Big Data for Labour Market Intelligence

Capacity development programme 2024

Workshop 1

Combine and compare Online Job Vacancy (OJV) data with traditional data

Speaker: Emilio Colombo
14/02/2024
Overview

- What do OJA represent?
- How are they related to labour market variables?
- Can they help to predict labour market variables?
- How can we extract LMI from OJAs?
- Conclusions
OJA and representativeness

Are OJA representative of the LM structure?

- We need to define what we observe
- We need to select a benchmark
- We need to compare OJA with the benchmark
OJA and the population

We observe some (large) number of OJA
Are they representative of the universe of OJA?
We do not know as the true population of OJA is unobservable
We need to estimate it

- Capture recapture models: need ads to be posted in different sites
- Multilevel modelling: individual response is a function of different covariates
OJA finding the right benchmark

Which labour market variable and statistics is closest to the concept of OJA?

1. Job Vacancy Statistics
   Advantage: JVS is the closest definition to OJA. Allows for sectoral dimension
   Problem: JVS are constructed from a survey of firms which has a relatively small sample size and
   has no information about occupations nor it allows a regional division. Moreover, JVS measure the
   stock of vacancies, OJA the flow.

2: LFS
   Advantage: by far the most detailed and granular survey on Labour Market in Europe
   Problem: LFS is a supply side variable. Measures the stock of employment, (not vacancies)
   We can restrict the analysis to job changes <3months but different concept wrt OJA

3. Administrative data: very promising not available across countries
OJA representativeness

If OJA represents a labour market phenomena and they are a biased description of it, they can be “corrected” to match the distribution of some benchmark variable. Post-stratification delivers just that, i.e. it reweights the population of OJA with weights obtained from the reference sample.
OJA representativeness: PS by ISCO 1D

[Graph showing the representativeness of OJA by ISCO 1D categories with bars for different IT levels (1-9).]
What about prediction?

Even if OJA are a biased description of some LM phenomena/variable they may be a good predictor of the same phenomena.
For a good fit both in sample and out of sample representativeness is not needed, what is needed is predictive ability (fit or MSE).
This is true both at cross sectional and at time series analysis.
Caution: OJAs suffer from source instability, thus stable sources must be used!
Predictive ability: OJA vs JVS
Predictive ability: OJA vs LFS
The real value of OJAs: skills

OJAs have several drawbacks when compared with other standard LM variables. They lack representativeness, coverage for the whole economy, stability over time etc. However they have a unique information: skills. As such they can be used to complement other type of statistics to define the skill content of occupations. This allows to capture the change that occurs along the intensive margin in the labour market as well as the one that occurs along the extensive margin.

**Extensive margin**: job creation and job destruction

**Intensive margin**: change in the skill content of occupations: changes that occurs within occupations
Examples and applications

- Research for Cedefop on measuring skill intensity and skill change
- Research on the role of market concentration for skill demand
Detecting the change in the skill content of OJAs

Changes in the occupations’ skill bundles can be decomposed in two dimensions:
- **extension**: the domain of unique skills we observe in OJA for each occupation changes through time

- **intensity**: the frequency of skills demanded in OJA for each occupation varies through time in absolute and, most importantly, relative terms.
RCA measures

- Approach pioneered by Balassa (1965) to measure trade specialisation. Used by Hidalgo et al. (2018) to measure skill specialisation of occupations.

- Is a given skill relatively more or less demanded in a given occupation compared to the overall market? Or, analogously, is the same skill more concentrated in a given occupation compared to the overall market?

