ISSN 1831-9424

European Commission

JRC TECHNICAL REPORT

Alternative data to monitor cultural occupations

An assessment of online job ads

Sanjuán, J., Montalto, V., Panella, F., Alberti, V.

Joint Research Centre

EUR 31416 EN

EU Science Hub
https://ec.europa.eu/jrc

JRC132055

EUR 31416 EN

PDF ISBN 978-92-76-99158-8 ISSN 1831-9424 doi:10.2760/008462 KJ-NA-31-416-EN-N

Luxembourg: Publications Office of the European Union, 2023

How to cite this report: Sanjuán, J., Montalto, V., Panella, F., Alberti V., *Alternative data to monitor cultural occupations*, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/008462, JRC132055.

Contents

Acknowledgements

This work was supported by the Spanish Ministry of Universities through the *Formación de Profesorado Universitario* Programme [FPU19/00182].

Authors

Jordi Sanjuán

Valentina Montalto

Francesco Panella

Valentina Alberti

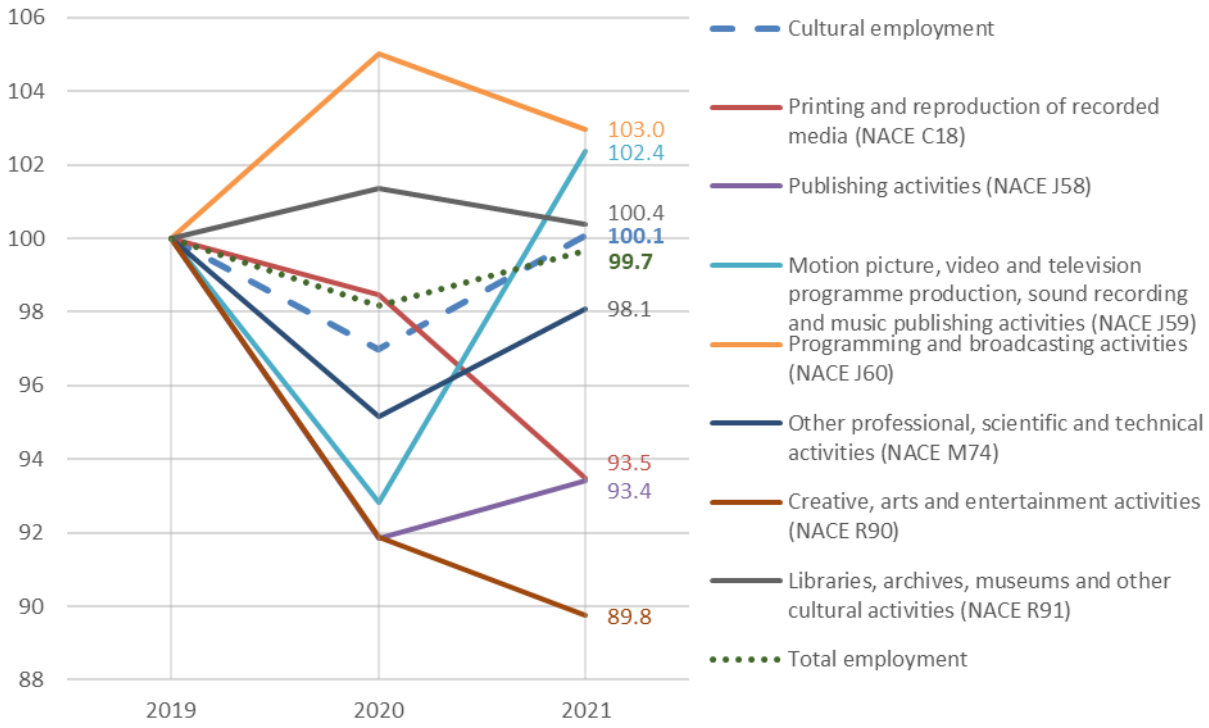
Abstract

The economic turbulence experienced in recent years as a result of the Covid-19 pandemic highlights the importance of obtaining complementary sources to official statistics for monitoring variables of socio-economic interest such as employment in real time. The need to obtain accurate and timely data is more important for domains such as the cultural and creative sectors (CCS), which have been highly affected by the social distancing rules introduced as a response to outspread of the pandemic. At the same time, the rise of the digital economy and the increasingly popular use of job posting platforms offer new and potentially very valuable sources of alternative data to monitor the labour market in real time. This technical report investigates the statistical and distributional features of online job posting data on cultural occupations in EU countries. Job posting data come from Online Job Adverts (OJA), a database developed by CEDEFOP: the European Centre for the Development of Vocational Training. The aim is to assess the representativeness of such data for policy-relevant work, using Eurostat cultural employment statistics as a benchmark. CEDEFOP collects data by web scraping a variety of sources including job search engines, public employment services, websites and agencies and online newspapers. This results in a database containing more than 134 million observations for the EU-27 and the UK.

1 Background

Covid-19 has caused a shock in the economy and in employment, which has affected all countries of the EU. The impact has been particularly significant in the cultural and creative sectors (CCS), due to the social distancing rules that have hit hard sectors whose main activities require physical presence. EY (2021) estimated that, for the EU, the fall in turnover in the cultural and creative sectors (CCS) in 2020 was 31%, and up to 90% for the performing arts or 76% for music. As for cultural employment in the EU, it decreased by 222,300 persons compared to 2019, according to Eurostat[1]. This represents a fall of -3.0 % compared to -1.8 % reported for the total employment. However, despite the more pronounced fall in 2020, employment in CCS recovered well by 2021, reaching the 2019 levels (Figure 1). However, this hides important differences within the CCS. While audiovisual-related activities have recovered rapidly or even experienced a surge in the midst of the pandemic (e.g. programming and broadcasting activities), others suffered major job losses that have not yet recovered. The sector most affected in this case is the performing arts.

