Katja Rutar, Statitical Office of Slovenia

An example of longitudinal Labour force survey weights

When in 1997 the Slovenian Labour Force Survey (LFS) was transformed from an annual cross-sectional survey to a quarterly longitudinal survey, a rotation pattern with five waves of data collection was selected, in three consecutive quarters, then one quarter of brake and finally two more quarters of data collection (3-1-2). We copied the Finnish LFS rotation pattern. This is a compromise solution which – in theory – provides 3/5 of overlapping in two consecutive quarters and 2/5 of overlapping units between corresponding quarters in two consecutive years. (Tables with numbers can be found in the presentation.)

Toward a longitudinal weighting scheme for two-quarter data sets

With the review of different literature describing longitudinal weights in longitudinal surveys we saw that different organizations emphasize very different details of the weighting process. One of the large issues is the bias arising from the sample attrition. This is a big problem already by the cross-sectional estimates, because from the initial sample of about 2,400 households after five waves of data collection we get only about 800 (1/3) responding households. The second big problem is the definition of the universe or target population in two consecutive time periods, because such demographic data are not published. This problem is smaller in quarterly comparisons and bigger in annual comparisons, because in a one-year period more people change households than in three months. An important issue is also coherence between cross-sectional and longitudinal estimates as we have two different estimates in two different points of time. A practical problem is also the number of final weights in the final datasets. Several different weighting factors produce confusion, complexity and cost for producing and using weights for longitudinal analysis. We will deal with each of the mentioned issues in the following chapters.

When we started thinking about longitudinal LFS weights we first looked at EU-SILC longitudinal weights (on year to year basis) which are defined by regulation and have already been used for several years. Their strategy is to make every rotation group in every year representative for the total population. But this is reasonable for sub-samples which are – in the case of Slovenia – more than two times larger than LFS sub-samples. For the sub-sample of 804 households (5th wave in the 4th quarter of 2013) weighting (calibration) at NUTS 3 level including gender and eight age classes this is not possible. So (first stage) nonresponse weighting will be done for each wave separately while (final stage) calibration will be done on all three overlapping rotation groups together.

Design weights (reciprocal of the probability of selection) for each rotation group are calculated when the group is included in the sample for the first time and they are then used in all subsequent data collection waves. So we will consider the same strategy in calculating longitudinal weights. As a basis we will use cross-sectional design weights.

Cross-sectional nonresponse weights also have to be considered, because, they correct the data also for the longitudinal attrition. But for the longitudinal estimates we will calculate additional longitudinal nonresponse weights at the level of strata and wave in the same way as cross-sectional nonresponse weights are calculated: $\frac{\text{responing households} + \text{nonresponding households}}{\text{responding households}}$.

As there are no auxiliary demographic data on longitudinal population in two consecutive periods of time, most studies about longitudinal weights, which we found, decided on the so-called cohort analysis, based on population at the beginning of the corresponding reference period. Big disadvantage of the cohort longitudinal analysis is late availability of longitudinal data – such longitudinal weights can be calculated only when the data for the following wave are collected and weighted. As rotation groups are drawing design weights from the first wave and nonresponse weights are cumulating by waves, in practice, there are no big differences between final weighting factors before calibration in two consecutive data collection waves. Pearson's Correlation Coefficient between those weights in overlapping population from 3rd and 4th quarters of 2013 is 0.96 (n=8,534 individuals). (For information – correlation of final weights for the same population is 0.89, so most of the difference comes from calibration.)

