EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Experimental OJA based indicators on labour demand changes: opportunities and challenges

Annalisa Lucarelli
(ISTAT- DCSW - *Social Statistics and Welfare Central Directorate*)

Francesca Inglese, Renato Magistro, Giulio Massacci and Giuseppina Ruocco
(ISTAT - DCME – *Methodology and Statistical Process Design Central Directorate*)

Kostadin Georgiev and Galya Stateva
(BNSI - Eurostat)

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

Contents

- Job Vacancies official Statistics (JVS) for analysing the demand side of labour market
- Online Job Advertisements (OJAs) *vs* JVS
  - different definitions and measures
- European projects on OJAs with Istat participation
- OJAs *vs* JVS: highly detailed information at high frequency
- OJA data source – CEDEFOP
- Quality issues
- OJA based experimental indicators
- Conclusion: opportunities and challenges

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

JVS for analysing the demand side of labour market

- In the context of the official statistics covering the demand side of the labour market, there are **official job vacancy surveys** that supply quarterly information on the **unmet labour demand**
- As job vacancies measure employment intentions that have materialized in candidates' searches, they can give "early warnings" on the dynamics of jobs in the near future
- It connects vacancies to short-run economics and makes job vacancy statistics **leading indicators of the economic cycle**.
- Data on job vacancies are used by the **European Commision and the European Central Bank** to monitor **short-term developments in the business cycle** and the labour market
- The vacancy indicator currently used at the European level – the job vacancy rate – is one of the **Principal European Economic Indicators PEEIs** on the labour market
- The production by the NSIs takes place on the basis a **EU framework Regulation (No 453/2008)**

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

JVS *vs* OJAs different definitions and measures (1/2)

- Online Job advertisements are one of the methods to notify a job vacancy

OJAs refer to advertisements published on web-sites/job portals

- OJAs cannot be assimilated to JVS as defined by the EU Regulation

"*a paid post that is newly created, unoccupied, or about to become vacant: for which the employer is taking active steps ….; and… intends to fill either immediately or within a specific period of time*"

- notifying the vacancy to public employment services and/or contacting a private employment agency/head hunters;
- advertising the vacancy in the media (for example the Internet, newspapers, magazines) and web-site/job portals;
- personal contacts, word-of-mouth;
- using internships;
- ……

It' s only one of the possible active steps

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

JVS *vs* OJAs different definitions and measures (2/2)

not all job vacancies are advertised online and employers can use methods other than an advertisement on the websites:

- even if there is a generally growing trend in the number of job vacancies being offered online, many vacancies continue to be filled through traditional channels, such as newspapers, employment agencies, noticeboards or personal contacts

- job advertisements and job vacancies from a probability-based survey are two different measures of job opening
- they cover different populations of job opening
- job vacancies are measured at a point in time, making it a stock variable
- job advertisements are measured per unit of time, making it a flow variable

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

European projects on OJAs with Istat participation

ESSnet projects: USE OF BIG DATA FOR STATISTICAL PURPOSES

THE IDEA:

Online Job Advertisements (OJAs)

TO IMPROVE

Job Vacancy official statistics (JVS): EU Regulations

BIG DATA EUROPEAN PROJECTS:

ESSnet Big Data I (2016-2018)

ESSnet Big Data II (2018-2020)

***ESSnet - Web Intelligence Network (WIN)* 2020-2025**

*OJAs **can not be used to directly replace JVS required by EU Regulation**..... but OJA data can provide **more granular insights** than JVS. OJA data within official statistics as the basis for producing **supplementary indicators**.."*

OJAs *vs* JVS: enrichment of strategic information

JVS from OFFICIAL SURVEY on the basis a EU Regulation		OJAs from WEB DATA
National, by NACE Rev.2 economic activity section, quarterly		National, by NACE Rev.2 economic activity section, quarterly
Stock of vacancies on the last calendar day of the quarter		Stock of vacancies on a specific day of the reference quarter
not available	ENRICHMENT	**Flow of OJAs collected on each day of the month**
Region on voluntary basis		**Geographical area/ region / province / city**
Occupation on voluntary basis		**Occupation, education, skill, contract type, working hours, salary**
not available		**Daily (weekly, monthly, quarterly)**