- If a skill is relatively more (less) frequent in a given occupation compared to the market average, the occupation is (under-)specialised in that skill.
RCA measures. Balassa RCA

There are many classes of RCA indexes. The most widespread is the Balassa index. For a given occupation \( i \) and skill \( j \):

\[
\text{rca}_{ij} = \frac{\frac{s_{f_{ij}}}{\sum_{j=1}^{t} s_{f_{ij}}} \sum_{i=1}^{t} s_{f_{ij}}}{\frac{\sum_{i=1}^{t} s_{f_{ij}}}{\sum_{i=1}^{t} \sum_{j=1}^{t} s_{f_{ij}}}}
\]

Relative skill \( j \) frequency in occupation \( i \)

Average relative skill \( j \) frequency

If \( rca_{ij} > 1 \), occupation \( i \) is specialised in skill \( j \)
Use case 1 – OJA shares and change

Skill change and OJAs share for top 5 most frequent High, Medium, and Low skill occupations. 2019 – 2022 Germany
Use case 1b – OJA shares and change

Skill change and OJAs share for top 5 most frequent High, Medium, and Low skill occupations - Italy, 2019-2022
Use case 2 – OJA shares and change determinants

Determinants of skill change. 2019-2022. Germany. Top 5 most frequent High, Medium, and Low skill occupations

<table>
<thead>
<tr>
<th>Code</th>
<th>Occupation Description</th>
<th>abs_change_digital%</th>
<th>abs_change_transversal%</th>
<th>abs_change_professional%</th>
</tr>
</thead>
<tbody>
<tr>
<td>4419</td>
<td>Clerical support workers not elsewhere classified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7422</td>
<td>Information and communication technology installers and services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4225</td>
<td>Enquiry clerks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5321</td>
<td>Health care assistants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5223</td>
<td>Shop sales assistants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8219</td>
<td>Assemblers not elsewhere classified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9112</td>
<td>Cleaners and helpers in offices, hotels and other establishments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8211</td>
<td>Mechanical machinery assemblers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9333</td>
<td>Freight handlers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9529</td>
<td>Manufacturing labourers not elsewhere classified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2431</td>
<td>Advertising and marketing professionals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3343</td>
<td>Administrative and executive secretaries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2149</td>
<td>Engineering professionals not elsewhere classified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2512</td>
<td>Software developers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2511</td>
<td>Systems analysts</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Use case 2b – OJA shares and change determinants

Determinants of skill change. 2019-2022. Italy. Top 5 most frequent High, Medium, and Low skill occupations

- 9412 - Kitchen helpers
- 7412 - Electrical mechanics and fitters
- 8219 - Assemblers not elsewhere classified
- 4312 - Statistical, finance and insurance clerks
- 5131 - Waiters
- 4311 - Accounting and bookkeeping clerks
- 9329 - Manufacturing labourers not elsewhere classified
- 2149 - Engineering professionals not elsewhere classified
- 2511 - Systems analysts
- 2431 - Advertising and marketing professionals
- 9112 - Cleaners and helpers in offices, hotels and other...
- 3343 - Administrative and executive secretaries
- 2512 - Software developers
- 5223 - Shop sales assistants
- 9333 - Freight handlers

- abs_change_digital%
- abs_change_transversal%
- abs_change_professional%
Use case 3. Country estimates

We aggregate at the country level by weighting each ISCO08 IV digit occupation by its OJAs share for the country labour market.

*Contribution to skill change of each skill group*
Use case 4. Occupation ISCO08 I digit

Analogously, we aggregate at the ISCO08 I digit by weighting each ISCO08 IV digit occupation by its OJAs mentions shares for the country labour market.
LM concentration and skill demand

Due to their granularity OJAs can be used to match local labour market variables. Colombo and Marcato (2023) use firm level data and OJA in the Italian labour market to analyse the effect of LM concentration on skill demand.

They show that employers in a highly concentrated labour market demand competencies associated with the ability of workers to learn faster (e.g., social skills) rather than actual knowledge. They also require less experience but higher education.

These results are consistent with the hypothesis that employers in more concentrated labour markets are more prone to train their employees. Instead of looking for workers who already have job-specific skills, they look for workers who can acquire them faster and more efficiently.
LM concentration and skill demand

An increase in concentration is associated with a lower demand in knowledge, organisational, digital and cognitive skills but higher demand in social and hard professional skills which are more likely to be taught on the job.
Thank you for your attention

Emilio Colombo
emilio.colombo@unicatt.it