Figure 1. Index change (base 100 = 2019) of employment (total, cultural and in some cultural sectors), EU27.



Source: Eurostat (lfsi_emp_a).

The effects of the pandemic on CCS risk to importantly reverberating on the entire economy and people's well-being. In the past twenty years, a growing corpus of studies and reports has increasingly reported on the relevance of culture as an economic sector per se (e.g. Throsby, 2008) which in turn contributes to a variety of aspects such as productivity, innovation, health, well-being or territorial regeneration (OECD, 2022). It is a sector that generates 4.4% of the EU's GDP (EY 2021), with above-average employment growth rates (0.83% on average in the EU27 over the last ten years, compared to 0.64% for total employment, according to Eurostat).

Building on this evidence, the European Commission has identified CCS as a priority investment sector whose development is intended to support Europe's shift towards a knowledge-based economy. Such vision is clearly reflected in the *New European Agenda for Culture* (European Commission, 2018), which comes as a renewed policy agenda for culture after the very first European Agenda for Culture adopted in 2008[2].

[1] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Culture_statistics_-_cultural_employment

[2] For an overview of the EU's cultural policy cornerstones, see: https://culture.ec.europa.eu/policies/strategic-framework-for-the-eus-cultural-policy

However, measuring the CCS and its broad economic and societal impacts remains a great challenge for a number of reasons that are explained in more details elsewhere (e.g. KEA, 2015). In particular, the contribution of culture to total employment is consistently underestimated due to the difficulty of grasping freelance professionals – which represent a notable share of workers in these sectors (about 30% according to Eurostat's data)– as well as cultural occupations done as secondary jobs – which is a quite common scenario in CCS (Eurostat, 2019).

New sources of data are emerging that could potentially open the way to more timely and detailed analysis of cultural employment to support policy-making. In particular, this work explores the Online Job Adverts (OJA) data from CEDEFOP, which collect information from a wide range of job posting platforms (CEDEFOP, 2019).

CEDEFOP's OJA database has already been the basis for a number of analyses and research projects, ranging from country-specific assessments (e.g., Giambona et al., 2021), to the development of tools to monitor changes in the labour market (e.g., Giabelli et al., 2021) to the exploration of the markets' demand for specific skills (e.g., Marconi & Vergolini., 2022). Nevertheless, to the authors' knowledge, no attempt has been made to use the OJA database to explore online advertisements for cultural occupations.

OJA data is coming from web sources and the validation and quality assurance processes are not performed by National Statistical Offices: therefore, the data may not be representative by design. It is important to assess their statistical properties and the extent to which analysis based on such data can be generalised for EU policy purposes, ahead of using them. This work aims to perform such an assessment, and to investigate the statistical and representativeness properties of Online Job Adverts data for cultural occupations and over time, using Eurostat cultural employment statistics as a benchmark.