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

OJAs data source: Cedefop DataLab

- A pan-European approach to collect OJAs in all European countries since 2018
- Cooperation between various institutions from the European Union
- Continuous collection and High frequency information (on a daily basis)
- Detailed information on the characteristics of the:
  - the job (occupation, location, contract, working time, and salary);
  - the employer (economic activity sector);
  - job requirements (education, skill, and experience);
  - advertisement (publishing and the expiring date of the ads)
- Data processing: data collection, data cleaning, standardization, classification, validation and plausibility rules
- Classification of information on the basis of International Standard Classifications (NACE Rev. 2., ESCO, ISCO-08, NUTS, ISCED)
- Quarterly release notes and a Blog

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

OJA quality issues (1/3)

Representativeness:

- OJAs over-represent certain occupations and skills and under-represent other
  - over-representation for workers with higher level of education and large size enterprises, the opposite for small enterprises (in Italy word-of-mouth is frequent for seasonal small enterprises)

Coverage and selection:

- Under-coverage
  - (selection bias) not all job vacancies are advertised online. Employers use methods other than an advertisement on the websites
  - the job portals/web-sites are not necessarily totally covered by data ingestion activities their coverage varies over time
- Over-coverage
  - job ads may not be removed from portals, even if the vacant position has been filled (delays in the communication)
  - job ads placed online to investigate the market not always refer to a position that actually exists in the scope of the official survey (ghost vacancies)
  - duplicates: job offers published on several web sources or published several times (on the same website)

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

OJA quality issues (2/3)

Quality of the selected sources (website/job portals)

- **Relevance** of the sources
- **Stability** w.r.t. the **existence of the sources** over time and **popularity** of the sources
- **Coverage** of the relevant information

Why are **dynamic** sources a problem?

Goal: capture the dynamics of the **labour market with the help of OJAs**

Problem: OJAs might (*additionally*) **capture the dynamics of the sources** (*concept drift*)

Assessing stability of the sources

- number of sources over time
- determine whether the sources are the same during the period under analysis
- determine if the most important sources at several points of time are present over the whole time span

