splitting 1 normal RG)
- Same 2 stage sampling for each of these 5 RGs, whence:

RG	*2*	*3*	*4*	*7*	*8*
initial # HHs =	6,695	3,965	2,132	6,695	6,695
$\sum_{jk} d_{jk}$ =	*M*	*M*	*M*	*M*	*M*
initial # INDs ≈	14,729	8,723	4,690	14,729	14,729
$\sum_i {d_i}/{\tilde{\pi}_i}$ ≈	*N*	*N*	*N*	*N*	*N*

$\sum_{jk}$ denotes summation over responding HHs
$\sum_i$ denotes summation over responding INDs

- **Problem**: too much impact of RG3 and RG4 on quarterly estimates
- **Solution**: forcing the impact of each RG being proportional to its original size, i.e.

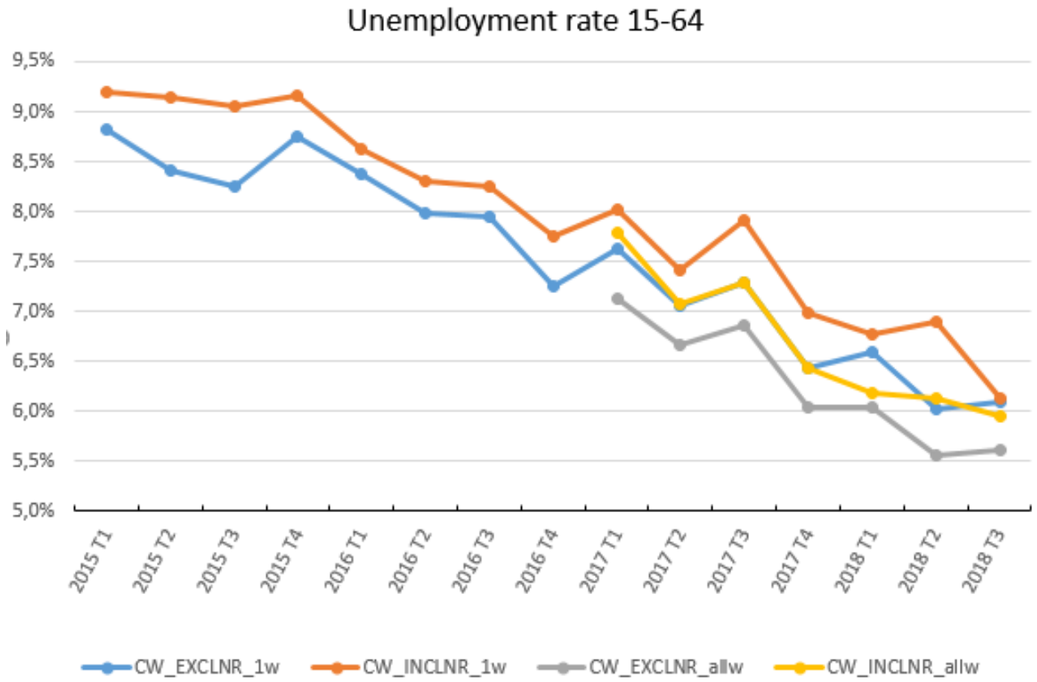
$$\frac{\sum_{i\in RG2} w_i}{6695} = \frac{\sum_{i\in RG3} w_i}{3965} = \frac{\sum_{i\in RG4} w_i}{2132} = \frac{\sum_{i\in RG7} w_i}{6695} = \frac{\sum_{i\in RG8} w_i}{6695}$$

→ → → Transforming into 4 contrast constraints ← ← ←

- RG2 <> RG3 : $\sum_{i\in RG3} w_i - \frac{3965}{6695}\sum_{i\in RG2} w_i = 0$
- RG2 <> RG4 : $\sum_{i\in RG4} w_i - \frac{2132}{6695}\sum_{i\in RG2} w_i = 0$
- RG2 <> RG7 : $\sum_{i\in RG7} w_i - \frac{6695}{6695}\sum_{i\in RG2} w_i = 0$
- RG2 <> RG8 : $\sum_{i\in RG8} w_i - \frac{6695}{6695}\sum_{i\in RG2} w_i = 0$

