# *Seminar on Registers in Statistics - methodology and quality 21 - 23 May, 2007 Helsinki*

## *Using Register information to estimate (early) monthly unemployment rates for EU aggregates*

*Carsten Olsson*[1]
*Eurostat F2 "Labour Market Statistics"*
*Carsten.olsson@ec.europa.eu*

# *1 Introduction*

The problem discussed in this paper is, in a nutshell, that registers may provide perfect information on the wrong population while statistical surveys may provide incomplete information on the right population[2]. To complete the picture, other factors such as timeliness and revisions must also be taken into account when deciding on the best approach in a given context. The current solution for the estimation of EU unemployment rates is currently a balanced approach where information from registers is benchmarked against survey results. In the future, it is possible that the approach will lean towards more use of survey information (that may integrate information from registers) and less direct use of register information.

The first section outlines the general aims and the current approach used by Eurostat to compile monthly unemployment data. Then, register data and survey data are compared as sources, and, finally, possible future developments are sketched out.

# *2 Compiling monthly EU unemployment rates*

One of the most accessed Eurostat indicators is the monthly unemployment. It measures the number of unemployed persons in the EU and it is presented as levels and rates as a percentage of the total active population. Several breakdowns are available: by country of residence, by sex, and by two age groups.

Although this indicator is already well established, it is by no means perfect. In order to evaluate the necessity for (and usefulness of) compiling this indicator and to clarify future developments, it has to be viewed in a wider context.

## *2.1 Aim*

The aim is to compile a timely, comparable, and reliable indicator of unemployment rates.

Historically, data on monthly unemployment levels have been available from Employment Agencies for a long time in most countries. These data are not necessarily comparable across countries, which make it difficult to use them directly for coordinating common EU policies, e.g. the development of a coordinated strategy on employment as outlined in Title VIII of the EC Treaty[i].

As mentioned in [ii], Eurostat undertook several steps in the 1990'ies to further improve labour market statistics in order to support the tasks assigned to the Community.

---

[1] The usual disclaimer applies. The opinions expressed herein are those of the author and do not necessarily represent those of Eurostat. The author would nevertheless thank his colleagues, including Remko Hijman, Omar Hardarson, Joachim Recktenwald, Africa Melis, and Veijo Ritola, for their feedback on the early drafts of this paper.

[2] The words 'right/wrong population' is used for illustrative purposes. Here, it covers the idea that aims and concepts may be different.

The EU Labour Force Survey (LFS) had already been introduced in 1983. It was carried out in spring (quarter 1 or 2 depending on the country). The framework Council Regulation defining the Continuous Labour Force Survey was adopted in 1998. Since then the transition to the continuous quarterly survey (where the reference weeks are spread uniformly throughout the year) has been gradually conducted by Member States. By 2005, all Member States are carrying out a continuous labour market survey and all (but one) producing quarterly LFS results.

However, comparability and quality may not always be sufficient by themselves. The "Communication of the Commission to the European Parliament and the Council on eurozone statistics"[iii] underlined the need for timely data of sufficient frequency for the short-term economic indicators:

> "The Barcelona European Council of March 2002 expressed the need for further progress to be made on Eurozone statistics […].
>
> Indeed, the availability of good quality Eurozone [and EU] statistics is essential for the co-ordination of economic policies […]. This applies to infra-annual macro-economic statistics, which represent a key instrument for business cycle analysis and short-term economic policies.
>
> […] A sufficiently detailed and comprehensive set of timely and reliable monthly and quarterly statistics for the EU and the Eurozone is indispensable […]."

The LFS already delivers comparable high quality information on – among other things – the labour status, but it only requires the production of quarterly results, which are transmitted to Eurostat with a delay of 84 days (12 weeks).

The latest report on information requirements in EMU[iv] sets the requirements for an indicator of short-term unemployment rates to monthly frequency with a delay of no more than 30 days after the end of the reference month. Thus, the LFS needs to be amended or supplemented with additional information in order to produce timely monthly unemployment rates.

The monthly unemployment rates published by Eurostat since the mid-eighties attempt to meet the needs.

## 2.2 Current method for compiling monthly unemployment rates

Eurostat is currently making use of a combination of sources for the compilation of the harmonised monthly unemployment rate.