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

Cedefop OJA data by source - 2019 - 2023, Italy

	2019				2020				2021				2022				2023			
Sources	**March**	**June**	**Sepetember**	**December**	**March**	**June**	**epetembe**	**December**	**March**	**June**	**epetembe**	**December**	**March**	**June**	**Sepetember**	**December**	**March**	**June**	**Sepetember**	**December**
Source_1	17,0%	20,9%	7,4%	6,5%	5,3%	8,5%	8,2%	8,2%	13,2%	13,9%	14,4%	10,8%	14,7%	12,9%	10,6%	10,3%	10,8%	9,9%	9,8%	22,9%
Source_2	3,2%	9,9%	4,4%	5,2%	6,5%	4,2%	6,9%	5,1%	10,6%	11,0%	6,4%	8,9%	5,9%	9,2%	11,2%	10,9%	3,0%	4,6%	9,0%	18,4%
Source_3	7,8%	4,3%	2,5%	2,1%	2,0%	1,7%	1,1%	2,1%	5,4%	5,9%	6,2%	5,5%	8,0%	6,2%	4,7%	6,4%	7,3%	7,6%	6,9%	11,4%
Source_4	0,6%	1,3%	0,4%	0,8%	0,5%	0,6%	0,6%	0,7%	2,0%	2,0%	2,6%	2,4%	3,1%	3,0%	2,6%	1,0%	3,2%	3,3%	2,9%	8,0%
Source_5									0,9%	2,9%	3,1%	3,5%	3,8%	3,5%	3,2%	3,7%	3,5%	3,2%	3,2%	8,0%
Source_6	0,3%	1,5%	0,6%	0,0%	0,1%	0,1%	2,1%	1,6%	2,5%	2,0%	2,7%	2,0%	1,5%	12,6%	9,5%	10,4%	12,0%	0,4%	0,6%	6,1%
Source_7	8,5%	0,6%	7,0%	8,6%	6,3%	4,0%	3,4%	3,6%	4,1%	1,1%	2,3%	1,9%	1,0%	2,4%	1,5%	1,4%	3,0%	1,7%	2,0%	5,8%
Source_8								0,4%	0,7%	0,5%	0,5%	0,7%	0,6%	1,1%	1,1%	1,2%	1,3%	2,1%	0,8%	5,7%
Source_9	1,8%	2,3%	0,8%	0,8%	0,7%	1,1%	1,0%	1,2%	1,8%	2,0%	2,2%	2,0%	2,4%	2,6%	2,4%	2,3%	2,2%	2,4%	2,2%	3,8%
Source_10									1,1%	1,2%	1,3%	1,5%	1,5%	1,8%	1,4%	1,4%	2,0%	1,3%	1,6%	3,5%
Source_11								0,0%	0,1%	2,9%	2,7%	1,5%	1,7%	1,9%	2,1%	0,8%	0,8%	1,1%	1,1%	2,2%
Source_12	1,0%	0,9%	0,3%	0,4%	0,5%	0,6%	0,6%	0,4%	1,3%	0,6%	1,1%	1,5%	0,9%	1,2%	0,9%	0,8%	0,5%	0,6%	1,0%	1,6%
Source_13						0,2%	0,1%	0,2%	0,2%	0,3%	0,4%	0,3%	0,4%	0,4%	0,3%	0,3%	0,5%	0,5%	0,5%	1,1%
Source_14	0,8%	0,9%	0,3%	0,2%	0,2%	0,4%	0,3%	0,3%	0,6%	0,5%	0,6%	0,5%	0,7%	0,5%	0,4%	0,5%	0,5%	0,4%	0,4%	0,8%
Source_15								0,1%	0,2%	0,1%	0,2%	0,2%	0,2%	0,2%	0,2%	0,3%	0,2%	0,2%	0,2%	0,6%
Source_16	0,3%	0,9%	0,3%	0,3%	0,0%	0,0%	0,3%	0,3%	0,6%	0,0%	0,0%	0,0%	0,0%	0,6%	0,6%	0,3%	0,1%	0,1%	0,1%	0,1%
Source_17	2,9%	3,1%	1,4%	1,6%	1,3%	1,3%	1,2%	1,5%	2,9%											
Source_18	1,0%	2,8%	52,1%	54,7%	51,5%	45,8%	50,6%	49,1%												
Source_19	2,2%	1,2%	0,7%	0,9%	0,4%	1,4%	2,2%	2,4%	4,4%	2,8%	4,4%	3,8%	4,9%	6,0%	3,8%	3,2%	4,2%			
Source_20	0,6%	1,1%	0,4%																	
Source_21	0,2%	4,4%	6,1%	6,4%	6,1%	7,4%	3,2%		1,7%	1,9%	1,0%	1,3%	0,9%	0,4%	0,8%	0,5%	0,9%	15,8%	11,0%	
Source_22																				
Source_23	32,8%	23,9%	2,8%	1,7%																
Source_24																				
Source_25				0,1%																
Source_26	0,7%	0,9%	0,3%	0,2%																
Source_27				5,5%	13,8%	17,5%	13,4%	16,2%	35,9%	39,6%	41,6%	43,4%	47,9%	33,7%	42,7%	44,4%	44,1%	44,8%	46,6%	
Source_28	11,0%	9,1%	9,6%																	
Source_29			0,2%	1,5%	0,4%	0,8%	0,7%	0,4%	0,0%											
Source_30	6,4%	8,6%	1,9%	2,2%	4,1%	3,8%	3,0%	5,2%	7,8%	8,7%	6,4%	8,2%								
Source_31	0,9%	1,3%	0,5%	0,3%	0,3%	0,5%	1,1%	0,8%	2,1%											
Over 80%	**83,4%**	**81,2%**	**86,6%**	**87,0%**	**83,4%**	**83,3%**	**82,3%**	**83,8%**	**81,5%**	**81,9%**	**82,4%**	**80,7%**	**81,3%**	**80,7%**	**82,5%**	**82,4%**	**81,9%**	**82,7%**	**83,4%**	**86,3%**
Total	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**	**100,0%**

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

OJA quality issues (3/3)

Quality of the classification

- extracting the classification variables is probably the most error-prone process step
- the classification variables play a major role for the production of OJA based indicators
- classification algorithms based on machine learning model to extract OJA information

Why are **classification errors** a problem?