More specifically, we first explore territorial differences in the job adverts collected, and whether the shares of cultural occupations out of total job ads correlate with the share of cultural employment in each country. Secondly, we analyse which educational levels are most in demand in the database and compare them, again, with those of official employment statistics. Thirdly, we do the same by distinguishing between the different cultural occupations to see which are more represented in the database with respect to their weight in employment. Finally, we analyse the evolution over time of job advertisements comparing it with that of official statistics on total and cultural employment.

We also look at the general characteristics of the database as a whole, so that the evidence obtained regarding its usability can also be applicable for other research purposes.

2 Methodology

For the purposes of this work, the statistical concept of cultural employment is defined following the methodology proposed in the ESSnet-Culture final report (2012). It encompasses situations where a person:

- holds a cultural occupation and works in the cultural sector (e.g. a ballet dancer employed by a ballet company or a journalist working for a daily newspaper);
- holds a cultural occupation outside the cultural sector (e.g. a designer working in the automotive industry); or
- holds a non-cultural occupation in the cultural sector (e.g. an accountant working in a publishing house).

To calculate cultural occupation statistics based on this definition, Eurostat cross-tabulates ISCO and NACE cultural codes (three digits for NACE, four for ISCO)[3]. However, as NACE and ISCO classifications have not been developed to capture cultural sectors and occupations, many NACE and ISCO codes encompass activities and jobs that cannot be defined as 'cultural'. The ESSnet-Culture final report thus puts forward a list of 'fully' and 'partially' cultural NACE and ISCO codes; this was later amended by the Eurostat working group on culture statistics. While 'fully' cultural codes are included in the cultural occupation statistics, 'partly' cultural codes are only in some cases included in order to avoid inflating the figures[4].

In this work, we only consider ISCO codes that are 'fully' cultural (Table 1). In order to correctly include 'partly' cultural codes, we would indeed need to cross-tabulate four digits ISCO codes with three digits NACE codes. However, the dataset makes available only one or two-digits NACE codes for job postings.

Table 1. List of ISCO 08 codes that retain 'fully' to cultural occupations

ISCO code	Occupation
216	Architects, planners, surveyors and designers
2353	Other language teachers
2354	Other music teachers
2355	Other arts teachers
262	Librarians, archivists and curators
264	Authors, journalists and linguists
265	Creative and performing artists
3431	Photographers
3432	Interior designers and decorators
3433	Gallery, museum and library technicians
3435	Other artistic and cultural associate professionals
3521	Broadcasting and audio-visual technicians
4411	Library clerks
7312	Musical instrument makers and tuners
7313	Jewellery and precious-metal workers
7314	Potters and related workers
7315	Glass makers, cutters, grinders and finishers
7316	Sign writers, decorative painters, engravers and etchers
7317	Handicraft workers in wood, basketry and related materials
7318	Handicraft workers in textile, leather and related materials
7319	Handicraft workers not elsewhere classified

Source: Eurostat (2018).

The policy relevance of alternative data on cultural occupations is assessed using the OJA dataset (CEDEFOP, 2019), developed by the European Centre for the Development of Vocational Training (CEDEFOP)[5]. Data are

[3] NACE (Nomenclature générale des activités économiques dans les Communautés européennes') classifies the employer's main activity; ISCO is the International Standard Classification of Occupations.

[4] For more details on how 'partly' cultural codes are treated, see the *Guide to Eurostat culture statistics* (Eurostat 2018), p. 17.

[5] A more detailed account of the technical and methodological aspects of these data can be found in Alexandru and Necula (2020).

obtained through web scraping, crawling or API data collection from several sources including job search engines, public employment services, employment websites and agencies and online newspapers.

These data are still experimental. In this case, we have used version 12 of the dataset, but it is still under review and new versions with corrections and adjustments are regularly being released. In the latest version (at the time of analysis), observations run from the third quarter of 2018 to the second quarter of 2021, i.e. 3 entire years, and the data cover EU 27 and the United Kingdom. Updates are made on a quarterly basis.

