

WORKSHOP ON LABOUR FORCE SURVEY METHODOLOGY

7 th WORKSHOP ON LABOUR FORCE SURVEY METHODOLOGY

DATA PROCESSING AND DATA QUALITY

Madrid, Spain, 10 – 11 May 2012

G. Data quality: Quality issues on monthly figures, timeliness and consistency between quarterly and monthly data

G2 – Monthly Unemployment Figures – Quality Issues

Daniela Gumprecht – Austria

Preliminary Notes and Background

Since reference period January 2011 Austria publishes monthly unemployment figures according to international definitions. Data stem solely from the Labour Force Survey. Monthly figures are based on an adopted weighting scheme derived from the standard weighting scheme for quarterly LFS data. This procedure allows computing flash as well as final monthly unemployment estimates for the whole population as well as for some subgroups. Flash estimates are available in time to be used for Eurostat's harmonized unemployment statistics. Eurostat makes the following demands on monthly unemployment figures:

- Timeliness: data have to be transmitted 3 working days (72 hours) before the publication of the press release at the latest.
- Scope of delivery: numbers of non-adjusted and seasonally adjusted employed and unemployed persons aged 15 to 74 by sex and two age groups. All in all four types of figures for four subgroups.
- Consistency: between monthly and quarterly data and between sub- and super-groups.
- Revision degree: If the revisions of previous monthly values are too big, Eurostat does not publish them anymore.

Consistency:

Consistency between sub- and super-groups means all super-groups should be cumulations of adequate subgroups of the lowest level (i.e. sex and age group). Austria follows a hierarchical approach, which means all calculations and adjustments are done on the deepest level, i.e. the four smallest subgroups young women, young men, old women, old men. All super-groups are built by cumulating the corresponding subgroups. This approach leads to consistency between sub- and super-groups.

Consistency between final monthly and quarterly values should be achieved; that is the mean of monthly unadjusted levels should be exactly the same as the unadjusted quarterly levels. This is generated by the following simple procedure. As soon as quarterly LFS data are available, preliminary monthly estimates of this quarter are adopted by multiplying monthly numbers of employment and unemployment by the factor q / \bar{m} , where q denotes the quarterly value and \bar{m} the mean of the months.

Time Series and Seasonality:

Eurostat's harmonized unemployment figures are seasonally adjusted (s.a.), and adjustment should be done by TRAMO/SEATS or ARIMA X12. For the calculation of seasonally adjusted monthly unemployment rates Eurostat recommends an indirect way. Employment and unemployment levels of the smallest subgroups, i.e. young and old men and women, should be adjusted. Adjusted subgroups are cumulated to adjusted super-groups. Rates of all different sub- and super-groups are estimated using the corresponding adjusted levels. If there is no seasonality inherent, there is no reason for seasonal adjustment. Original series can be used as seasonally adjusted ones.

Austrian time series show rather different seasonal patterns, see Figure 1. Series of employment and total labour force show clear seasonal patterns whereas unemployment series show no clear regular pattern, some subgroups are seasonal, some are not, and most of them are highly erratic and volatile. The simplest case of using original series as s.a. ones is not really satisfying due to the fact that series are highly erratic. As some kind of compromise between classical seasonal adjustment and no adjustment at all, it was agreed with Eurostat to

estimate the trend values and use them as adjusted values. However, trend calculation does not solve the problem of really small sample sizes in some subgroups, e.g. in April 2010 there were only 29 unemployed women at the age of 15 to 24 interviewed for the microcensus. It is problematic to publish monthly unemployment estimates for young men and young women separately. Therefore these values are flagged as “unreliable/ uncertain data” in the Eurostat database. If users are interested in further details, they have to consult quarterly LFS results.