Goal: capture the dynamics of the **labour market by occupation, skill, territorial area..**

Problem: OJAs might **not accurately capture** labour demand **dynamics by strategic variables of interest**

Assessing and improving classification accuracy

- data annotation exercise can be adopted
  - labelling (manually classifying) of job ads according to a classification
- two “labelling exercises” coordinated by Eurostat happened in 2022 and 2023
  - (**first exercise, occupation, Italy**) in a rather high percentage about 45% the two classifications provided the same results
  - (**second exercise, occupation, Italy**) improvement to 62%
  - (**second exercise, NUTS2, Italy**) the distribution of correctly classified reaches a quite similar percentage of about 60%

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

OJA based new experimental indicators (1/4)

GOAL:

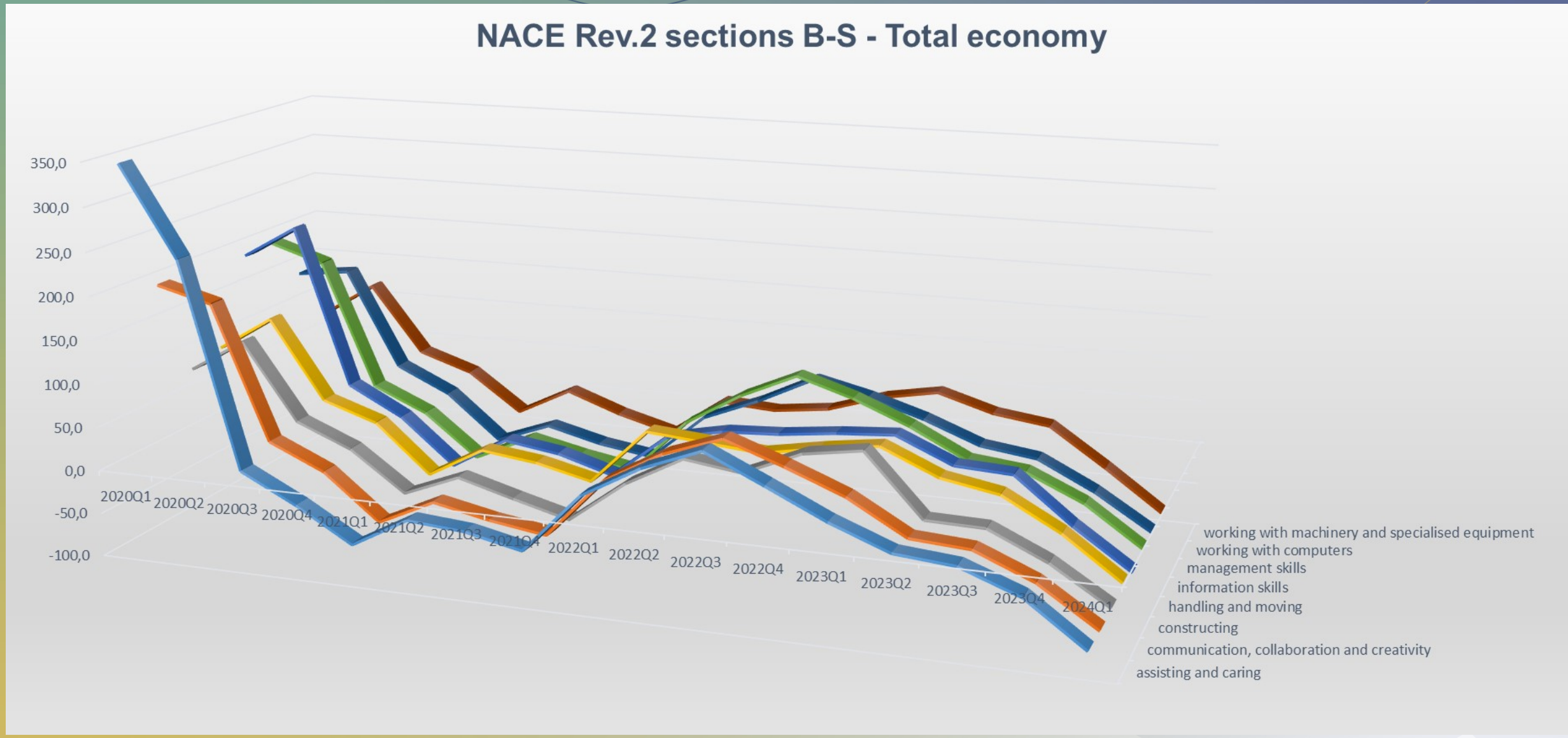
production of indicators on unmet labour demand dynamics