Our research consisted of a first exploratory analysis to observe the statistical properties of the database, so the results reported here are mainly descriptive.

An important aspect to bear in mind before starting to work with the database is that it contains duplicate records. This is because the same advertisement with identical information can be collected through different sources. However, these cases can be identified through the general identifier (*general_id*) variable, as repeated records have the same identifier code. From the initial 310,414,865 records, we were left with 134,484,944 after cleaning up duplicates.

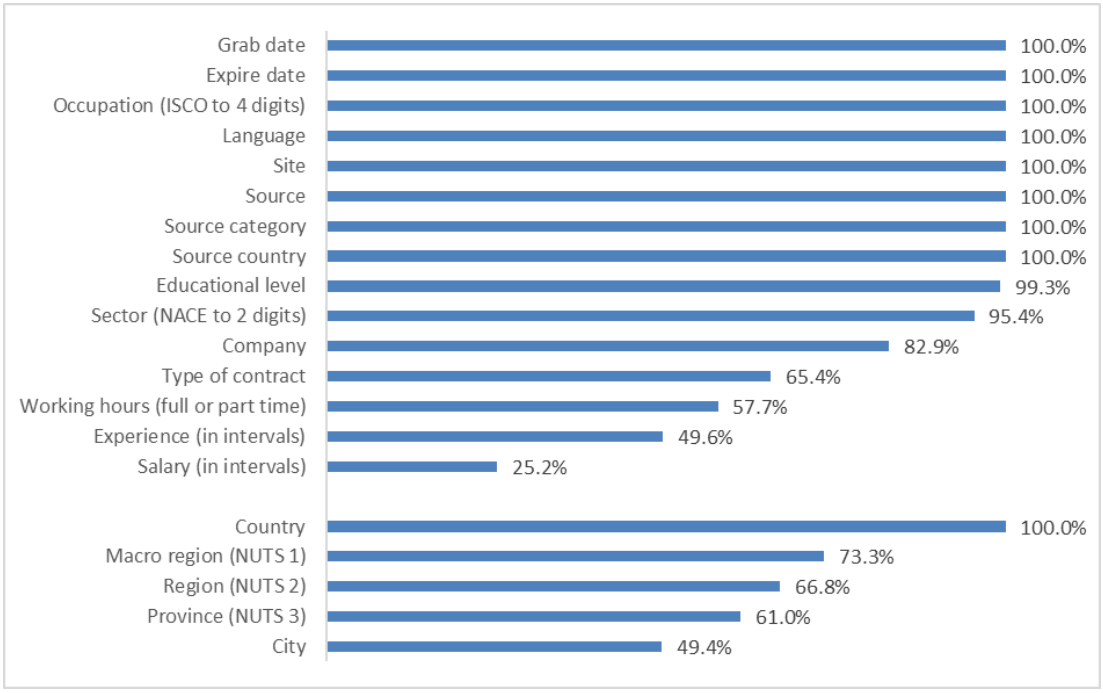
Each job advert record comes with a number of information:

- grab and expire date of the advert, and the language in which the advert is written;
- ISCO code of occupation (up to 4 digits);
- geographical variables (from country to city level);
- job characteristics such as type of contract (permanent, temporary, self-employed or internship), salary (in intervals) and working hours (full-time or part-time);
- job requirements such as educational level (according to ISCED classification) and experience (in time intervals);
- the sector (NACE Rev2 code up to 2 digits) and the company name; and
- the website where the advert is published (site) and the job portal where the vacancy is retrieved from (i.e., the aggregator portal is the source of the vacancy, the company website is the site where the vacancy is published. Not necessarily source and site of a vacancy coincide), along with the source category (job search engine, public employment service, employment website, employment agency, online newspaper, classified ads portal, company or education) and source country.

However, the data coverage is very much variable across job ads, as shown in Figure 2. For example, the 4-digit occupation variable has full coverage, which is the one we use to identify cultural occupations, but salary only appears in about a quarter of the advertisements, so that the analysis of salary data could lead to biased conclusions. For instance, it may be that jobs offering higher salaries are more likely to advertise this information to attract more applicants, while lower paid jobs may prefer to omit it in order not to deter applicants.