The basic building blocks for compiling harmonised monthly unemployment rates are the continuous LFS, which is the level reference for the Eurostat estimations of quarterly unemployment rates, supplemented by monthly information on the number of people registered with national Employment Agencies (short-term evolution).

Currently, Eurostat combines these two sources at the level of Member States by taking the best from both of them. For a large number of countries, LFS results are available for four years[3] or more. For most of these countries, the "standard" approach is applied. In this approach, the LFS data are used for setting the quarterly levels and the Employment Agency data for setting the infra-quarterly development. For the most recent months, where LFS may not yet be available, it is assumed that the relationship between the LFS and the Employment Agency data is similar to the most recent past[4].

For the remaining countries with LFS series of less than four years length, the unemployment figures are benchmarked against a moving average of LFS data. This means that the short-term development is

[3] It is generally agreed that seasonal adjustment procedures produce acceptable results when at least four years of data are available. In order to avoid breaks in the series that are to be seasonal adjusted, the benchmarking procedures are different for series where the underlying LFS series is shorter or longer than four years.

[4] Several variations of this procedure exist depending on the length of the quarterly LFS series and the nature of the relationship between them.

dominated by the Employment Agency data series while the long-term development is still closely linked to the LFS series.

Finally, a small number of countries provide 3-months moving averages of LFS data as an approximation for monthly data and some few countries produce monthly figures from the LFS.

By aggregating the estimates for the individual Member States, is possible to produce timely estimates of monthly unemployment rates for the euro-area and the EU total, which are (asymptotically) based on harmonized definitions.

## 2.3 Strengths and deficiencies of current method

The strengths of the current method for compiling monthly unemployment rates are that it is available with an acceptable delay, i.e. around 30 days after the end of the reference month, and it is benchmarked against the LFS, which is well defined and comparable across countries.

The deficiencies are primarily that

- infra-quarterly developments are non-harmonised;
- there is an important delay between the publication of the current month's unemployment rate and time when the benchmark results become available; and
- for those countries that provide moving averages, the time period covered is not equal to the reference period.

In practice, this means that the estimation of the current month's unemployment rate, which for most users is the most important one, may be revised when the LFS results become available three-six months later. It also means that turning points may be detected too late or shifted in time.

# 3 Comparison of sources for unemployment levels

As already mentioned, the two most common sources for measuring unemployment levels are the LFS and data from Employment Agencies.

Employment Agencies exist in almost all countries and their role is normally (but not exclusively) to ensure that employers and potential employees are matched and to distribute financial compensation to people who are temporarily out of work. The Employment Agencies are able to compile time series of the level of unemployed based on their registers. The definitions used are evidently resulting from national register practices and legislation. Thus, comparability across countries is an issue.

The other alternative set of indicators is the data from the European Labour Force Survey (LFS). The LFS is carried out in all EU Member States using common definitions and approach. The LFS is available at EU level with a quarterly frequency.

## 3.1 Definition differences

The existing LFS is designed to provide comparable indicators of good quality. This is highlighted in the preamble of the framework Council Regulation defining the LFS[v], which says