by strategic information not existing in official statistics

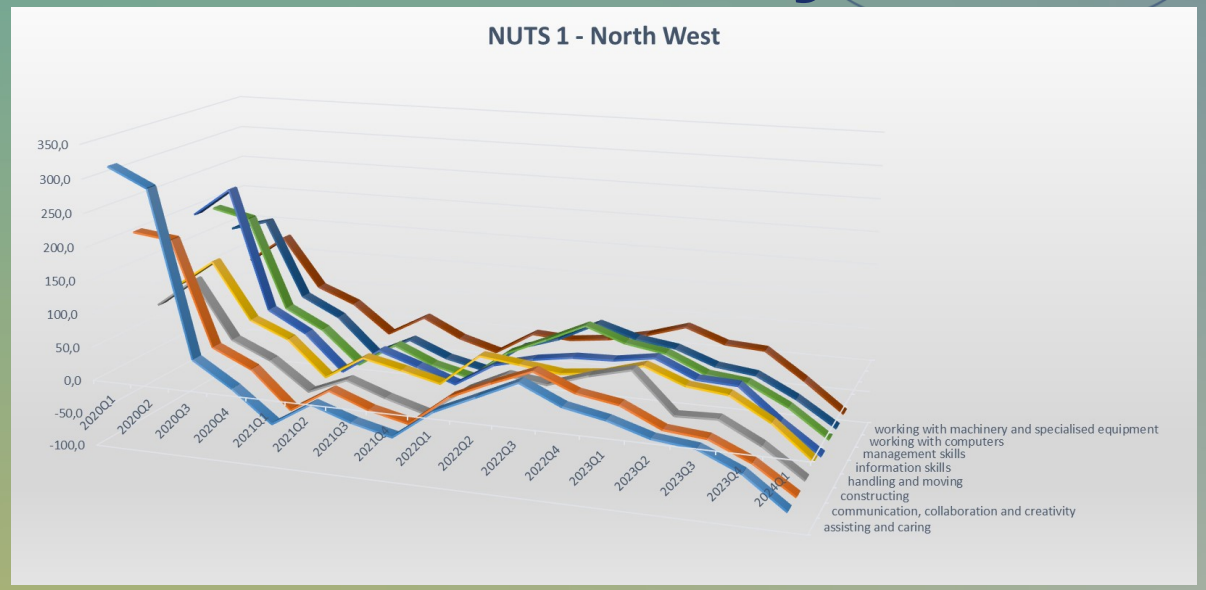
DONE SO FAR WITHIN THE COOPERATION WITH BULGARIA:

