November 2024

# Online Job Advertisement Rate: methodological note

As was highlighted at the Summit of Social Partners that took place on 31 January 2024 in Brussels-Val Duchesse, there is a need for skilled workers and for quality jobs to boost EU's competitiveness. Labour and skills shortages hold back business activities and limit the potential for EU growth.

In this context, it is important to provide policy makers and social partners, for the European Union as a whole, with indicators on labour demand and to monitor their evolution over time.

This need is partly filled with Job Vacancy Statistics that have been collected since 2008, at aggregate level, based on Regulation (EC) No 453/2008 of the European Parliament and of the Council. The Job Vacancy Rate (JVR) is published every quarter as a euro indicator.

By relating the stock of unfilled vacancies to the total labour demand, i.e. filled and unfilled posts, the JVR provides an indicator on the level of difficulty faced by employers to recruit labour force, i.e., labour tightness. In periods of relatively high demand, the labour market is said to be 'tight'. Unemployment will be low and there will be many unfilled job vacancies. When the supply of labour is relatively high, the market is 'slack' with few vacancies and many jobseekers.

The Online Job Advertisement Rate (OJAR) complements the published Job Vacancy Rate with breakdowns by occupations obtained from Online Job Advertisements (OJA) data, as collected through the Web Intelligence Hub (WIH) infrastructure of Eurostat.

Publishing labour market demand indicators based on OJA data, as experimental statistics, is therefore directly relevant to fill the information gaps of the existing official statistics on the Job Vacancy Rate.

# Definition and calculation

To calculate an indicator on labour market demand, the first idea is to mimic the Job Vacancy Rate (JVR) which is defined as:

```
JVR = [Number of vacant posts / (number of vacant posts +
number of occupied posts)] × 100
```

By replacing the number of vacant posts by the number of OJAs and the number of occupied posts by the number of employees, collected from the EU-LFS, we obtain a first specification of the Online Job Advertisement Rate as follows:

```
OJAR1 = Number of OJAs / (number of OJAs + number of employees
[from EU-LFS])
```

However, despite the continuous improvement of algorithms, the WIH source may be affected by the multiple counting of the job advertisements posted on different job portals.

This would impact on the denominator, hence the comparability of the OJAR indicator across countries.

Therefore, Eurostat computes the OJAR as a simple ratio between the number of OJAs and the number of employees, i.e.:

```
OJAR = [Number of OJAs / number of employees (from EU-LFS)] ×
100
```

With this specification, the possible multiple counting of OJAs only has a 'scaling effect', increasing the ratio by a given factor, without impacting the other source (EU-LFS).

# Data sources

The Online Job Advertisement Rate is calculated from two different sources: the Web Intelligence Infrastructure (WIH) for OJAs and the EU Labour Force Survey (EU-LFS) for the number of employees.

## First source: online Job Advertisements

The number of OJAs is counted from the information collected through the WIH infrastructure. These data cover millions of ads posted in EU countries, collected from more than 500 web sources, including job search engines and public employment services' websites. The information retrieved from these ads is processed and classified according to variables, such as occupation or location. As an approximation of the stocks of open online advertisements, we count the flow of OJAs posted from the first to the last calendar day of

each quarter within the reference year. Those quarterly flows are then averaged over the reference year.

It is important to note that online job advertisements differ from job vacancies. On the one hand, there may be vacancies that are not advertised on the web, in particular in the public sector (e.g., teachers or healthcare workers). On the other hand, a job advertisement might refer to several vacancies at the same time or might not correspond to a vacant post at all. Indeed, some employers post online job advertisements although they do not intend to fill them within a specific period of time e.g. because they wish to explore possible opportunities.

Data on the number of OJAs are obtained through a number of processes that improve the quality of the data. This includes selecting web portals, assessing the coverage of job ads for which the population or region could be retrieved and limiting duplicates.

This source dataset has the following strengths:

- OJAs provide very detailed information on the occupation and skills sought by employers;
- OJAs also frequently include additional information about the job for which the vacancy is advertised, such as location, education requirements, working conditions, and characteristics of the employer;
- OJAs are particularly appropriate to measure skills demand because they reflect the exact names and labels of the skills sought for by the employers posting the job advertisement;
- Timeliness is a big advantage of this data source, as the data is available as early as 1 month after the end of the reference period.

However, OJA data and the resulting Online Job Advertisement Rate remain experimental, as the current output does not meet the required standards of official statistics for certain quality dimensions.

First, the current approach does not produce figures on the completeness of the scraped web portals posting job advertisements. Thus, the selection of web portals might have a significant impact on the quality of the data output and the comparability across time and countries.