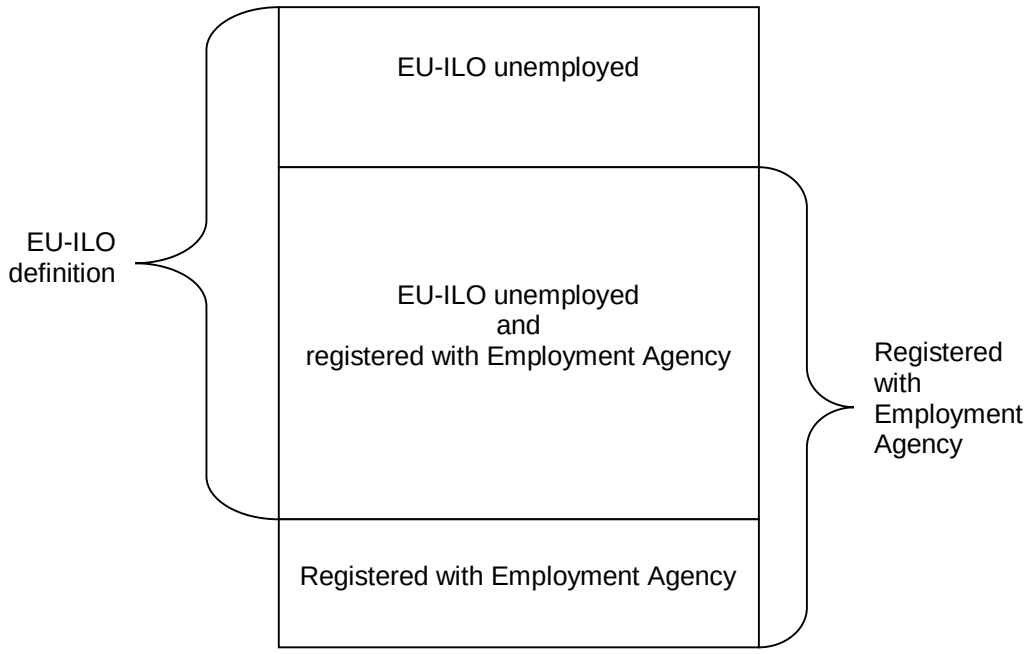
> "… in order to carry out the tasks assigned to it, the Commission needs comparable statistical information on the level and pattern of and trends in employment and unemployment in the Member States".

The LFS is as far as possible based on internationally agreed statistical concepts and definitions. For example, the definition used to describe the employment status[vi] (employed, unemployed, inactive) is derived from the corresponding ILO definition. It says, slightly simplified, that a person is recorded as unemployed if he/she has been without work during the reference week, has been available for work, and is

actively seeking work. The LFS covers the resident population on the economic territory of each Member State and may include institutional households if possible. The LFS is carried out as a continuous survey, i.e. each week of the quarter is covered. The survey method is not defined but interviews have to be with at least one person of the household, who may answer on behalf of all members of the household. The principles that define the formulation of questions on labour status are laid down in the Regulation. The sample size is specified in relation to the relative error in a quarter and the change between successive quarters.

Employment Agencies also compile their statistics on number of job seekers following a strict set of rules. However, these rules are not harmonised since - as is obvious - the purpose of the Employment Agencies is generally to implement the national policy of matching the needs of employers and potential employees as well as distribute financial compensation to people who are temporarily out of work.

There is frequently a good numerical correspondence between the two sources. Some of this is due to a common core of people both being registered with the Employment Agencies and at the same time satisfying the EU-ILO definition of unemployment. In addition to the common core, one group of people only satisfies the EU-ILO definition while another group only satisfy the Employment Agency definition. If these two groups are numerically similar in a given period, the total number of unemployed persons reported by the LFS and by the Employment Agencies will also be similar, but that does not necessarily mean that the short-term (or long-term) developments will be identical.




## *3.2 Definition differences in practice*

Since Registers and the reference survey (LFS) are designed with different purposes in mind (at least in the details), there are clearly a number of areas where this will show up in practice.

It is well known that stickiness is a problem when using registers of unemployed. Stickiness is the fact that people become registered as being unemployed as soon as they have the (financial) need but may not necessarily unregister at all because there are no obvious benefits or needs to unregister as soon as they have a job. This creates some stickiness in most registers on unemployed.

People will mainly register in employment agencies if they have an incentive. The incentive is usually that the agency will actually help people finding a job. However, for large population groups, the agencies are not necessarily the most efficient mean to find a job and other methods may be just as useful or effective.

Another reason to register is that unemployment benefits or other financial compensation is linked to the registration. But for those parts of the population that may be excluded from such benefits (e.g. students or better-off people) there are no such incentives to register. On the other hand, people may be registered in order to be able to get access to those financial benefits although they have no immediate need for a job or are unable to start working for health reasons.

It is also possible that a person that has a temporary job or a job, which is too small to support him, may be registered with the Employment Agency (because he is looking for a better job or get additional financing) while the same person considered as being employed in the ILO sense.

This shows that the registers are not necessarily representative for the full population (in the ILO sense) or there may be delays in time needed to react on changes in the labour market. Furthermore, changes in the rules governing the Employment Agencies may lead to changes in the number of people registered as being unemployed without a corresponding change in the number of unemployed persons in the ILO sense.