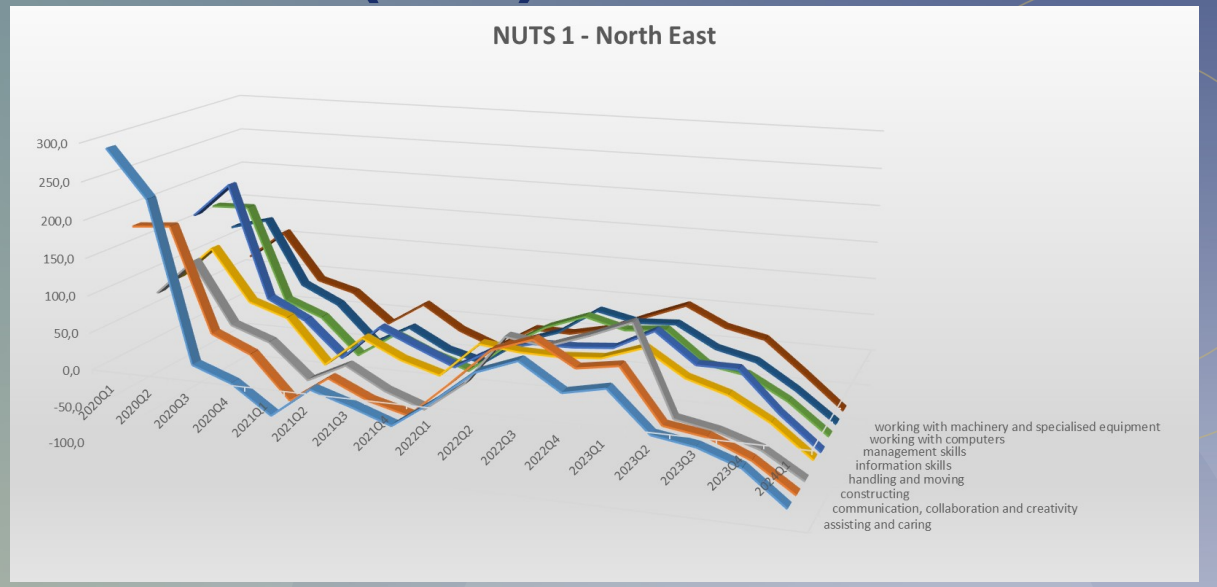
- selection of the type of indicator (level or changes)
- definition of a methodological framework
  - OJA stock and/or flow variables
  - breakdown (economic activity, occupation, skill, territorial area, etc.)
  - level of detail: to which digit of the International Standard Classifications
- production of year-on-year changes of OJA stock at the end of each quarter by:
  - ISCO-08 Occupation (level 1, level 2),
  - ESCO- Skill (level 0, level 1),
  - NUTS2 for Regions

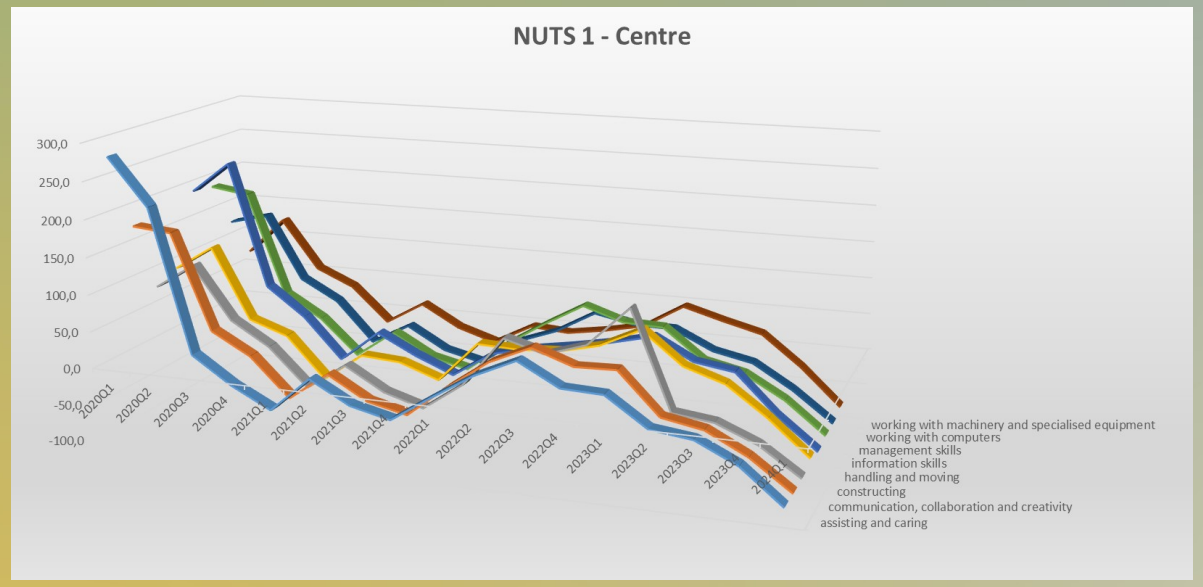
OJA year-on-year changes by skill
Italy, 2020Q1-2024Q1 (2/4)