Moreover, the OJA source is biased regarding economic activities and occupations: certain positions are likely to be advertised online more than others. For example, IT positions are more likely to be advertised on the internet than jobs in small shops, such as butchers or bakeries, or in the public sector (e.g. teachers, care workers). This may lead to the under coverage of some occupations.

The same advertisement may be counted multiple times in case it was posted in several web portals. However, the algorithms are progressively improved to eliminate redundant OJAs.

The quality of the output data is also heavily dependent on the methodology used for classifying relevant characteristics in the raw data, i.e., keyword selection as well as language and quality of the classifications' translation into national languages. This is because the data is built using natural language processing of the information retrieved from the web, classified into several variables, of which some according to existing classifications.

There are large and complex amounts of data to be classified in the processing phase.

Therefore, some misclassifications may exist regarding the:

- occupation (based on the [international standard classification of occupations](#), ISCO)
- skills (based on the [European skills, competences, qualifications, and occupations](#), ESCO)
- geographical location (based on the [nomenclature of territorial units for statistics](#), NUTS)
- economic activities (based on the [statistical classification of economic activities in the European Community](#), NACE).

Consistency issues between classification levels or data releases may arise.

Continuous work on the evaluation and quality improvements of the used algorithms is ongoing.

Details on accuracy of the source data used can be found in the metadata of the source dataset at: [WIH-OJA](#)

## Second source: number of employees

The number of employees is estimated from the EU Labour Force Survey (EU-LFS) for the population aged 15 to 64. In the EU-LFS, employees are defined as persons who work for a public or private employer based on a written or oral contract and who receive a payment in cash or in kind. This payment in cash or in kind is not directly dependent upon the revenue of the unit for which they work. Members of the armed forces are also included if they are part of the survey target population. Employees normally work under direct supervision of, or according to guidelines set by the employing organisation or enterprise (see [explanatory notes](#)).

The number of employees collected from the EU-LFS is published as official statistics which ensures a high level of quality including comparability across countries.

# Data to be published

## Breakdowns and variables

Both OJA and LFS data are available at the 4-digit level of the ISCO classification. However, LFS data might not be publishable for many ISCO4 positions, in particular for smaller countries. For this reason, the OJAR indicator is published for the EU as a whole (EU27), at the 3-digit level of the ISCO classification.

To summarize, this experimental dataset includes information on labour demand:

- for EU27 totals

- by occupation, at the 3-digit level of the ISCO classification

- for a group of occupations that includes specialists in Information, Communication and Telecommunications (ICT). This corresponds to ISCO codes: 133, 2152, 2153, 2166, 2356, 2434, 251, 252, 351, 352, 3114, 742.

The dataset also includes the corresponding number of OJAs and employees.

The released variables are the following:

- the Online Job Advertisement Rate itself (OJAR), in % (ratio)
- the components of the Online Job Advertisement Rate, namely the number of OJAs and employees, in thousands.

## Reference period and flags

The data are released for annual reference periods, starting from 2019 when OJAs started being collected by the WIH infrastructure, up to the last year available for LFS data, i.e. 2023.

In addition to the annual series to be published, quarterly series are also available for OJAs and the number of employees.

To verify if quarterly data were interpretable, Eurostat checked the correlation coefficients between both series, the volatility of the quarterly changes in the OJAR ratio as well as the lagged correlations:

```
CORREL [(OJAR_{T+1} − OJAR_T); (OJAR_T − OJAR_{T-1})]
```

Lagged correlations were significantly negative at EU level, in most cases, indicating that quarterly movements in the Online Job Advertisement Rate could not be interpreted, thus confirming the choice to publish annual data.

Moreover, to avoid excessive volatility due to the low number of observations, Eurostat excluded series with less than 40 OJAs in any year over the reference period. Occupations with less than 40 OJAs were aggregated into an ISCO category: 'OTH' (Other occupations) in order to restore the additivity constraint: the EU total should be equal to the sum of all ISCOs.

The same treatment was applied to the variable 'number of employees' when the number of observations was below the EU-LFS dissemination threshold. There again, the occupations concerned were aggregated into the ISCO category: 'OTH' (Other occupations), together with the missing ISCO, thus meeting the additivity constraint.

# Coverage

The data cover the European Union as a whole (EU27).

# How to interpret the data

The Online Job Advertisement Rate should not be interpreted as a proxy for the Job Vacancy Rate. The calculation formula differs and OJA can provide additional breakdowns, not a substitute for job vacancies that still provide the reference/benchmark totals.

The OJAR rather provides a relative ranking of occupations in terms of labour market demand measured from online job advertisements. This is expected to fill important data gaps for e.g. training policies that require identifying the occupations, hence skills, with the main bottlenecks.