Regarding the territorial level, the country where the job ad has been published is always given, but the coverage decreases when moving to NUTS 1, 2 and 3 levels and it reaches about half of the coverage at the local level (LAU code). It should be noted that – at the time of the analysis presented in this report – the CEDEFOP database uses the outdated NUTS 2013 classification.

Figure 2. Valid data (%) of the variables in the database.



Source: Own elaboration based on OJA database.

Unlike official statistics, this database does not follow a probability sampling process. Therefore, it is to be expected that, when it is compared to Eurostat's data, differences in representativeness will appear in different countries, occupations, educational levels, etc., as we shall now see. Consequently, corrective factors should be applied before drawing direct conclusions from these data. Some work has previously analysed this database, pointing out some of the limitations and potential ways to overcome them (see e.g. Alexandru and Necula, 2020; Beręsewicz and Pater, 2021).

The method to be followed will be a comparative analysis based on a descriptive observation of the data stratified by variables (country, educational level, occupations and time) against official Eurostat employment data. This will allow us to assess the main statistical properties of the data and the extent to which they are representative. We are aware that the preferred option for assessing the statistical properties of job ads should be official statistics on job vacancies (i.e. flows) rather than employment (i.e. stocks). However, official job vacancy statistics have also some limitations. On the one hand, because Eurostat statistics on job vacancies are not detailed enough for our purposes (i.e. it is not possible to extract cultural occupations only). On the other hand, Cammeraat and Squicciarini (2021), with similar data for the United States, point out that comparing online job adverts with official vacancy statistics might not be a good idea. Although they start from the premise that this would be their preferred option, they conclude that the challenges that need to be overcome for such a comparison make it preferable to compare online job ads with employment data. The reasons given include, among others: the number of vacancies in each occupation and also in each country varies according to the frequency with which people change jobs; the measurement methodology for job vacancies varies by country; statistics on vacancies are considered in a period of time while adverts are considered only on the day they are posted; depending on the economic structure of the country, direct recruitment may be more important than in others (so that it is not counted as a vacancy). They argue that, in light of the constraints posed by the vacancy statistics, the comparison of OJA with employment statistics may be «even more relevant and/or accurate when addressing policy-relevant questions related to industry, innovation, employment and skills dynamics».

Comparing job advertisements with employment also has limitations. It assumes that there is correlation between two phenomena: flows and stocks of jobs. However, this is not necessarily the case. For example, if in an occupation the frequency with which people change jobs is low, there may be few job advertisements for that occupation, irrespective of the fact that there is a high level of employment in that occupation. Aware of this limitation, the comparative analysis can still provide useful insights as regards the potential of job ads to unveil timely changes in cultural occupations, as we shall see in the following pages.

3 Assessing the representativeness of job ads for cultural occupations

3.1 Country-level analysis

To investigate the representativeness of cultural occupations ads, cultural job openings figures are compared against cultural employment figures calculated by Eurostat using data from the EU-Labour Force Survey (EU-LFS, 2020). We do so as we are interested in understanding whether these data are representative enough for policy-relevant analyses.

To do this, we first assess data distribution by country. Table 2 summarises the available samples for each country (after deleting duplicities), providing the total number of ads, the number of ads for cultural occupations and its percentage share. The last two columns are the samples of adverts in relation to actual employment in each country. This is the annual average number of advertisements (i.e. the total divided by three) divided by the average total employment level of the country for the last three years (2018, 2019 and 2020). The penultimate column gives this figure for the entire sample of advertisements, while the last column is for cultural occupations ads only.