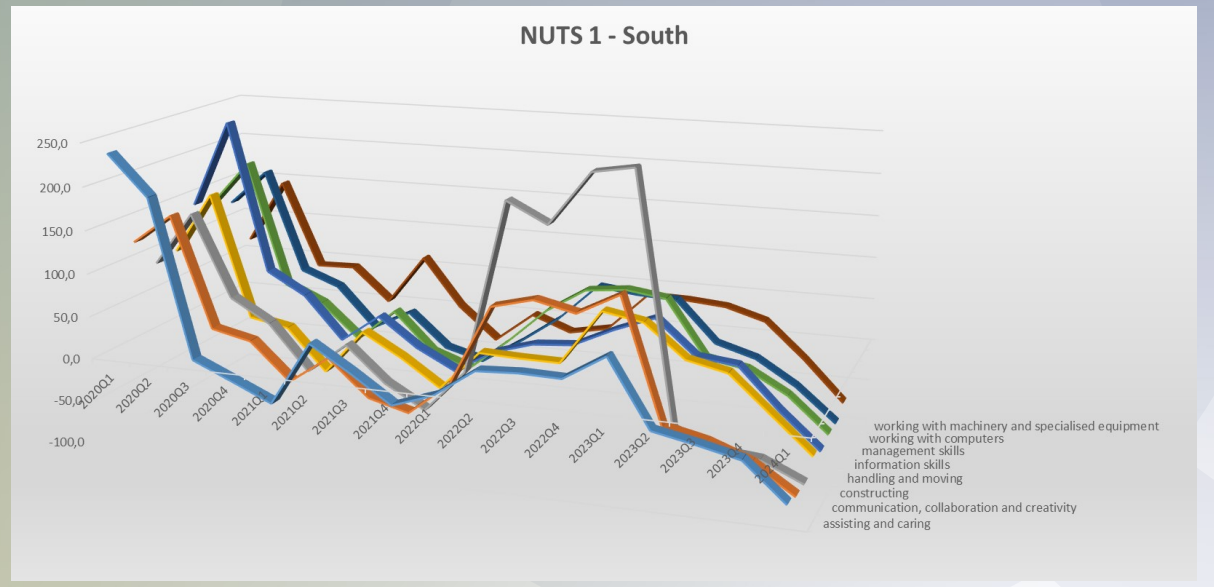


OJA year-on-year changes by skill and NUTS 1 area Italy, 2020Q1-2024Q1 (3/4)