Surveys are not without problems either although they may show up in other areas. In particular, the sample of a survey only covers a small part of the population and the methods used to gross up depend on whether the selection of households is representative, whether answers are accurate, whether non-response can be handled correctly, etc.. The result may be that the estimates have a high margin of error or – in the worst case – the results are biased.

## 3.3 Which one to trust the most?

Both sources deliver reliable information but it is only possible to compare them up to the level of definition differences since the aims are different.

Generally, Eurostat assumes that the labour force surveys deliver the most correct picture of the employment status across EU countries because it is carried out using common methodology and aims.

Empirically, the long-term developments in registers and LFS are similar for almost all countries, although levels typically are different. The short- and medium-term developments may be different, which reflects differences between the sources in aims, definitions, and methodology.

## 3.4 Summary

The table below shows the strengths and weaknesses of these two main sources, and the last column summarises the effect of combining the two main sources (the monthly Eurostat indicator). The sources are compared using the ESS quality definitions[vii] as a point of reference.

The main results are that

- Employment Agency data are heterogeneous and may have significant bias (compared to EU needs) because their figures are the result of national register practices and legislation, while the LFS data are comparable across countries but subject to sampling error, and – in the worst case - measurement errors.
- LFS data are to a large extent coherent in concepts with other official statistics, for example with respect to coverage of population and reference periods.
- Employment Agency data has the desired frequency and delays, and long series are usually available

The monthly unemployment rate produced by Eurostat combines most of the LFS strengths with the strengths of the Employment Agency data. The most important weaknesses are, as already mentioned, linked to the estimation of the most recent data and to the infra-quarterly evolution of the unemployment rates.

| | *Employment Agencies data* | *LFS* | *Combined monthly indicator* |
|---|---|---|---|
| Relevance – compared to EU needs | Weak – satisfies specifically national needs | Strong, designed for international comparison | Strong as benchmarked against LFS |
| Comparability | Weak | Strong | Almost strong |
| Accessibility and clarity<br>- simple to use, break downs | Medium | Good | Good |
| Accuracy – Levels and rates | Significant bias (compared to EU ILO-definition), low variance | Insignificant bias, some variance | Medium |
| Accuracy – Frequency | Strong - monthly | Weak - quarterly | Strong – monthly |
| Timeliness – Delay | Strong – delays between some few days and one month | Weak – almost three month delay | Better – delays almost as Employment Agencies, for EU total about one month |
| Coherence with other official statistics | Weak | Strong | Almost strong |
| Completeness - coverage of reference period | Mixed – Most Employment Agencies refer to specific day of month | Strong – represent average of quarter | Almost strong – mix of the two |
| Completeness - coverage of population | Weaker – willingness to register not always linked to employment status | Strong – represent average of population | Almost strong – mix of the two |
| Completeness - Length of series | Strong | Mostly weak | Better – mix of the two |
| Costs | Strong – low | Weak - expensive | Medium |

# *4 Can we overcome definitions differences?*

The subject of the seminar is the use of registers in official statistics. Eurostat is already making good use of information from registers to improve the compilation of harmonised monthly unemployment rates. The only way to improve the quality further would be if registers could provide additional information needed for the official statistics purposes. This is unlikely for a number of reasons.

First of all, there is a political need for both Employment Agency registers as well as for harmonised unemployment rates. Employment Agencies will always need to produce statistics that could highlight the effect of specific political interventions in the labour market as well as for internal management needs. Likewise, there is a political need for the harmonised unemployment rates in order to compare and evaluate the labour market across EU Member States.

Secondly, it is usually a direct or indirect requirement that persons are registered with an Employment Agency in order to receive unemployment benefits or other social assistance. This does not necessarily mean that they are actively seeking work. Similarly, persons that are seeking a job may not necessarily be registered with an Employment Agency.

Because of these inherent differences and because the willingness to register with an Employment Agency may depend on other factors than the labour status, it is excluded that the harmonised unemployment levels can be derived directly from the registers of Employment Agencies.

# 5 What then?

## 5.1 How can we and how should we compile harmonised monthly unemployment rates?

The difficulties outlined so far point at the delays and the frequency of the quarterly LFS as the main problems for producing high quality monthly unemployment rates.