As shown in Table 2, the number of observations available for both total employment and cultural occupations is very diverse across countries, going from 30,530 (total ads) and 785 (cultural occupations ads) in Malta, up to 36,261,389 (total ads) and 1,037,895 (cultural occupations ads) in the UK – the country that features the highest number of ads for the two occupational categories. Naturally, this has a lot to do with the size of the countries. In order to really make comparisons regardless of size, we must take the last two columns, which are divided by the number of workers in each country. We see that in this case, the top-performer is still the UK (386.83 annual job advertisements per 1,000 workers), followed by France (344.21) and Belgium (305.39). While Greece (16.17), Romania (29.37) and Slovenia (39.93) are at the bottom. If we turn strictly to adverts for cultural occupations, the UK is again in the lead with 11.07 annual adverts per 1,000 workers, but followed this time by Germany (7.13) and Ireland (6.31). At the bottom are Greece (0.75), Romania (0.86) and Slovakia (1.04). It follows from these data that Western and Northern countries, and particularly the UK, are likely to have a greater presence of online platforms in the labour market[6]. In other words, job ads (also cultural ones) seem to be more to be posted online in the UK compared to all EU countries. However, in relative terms, Bulgaria has the highest share of cultural occupations ads over total job ads (6.02%), followed by Croatia (5.39%), Greece (4.65%) and Lithuania (4.02%). That is, in countries where fewer job adverts are posted online, a higher proportion of these adverts are for cultural occupations. This may seem counter-intuitive, but it might actually be indicative that in those economies, the cultural and creative sectors are leading the way compared to other sectors of the economy, and are taking the lead in leveraging digital platforms.

Table 2. Summary of job ads observations by country: total employment and cultural occupations (2018-2020).

Country	N total job ads	N cultural occupations ads	Share of cultural occupations (%)	Annual job ads per 1,000 workers	Annual job ads in cultural occupations per 1,000 workers
Austria	2,604,023	62,562	2.40	204.94	4.92
Belgium	4,339,837	74,600	1.72	305.39	5.25
Bulgaria	619,487	37,293	6.02	67.95	4.09
Croatia	222,249	11,975	5.39	45.26	2.44
Cyprus	51,954	1,933	3.72	43.37	1.61

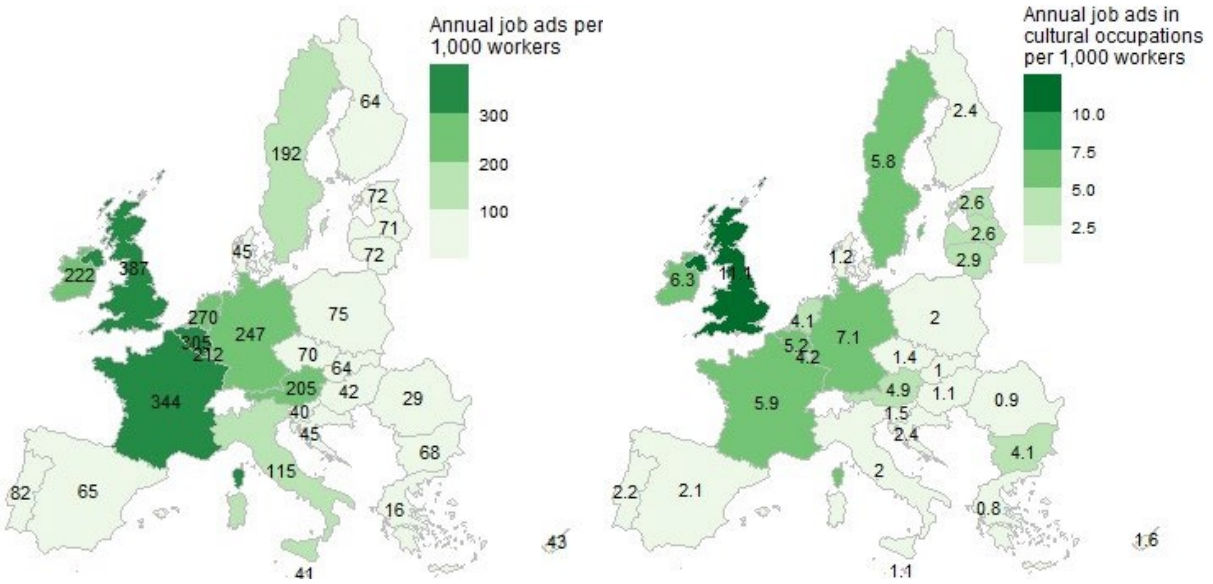
[6] Eurostat data on individuals' internet activity (isoc_ci_ac_i) largely confirm this impression. Relating annual job ads per 1,000 workers to the percentage of unemployed people who have used the internet for job search or sending an application, a fairly high Pearson's correlation coefficient of 0.53 is obtained (data from 2021, except 2019 for the UK). This implies that, generally, the countries with the highest number of adverts in the database are those where the use of digital platforms for job search is most widespread.