After longitudinal nonresponse adjustment weights of longitudinal subsample were grossed up to the population total at the beginning of the 3rd quarter (this is sometimes called base longitudinal weight). The second steep was standard calibration (used also as the last step in the production of cross-sectional weights) at the level of households to gender – (8) age classes distribution by NUTS 3 regions. An additional step, special for calculation of longitudinal weights, was calibration of the longitudinal subsample to main employment status estimates (unemployed, employed, inactive) at the beginning quarter, by gender and NUTS 2 regions. (An experiment with NUTS 3 regions was done but the variability of weights was too high.) The last step destroyed gender–age classes distribution, so we repeated calibration to demographic totals once again. And as the last steep we repeated calibration to main employment statuses. When we checked final longitudinal weights after this step, they converged to total population demographic characteristics and employment statuses estimates for the 3rd quarter. These are **basic descriptive statistics for the longitudinal weight**, compared to standard quarterly weight for 3rd quarter 2013:

	n	mean	stddev	min	q1	q3	max	population
W_LONG13q3	8.534	241,28	165,47	14,28	139,87	290,16	1.675,99	2.059.114
W_CROSS13q3	15.486	132,97	90,25	10,79	76,44	159,86	1.100,41	2.059.114

In the selected case the longitudinal sample represents around half of the quarterly cross-sectional sample. As a consequence the longitudinal weight is almost twice as big as the cross-sectional weight and the same is with its standard deviation. Consequently also relative standard errors for longitudinal estimates will be considerably higher than for quarterly cross-sectional estimates. The maximal longitudinal weight is also extremely high but detailed analysis shows that 99% of longitudinal weights are lower than 1,000.

Here we present results only for **changes in main (employment) statuses for the cohort** interviewed in the 3rd and 4th quarters of 2013 (change in one quarter), in 1,000 and in %:

2013Q3 / 2013Q4	**Unemployed**		**Employed**		**Inactive**		**Under 15**		**Total**	
	in 1000	%	**in 1000**	%	**in 1000**	%	**in 1000**	**%**	**Long. in 1000**	**Crossec in 1000**
Unemployed	**48**	50	**22**	23	**26**	27			**96**	**96**
Employed	**13**	1	**865**	94	**45**	5			**922**	**922**
Inactive	**21**	3	**45**	6	**677**	91			**743**	**741**
Under 15					**4**	1	**294**	99	**298**	**299**
Total – longitude.	**82**		**932**		**752**		**294**			
Total – crosssec.	**97**		**910**		**753**		**299**			**2.059**

Obvious is the incoherence between longitudinal and crosssectional estimates of employment statuses in the 4th quarter which we didn't use in the calibration process.

Precision of longitudinal weights for two-quarter data sets

Let's also look at the **relative sampling errors** for selected estimates of changes. As longitudinal weights in this exsample are calibrated to crosssectional totals at the beginning quarter, also variance was estimated as simple crossectional variance for all combinations of employment flows. We see that the majority (6) of estimates could be published without any warning. For changes between Unemployment to Employment and Inactive population or from Employment to Unemployment or from Inactivity to Unemployment a sign for less precise estimate should be added (10 % < cv < 30 %). Regarding Population under 15 years only the number of this population remaining in the same category could be published.

2013Q3 / 2013Q4	Unemployed	Employed	Inactive	Population under 15
CV	%	%	%	%
Unemployed	**6**	12	10	
Employed	15	**1**	**9**	
Inactive	13	**8**	**1**	
Population under 15			23	**2**

When we calculated relative standard errors for the same table separately for men and women, all estimates are less precise (cv > 10 %). The longitudinal sub-sample seems to be too small for a more detailed analysis of changes.

Conclusions & Open questions

There seem to be really many possibilities to calculate longitudinal weights. At each separate stage of weighting process some assumptions have to be taken. Even if we can calculate longitudinal population, it doesn't seem very useful for the purpose of longitudinal weights. Would it be appropriate to calibrate to employment status distribution in both quarters? Is it more reasonable (because of timelines) to adapt the data to the last quarter (without using nonresponse weighting)? Longitudinal subsamples in countries with small population are very small. They are subject to high (selective) attrition. But users are interested in flow estimates.

Literature

Boschetto, B., Discenza, A., Fiori, F., Lucarelli, C. and Rosati, S. (2010): Longitudinal data from Italian Labour Force Survey, presented on European Conference on Quality in Official Statistics.

Jenkins, J. and Chandler, M. (2010): Labour market gross flows data from the Labour Force Survey, Economic & Labour Market REview, Vol 4, No 2, Office for National Statistics.