Competition Among Employers In Urban Hiring Markets: Evidence From Online Job Advertisements

IPS45 - Web intelligence for official statistics

IAOS 2022

26-28 April 2022 Kraków, Poland

Outline

- Online Job Advertisements (OJAs)

- Labour Market Concentration Index

- Methdology

- Results and Future Work

Online Job Advertisements (OJA) data

- **Online job advertisements (OJAs)** refer to advertisements published on the Web revealing **an employer's interest in recruiting workers**

- OJAs usually include **data on the characteristics of the job** (e.g. occupation and location), characteristics of the employer (e.g. name of the employer, economic activity) and requirements (e.g. education/skills).

- Potential for complementing official labour market statistics with more granular data (and perhaps more timely)
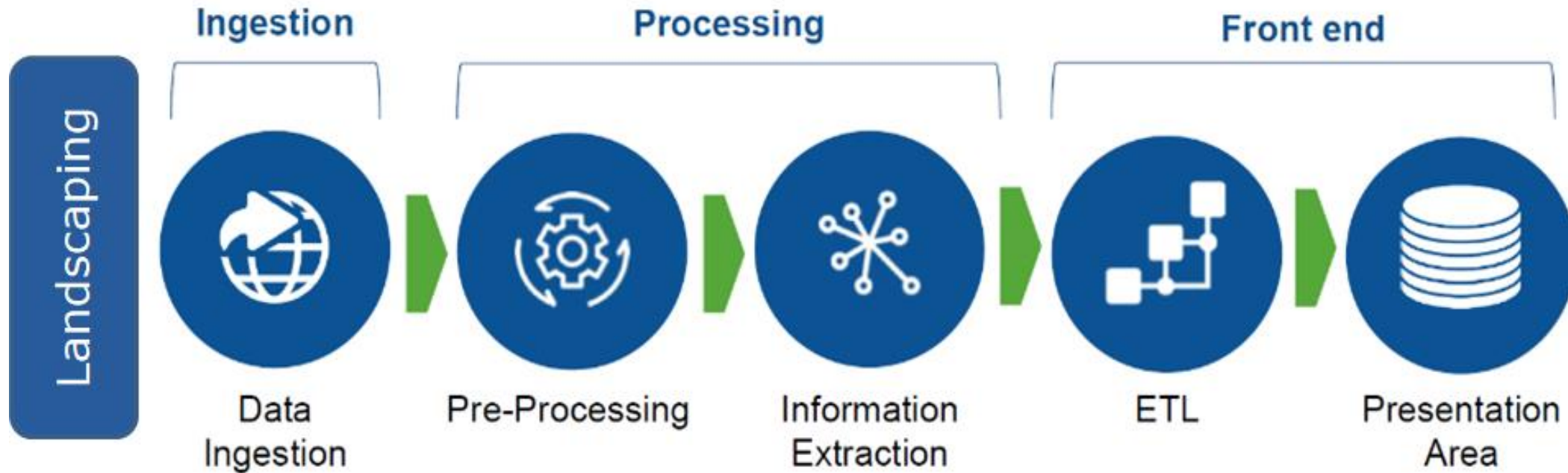
OJAs data @Eurostat

- Result of a joint project **between the ESS** and **The European Centre for the Development of Vocational Training (Cedefop)**

- OJAs are collected from the Web via APIs and scraping

- At Eurostat, OJAs represent one of the data available from the **Web Intelligence Hub** within the Trusted Smart Statistics (TSS) Initiative.

- OJAs data are scraped/downloaded for all 27 EU countries + UK. More than 140 millions distinct online jobs advertisements are collected from 316 distinct sources.

- Data from 2019 and 2020 are used for this study.

OJA data: Variables

Variable	Missing values	Variable type
economic activity of the employer	2%	Categorical (NACE at 2. level)
type of contract	29%	Categorical ("permanent", "self-employed" and "temporary")
working hours	38%	Categorical ("full-time", "part-time")
education level required	1%	Categorical (ISCED 2011)
salary	74%	Categorical (13 levels)
experience	51%	Categorical (8 levels)
place of employment (region)	36%	Categorical (NUTS3)
place of employment (city)	47%	Categorical (LAU)
occupation	0%	Categorical (ISCO level 4)
skills	1%	Categorical (ESCO level 3)
time (grab and expired dates)	0%	Date
company names	20%	String