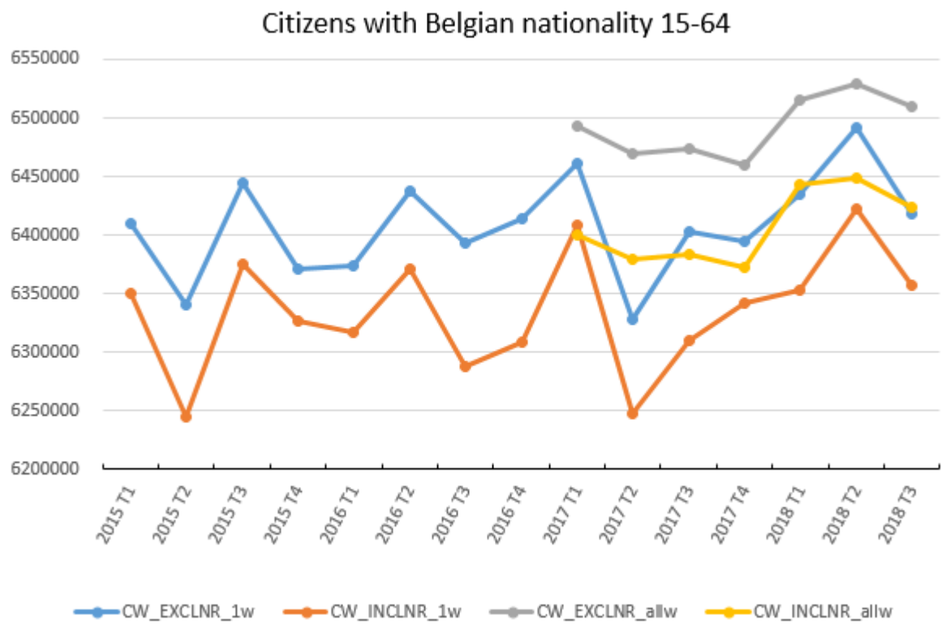
3. Result : effect on UR for age class 15-64

- Cross-sectional versus panel survey
- Old 1-step versus new 2-step calibration
- Legend:
  - estimation using calibrated weights excluding preliminary nonresponse correction, based on wave 1 data
  - estimation using calibrated weights including preliminary nonresponse correction, based on wave 1 data
  - estimation using calibrated weights excluding preliminary nonresponse correction, based on all data
  - estimation using calibrated weights including preliminary nonresponse correction, based on all data

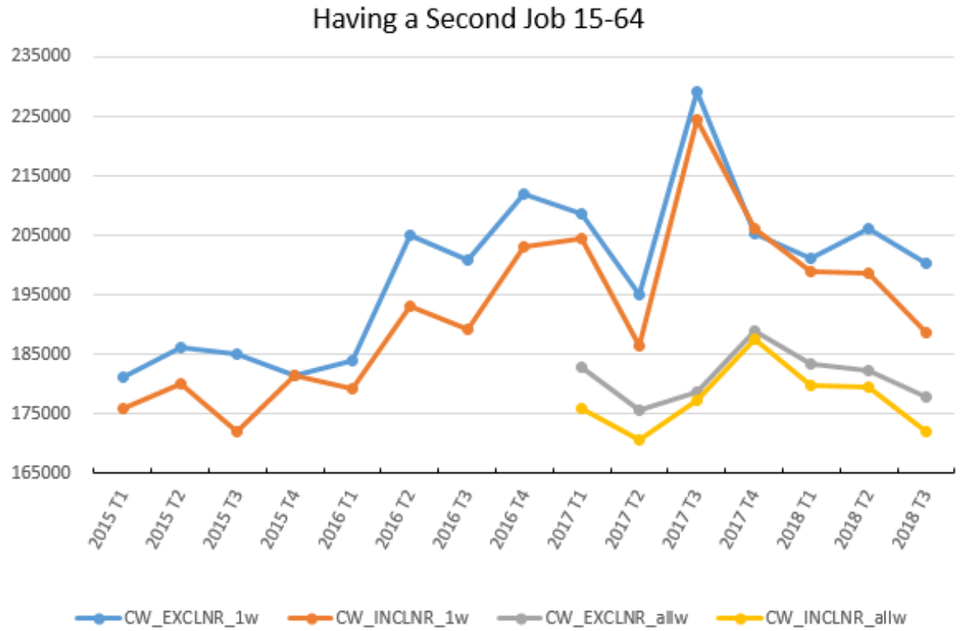


- Break not visible in time series for UR 15-64
  - o If (1) continued with cross-sectional survey (1 wave) and (2) **new weighting method** (2 step)
    - ➢ a break would have occurred: ● ↑ ●
  - o If (4) **switched to panel survey** (4 waves) but (5) continued with old weighting method (1 step)
    - ➢ a break would have occurred: ● ↓ ●
  - o Combination of (4) switching to panel survey (4 waves) and (2) new weighting method (2 step)
    - ➢ Compensation: ● ≈ ●

Some other results



- Compensation: ● ≈ ●
- Citizens without Be. nat. have lower response propensity
- NR correction is grossing up citizens without Be. nat. :
  ● ↓ ● and ● ↓ ● (no mode/wave effect)



- Significant mode/wave effect: ● ↓ ● and ● ↓ ●
  Significant break: ● ↓ ●
- No significant impact of explicit NR correction
- Problem: **underreporting** second job in follow-up waves

Read more in : **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 at** https://statbel.fgov.be/nl/over-statbel/methodologie/analyses**)**

https://statbel.fgov.be