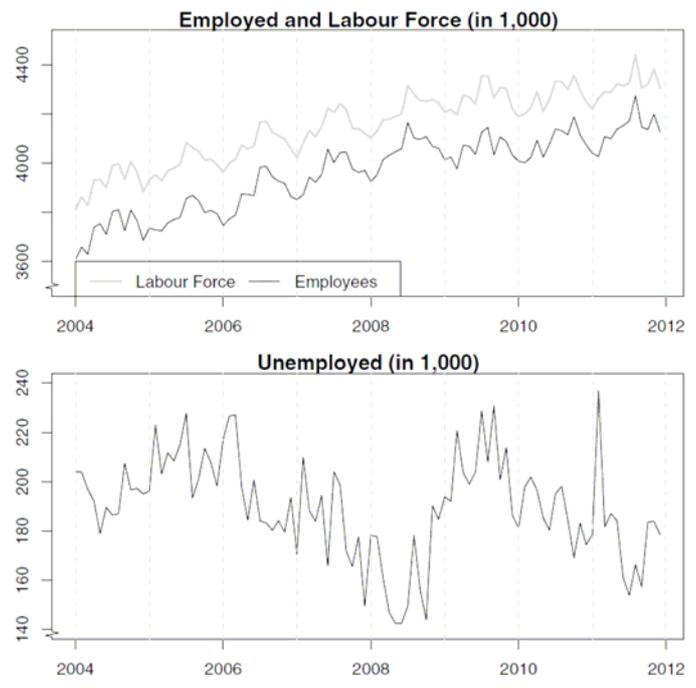


Figure 1: Austrian Labour Market, Monthly Totals since 2004.

Evaluation of Flash Estimates, Late-Response Analysis

Besides fulfilling the standards from Eurostat, there are other important quality aspects to be aware of whenever flash estimates are used. The difference between flash and final estimates of unadjusted and trend values is one of them. In this context sample sizes of subgroups, early response rates, late-response bias etc can be interesting. Austrian flash monthly estimates are based on approximately 90% of all monthly data. A late response bias is corrected via weighting, just as in case of non-response. To evaluate the reliability of flash estimates various indicators, partly discussed by LAMAS, can be used.

- Average difference in percentage points between final and flash estimates.
- Average absolute difference in percentage points between final and flash estimates.
- Maximum absolute difference between flash and final estimates.
- Percentage of correct direction (up or down) of provisional month-to-month changes.

Another important issue is the behaviour of the adjusted series, especially the behaviour when new values enter the calculations and the behaviour when flash estimates are replaced by final estimates which are available together with final quarterly LFS data. Very interesting in this context is the behaviour of turning points.

Figure 2 shows how long it takes for the trend to capture shifts in directions. The black lines show the trend lines calculated at different points in time, i.e. simulating the production process for monthly unemployment figures since January 2009. Each trend for each month is shown by its own black line. The bold gray line shows the trend estimation based on the whole series from January 2004 until January 2012, i.e. the best, in the sense of most current and ‘all inclusive’, trend series.

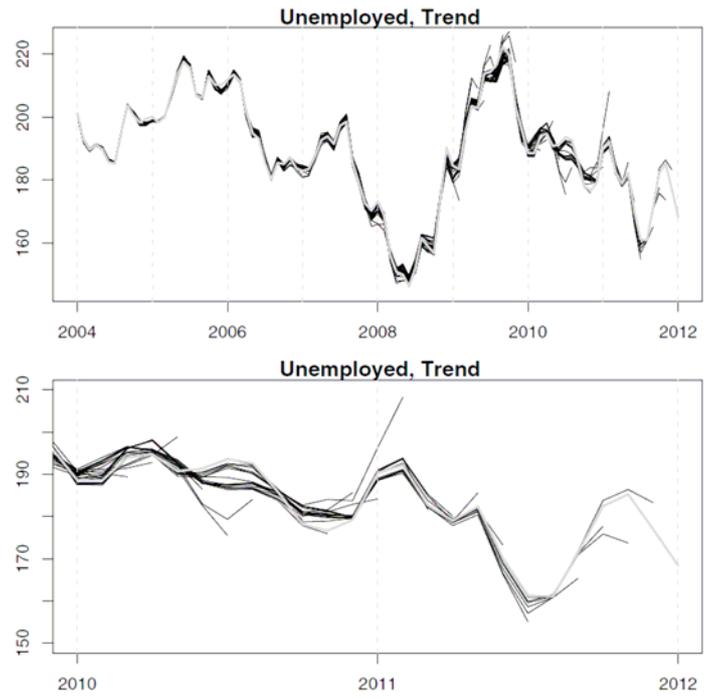


Figure 2: Total Unemployment, Trend Revisions.

References:

Gumprecht, D., Haslinger, A., & Kowarik, A. (2011). Austrian LFS Monthly Unemployment Rates. *Austrian Journal of Statistics*, 40, 297-313.