European Commission

Data pipeline

Ingestion   **Processing**   **Front end**

Landscaping

Data Ingestion | Pre-Processing | Information Extraction | ETL | Presentation Area

AWS
Data Lab

Interested? ESTAT-WIH@ec.europa.eu

R | Python

Source: adapted from CEDEFOP (2019a)

European Commission

Methodology

Background

- [ESSnet Big Data II Workpackage B](#) – Case study on using OJA data for calculating a Labour Market Concentration Index

- First Experimental results for Germany: https://github.com/OnlineJobVacanciesESSnetBigData/Labour-market-concentration-index-from-CEDEFOP-data

- Research paper from Azar et al. on the calculation of the concentration index using Burning Glass data (OJA): https://www.sciencedirect.com/science/article/pii/S0927537120300907

Labour Market Concentration (LMC)

- How much choice people have when looking for employment?

- Labour market: job search for a given **occupation (**ISCO 4-digits classification)**, region (**Commuting zone - Functional Urban Area**) and time (**Quarter**)**

- In the case of few employers with open positions: Little competition between employer, Possible Monopsony, Low bargaining power for workers

- No published measures available (so far)

- **Herfindahl-Hirschmann Index** – used to measure concentration. Sum of the squares of market shares (ranges from 0 to 10000 monopolistic situation)

- One index for each local, occupational labour market {Quarter X Occupation X FUA}

The Herfindahl–Hirschman Index (HHI)

- Sum of the squares of the ad shares of each firm advertising in a city's labour market

- Ranges from close to 0 (perfect competition) to 10 000 (monopoly/monopsony), with 2500 generally considered as highly concentrated (US Department of Justice, 2018)

- Lower index ⇔ more competitive market ⇔ better for workers

HHI	Market shares' distribution in the special case of equal shares	Other example of market shares' distribution
0	Perfect competition among firms	NA
1000	10 companies with equal shares (10%) of the ads	2 companies with a share of 20% each, 18 smaller competitors sharing the rest of the market
2500	4 companies with equal shares (25%) of the ads	1 dominant firm with a share of 40%, 4 competitors with a share of 15% each
5000	2 firms posting 50% of the ads each	1 dominant firm with a share of 70%, 9 smaller competitors sharing the rest of the market
10000	Monopsony of one firm posting 100% of the ads	NA