Czech Republic	1,080,355	21,324	1.97	70.22	1.39
Denmark	371,436	10,218	2.75	44.91	1.24
Estonia	134,809	4,852	3.60	71.55	2.58
Finland	473,726	17,549	3.70	64.00	2.37
France	27,771,830	476,350	1.72	344.21	5.90
Germany	30,255,509	871,564	2.88	247.41	7.13
Greece	183,675	8,542	4.65	16.17	0.75
Hungary	570,469	15,262	2.68	41.84	1.12
Ireland	1,470,559	41,764	2.84	222.07	6.31
Italy	7,762,617	132,723	1.71	115.01	1.97
Latvia	183,618	6,632	3.61	70.81	2.56
Lithuania	284,117	11,420	4.02	72.00	2.89
Luxembourg	181,039	3,590	1.98	211.74	4.20
Malta	30,530	785	2.57	41.31	1.06
Netherlands	6,986,350	105,398	1.51	269.61	4.07
Poland	3,616,493	96,862	2.68	74.91	2.01
Portugal	1,121,681	30,453	2.71	81.82	2.22
Romania	736,057	21,640	2.94	29.37	0.86
Slovakia	486,678	7,902	1.62	64.31	1.04
Slovenia	115,682	4,394	3.80	39.93	1.52
Spain	3,745,285	121,748	3.25	64.95	2.11
Sweden	2,823,490	86,014	3.05	191.93	5.85
United Kingdom	36,261,389	1,037,895	2.86	386.83	11.07
Total	134,484,944	3,323,244	2.47	199.29	4.92

Source: Own elaboration based on OJA database and Eurostat (lfsi_emp_a).

The distribution of data coverage by country can be better appreciated from the following maps. We see that the total, relative to the actual level of employment of the country, are very different (Figure 3). On the left, we consider all the adverts, while the map on the right represents only those for cultural occupations (i.e. the 5th and the 6th column of Table 2). In both cases, the UK has the largest sample and Greece the smallest, with a very significant difference. We go from 387 annual ads per 1,000 workers in the UK to only 16 in Greece or 11.1 versus 0.8 if we focus just on cultural ones.

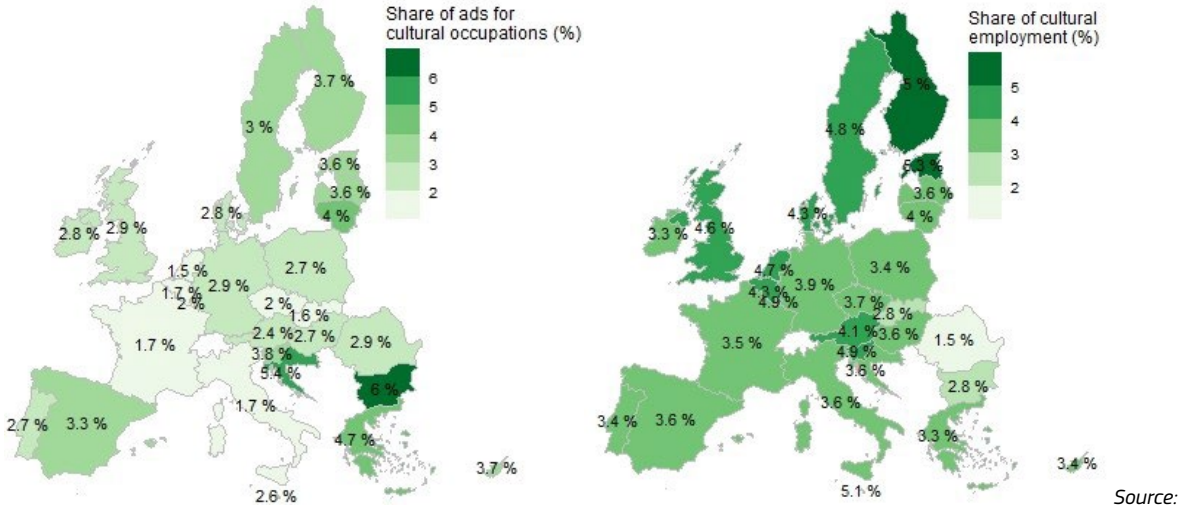
Figure 3. National samples of total job ads (left) and cultural occupation ads (right) relative to the total employment.