The estimation methods currently used by Eurostat could probably be improved but not beyond the level of information actually available in the basic sources. The other option is to improve the LFS with respect to both delays and frequency.

To reduce the delays, it may require important administrative changes in the way the LFS is organised in most countries. Several Member States are undertaking studies on how to speed up the data collection and estimation procedures. Due to the nature of such large surveys, it may not be possible to achieve 100% coverage with short delays but it may be possible to achieve significant cuts in the delays[5].

To improve the frequency, it may be needed to review the sample design. In many Member States, the current sample design is constructed so that reliable quarterly results can be achieved but not necessarily reliable monthly results. The change of the sample design would again have an impact on the data collection procedures used by the LFS, but several Member States have already shown that it is possible and feasible to implement such changes. Does that mean that there will be no need for the use of administrative data in the future? The answer is both yes and no. At the level of Eurostat, substantial quality improvements could only be achieved if more Member States would be able to produce LFS data with short delays and monthly frequency. At the level of Member States, it is likely that a better usage of register information would allow the production of monthly unemployment levels (and rates) of a sufficient degree of precision without increasing the sample size significantly[6]. This would then subsequently enable Eurostat to produce unemployment figures with a monthly frequency from the LFS, without using the registers themselves.

## 5.2 Publication

One outstanding issue is related to the coherence between data for EU aggregates and for individual Member States. Monthly LFS results will evidently have less precision (higher variance) than quarterly results. It could be expected that the precision would be very good for EU aggregates but not necessarily for individual Member States. In that case, one option to consider would be to publish the estimated EU aggregates directly from the monthly Member State results but to smooth (moving average) the published results for individual Member States. The reaction from users should be positive since the published data will all have an acceptable reliability but some users may have less positive reactions because the direct link between the published Member State' data and EU aggregates is not transparent. It will be a challenge to solve this problem.

---

[5] For example, Spain, the Netherlands, Finland, and Sweden have already shown that it is possible to organise the LFS data collection so that LFS results can be delivered less than 30 days after the end of the reference period.

[6] For example, the sample selection in the Danish LFS takes account of information already available in registers: Persons which were recorded as being unemployed in a given period have a higher probability of being selected for the survey. Thereby the stratification becomes more efficient.

# 6 Concluding remarks

In this paper, it has been shown how Eurostat makes use of information from Employment Agencies to reduce delays and improve the frequency of the EU Labour Force Survey in order to produce harmonised monthly unemployment rates for individual EU Member States and for EU aggregates.

While the estimation methods used by Eurostat probably could be improved, they cannot be improved beyond the level of information actually available in the source data.

The only way whereby substantial improvements of the harmonised monthly unemployment rates could be achieved, would be if Member States made important changes to their national LFS. The first steps would be to speed up the data collection and compilation procedures and to change (if necessary) the sampling design so that the LFS becomes representative on a monthly basis. The next steps would be to ensure that the sampling variance is reduced sufficiently (e.g. by making use of information from registers for the sample selection) so that reliable monthly unemployment rates could be published for individual Member States as well as for EU aggregates.

---

[i] "Treaty Establishing the European Community", http://www.cc.cec/home/treaties/index_en.html

[ii] Melis, Africa – "Methodological experience with the labour market policy data collection", Seminar on registers in Statistics – methodology and quality, 21-23 May 2007, Helsinki

[iii] "Towards improved methodologies for eurozone statistics and indicators", COM(2002) 661 final, see http://www.cc.cec/sg_vista/cgi-bin/repository/getdoc/COMM_PDF_COM_2002_0661_F_EN_ACTE.pdf

[iv] "2005 Status report on Information Requirements in EMU by the EFC", see e.g. "Statistical requirements in stage three of EMU" at http://www.consilium.europa.eu/docCenter.asp?lang=en&cmsid=245

[v] Council Regulation (EC) No 577/98 of 9 March 1998 on the organisation of a labour force sample survey in the Community (OJ No L 77/3)

[vi] Commission Regulation (EC) No 1897/2000 concerning the operational definition of unemployment

[vii] "Definition of Quality in Statistics", see http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/DEFINITION_2.PDF