R code

• Code available on Github https://github.com/eurostat/oja_hhi

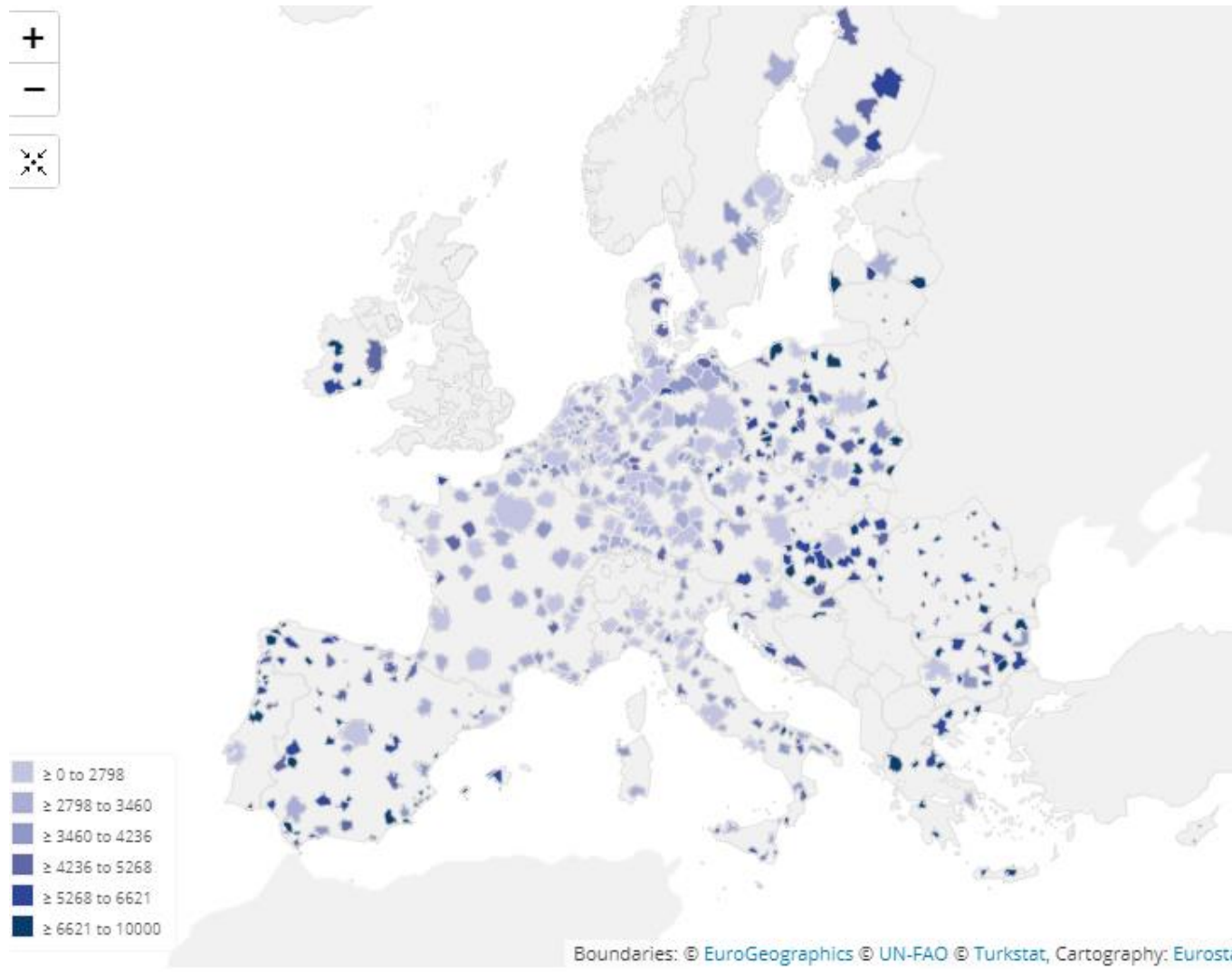|	
Data Query	• Querying all OJAs data for jobs located in the 27 Member States
Cleaning and Filtering	• Filtering all OJA where variables are not available (geoinformation, ISCO code) • Eliminating duplicates
Cleaning *companyname* variable	• Consolidating names of companies (basic string cleaning and consolidation of different spellings of the same company name) • Filtering OJA posted by intermediary agencies based on (i) keywords and (ii) classification model • Imputation of missing company names
Geodata Merging	• Assign OJAs to FUAs where possible (by city or NUTS3) • Download georeferenced data for visualization • Handle country exceptions
HHI index calculation	• Calculate HHI index – alternatives and aggregates
Results Visualization	• Aggregate country results • Prepare output tables and visualizations (gisco+ggplot)

European Commission

Results and Next Steps

The HHI across EU urban areas, fourth quarter of 2020

Boundaries: © EuroGeographics © UN-FAO © Turkstat, Cartography: Eurosta

Legend:
- ≥ 0 to 2798
- ≥ 2798 to 3460
- ≥ 3460 to 4236
- ≥ 4236 to 5268
- ≥ 5268 to 6621
- ≥ 6621 to 10000

Interactive map: https://ec.europa.eu/eurostat/cache/RCI/rcit/lmci.html

Lowest levels of the HHI:

Functional Urban Area	HHI, 2020 (equivalent number of companies)
Berlin	1 022 (9.8)
München	1 218 (8.2)
Hamburg	1 238 (8.1)
Paris	1 238 (8.1)
Milano	1 291 (7.7)
Ruhrgebiet	1 367 (7.3)
Stockholm	1 416 (7.1)
Bruxelles	1 489 (6.7)
Amsterdam	1 525 (6.6)
Düsseldorf	1 689 (5.9)
Lisbon	1 692 (5.9)
Köln	1 696 (5.9)
Leipzig	1 728 (5.8)
Lyon	1 739 (5.8)
Utrecht	1 743 (5.7)
Madrid	1 748 (5.7)
Dresden	1 759 (5.7)
Barcelona	1 772 (5.6)
Roma	1 794 (5.6)
Stuttgart	1 796 (5.6)

Future Work

- Extend data coverage to all EFTA countries

- Improve the cleaning and classification of company names

- Calculate the index using a different (broader) definition of labour markets

- Integrate/Correlate the results with more Eurostat statistical data collected at NUTS3 level

- Further analyse the results of specific occupation code

Thank you

European Commission