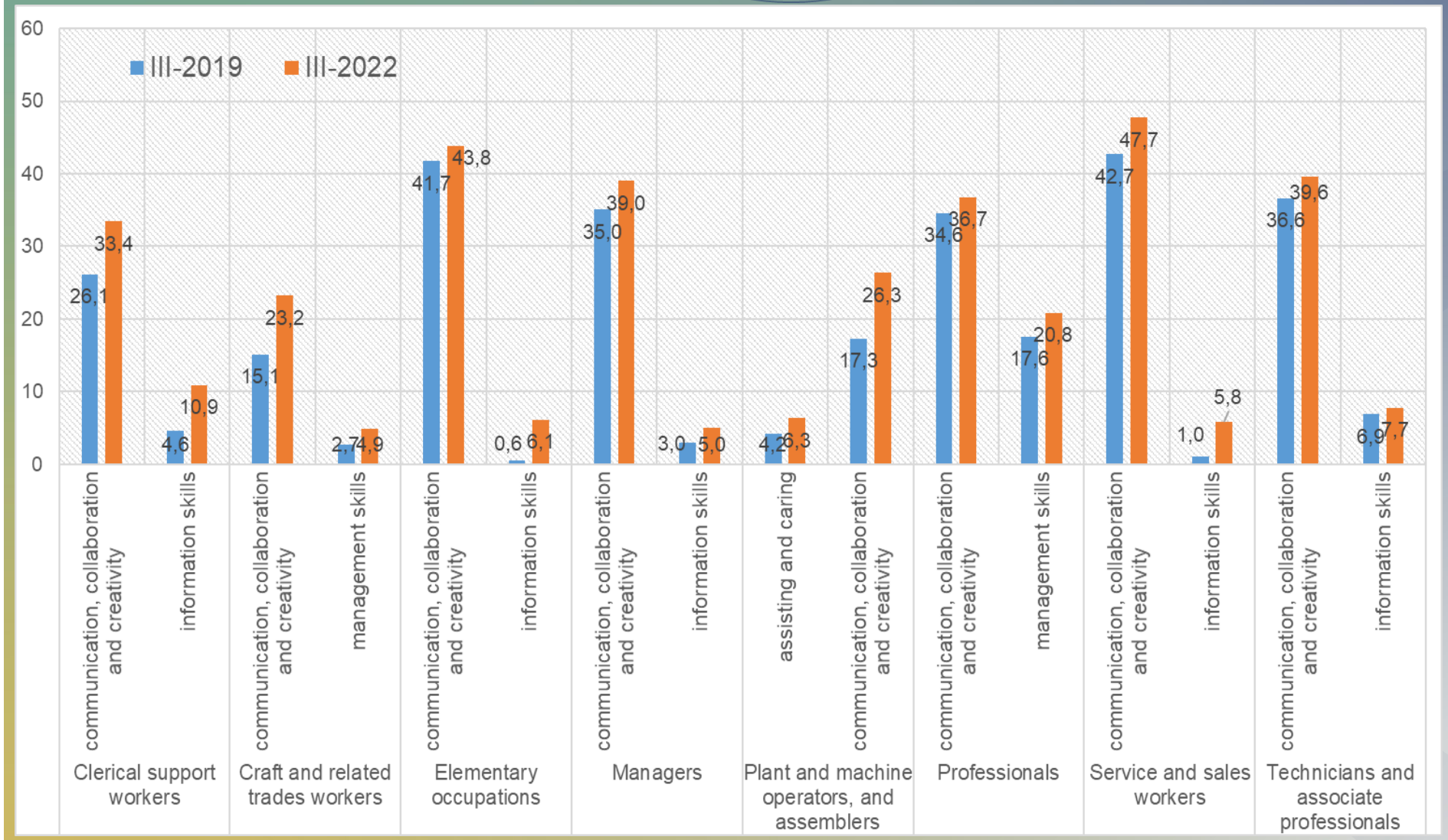








Skill increase by occupation – Italy, 2019Q3 and 2022Q3 (4/4)



- *communication, collaboration and creativity* the largest **increase** across all occupational groups
- except for Elementary occupation and Professional
  - the greatest increase in *information skill* and *management skill*
- *information skill* and *management skill* show the second largest increase of all groups

The conference is partly financed by the European Union

Conclusions

- Evidence consistent with the period under analysis characterized by the pandemic crisis
- the potential of the OJA data source is evident, but some methodological problems need to be addressed
  - classification accuracy
  - source stability

Challenges

- (*open issues*) Representativeness assessment:
  - comparing information collected from online sources with information from traditional data sources (Labour Force Survey and Vacancy Survey)
  - inserting a question in the current official Istat JV survey questionnaire on the share of vacancies that pass through the online recruitment channel

The conference is partly financed by the European Union

Opportunities

- Production of experimental statistics to integrate and supplement official statistics
  - complemented by quality indicators (metadata)
- Analysing how labour demand by enterprises changes in relation to changes in labour market characteristics
  - changes in the skills required by companies in relation to artificial intelligence development, green economy
- Meet new requirements from EU Regulations for unsatisfied/emerging needs of information
  - new EU Regulation on LMB statistics allows the use of innovative sources
- (**Italy**) Exploit the experience gained in the use of OJAs to fill the information gap on job vacancies in the public sector
- Extremely useful indicators on skill demand for the orientation of:
  - professional training/resolution of job matching problems
  - the funds provided by the National Recovery and Resilience Plan (NRRP)

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Thank you for your attention

anlucare@istat.it

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed
by the European Union