Source: Own elaboration based on OJA database and Eurostat (lfsi_emp_a).

Now, regardless of the size of observations by country, does the share of adverts for cultural occupations correlate with the share of cultural employment in each country? Not really (Pearson's correlation coefficient: -0.144, not significant), as we can see from the maps in Figure 4. There might be several reasons for that: first, the two aggregates are based on different definitions (e.g. Eurostat's cultural statistics include also non-cultural professionals working in cultural sectors, while here we are measuring only fully cultural occupations); second, some professions may have a higher likelihood to be advertised online, compared to others; third, there may be differences in the degree of development of the online job markets across countries.

Figure 4. Share of ads for cultural occupations in the database (left) and actual share of cultural employment according to Eurostat (right).



Source: Own elaboration based on OJA database (left) and Eurostat (cult_emp_sex) (right).

In what follows, we try to understand what may contribute to explain the observed differences in the 'representativeness' or coverage of (cultural) occupation data by job ads. To do so, we look at the level of education required for the advertised jobs as well as at data coverage for different (cultural) occupations and over time.

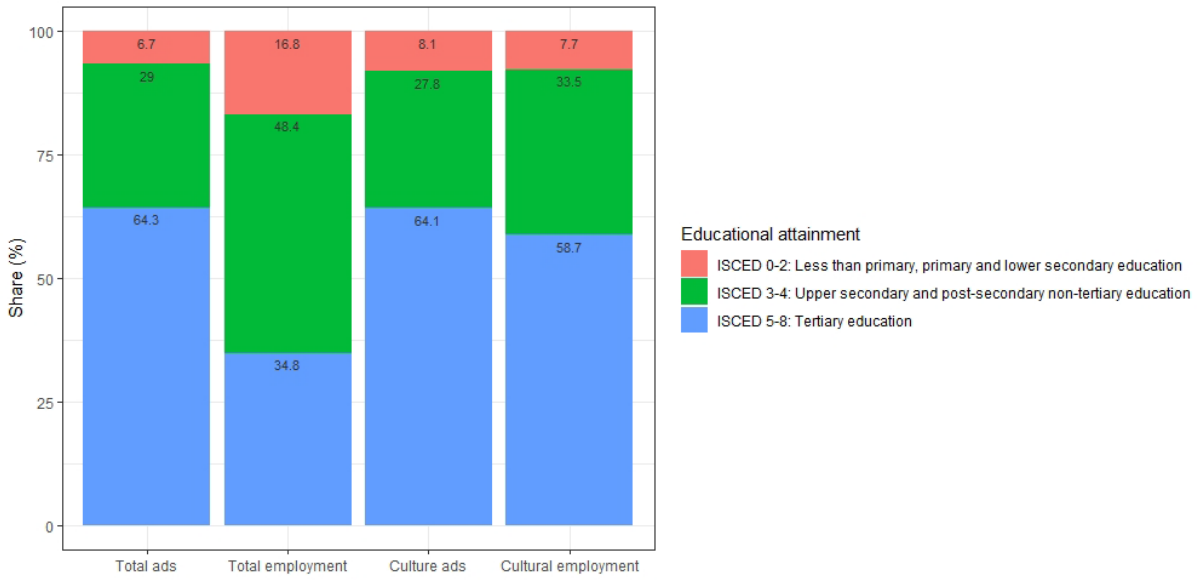
3.2 Educational level analysis

Figure 5 shows, in the first two columns, that jobs requiring tertiary education are much more represented compared to the share they represent of actual employment: job ads requiring a tertiary education degree are nearly double (64.3%) the share of workers with tertiary education in total employment (34.8%). This is very much in line with expectations as jobs requiring relatively higher levels of skills and more advanced educational background are more likely to be (also) posted online (Cammeraat & Squicciarini, 2021).

For cultural occupations, the share of job ads requiring highly educated workers is again higher (64.1%) compared to the share of highly educated cultural workers (58.7%), although in this case the difference is much smaller compared to total employment figures.

Thus, although there is a general bias in online advertisements towards jobs requiring higher education, given that cultural occupations also generally attract workers with high levels of education, the advertisements are in this sense more representative of cultural employment than of general employment.

Figure 5. Share of job ads and employment per educational attainment level, EU27.

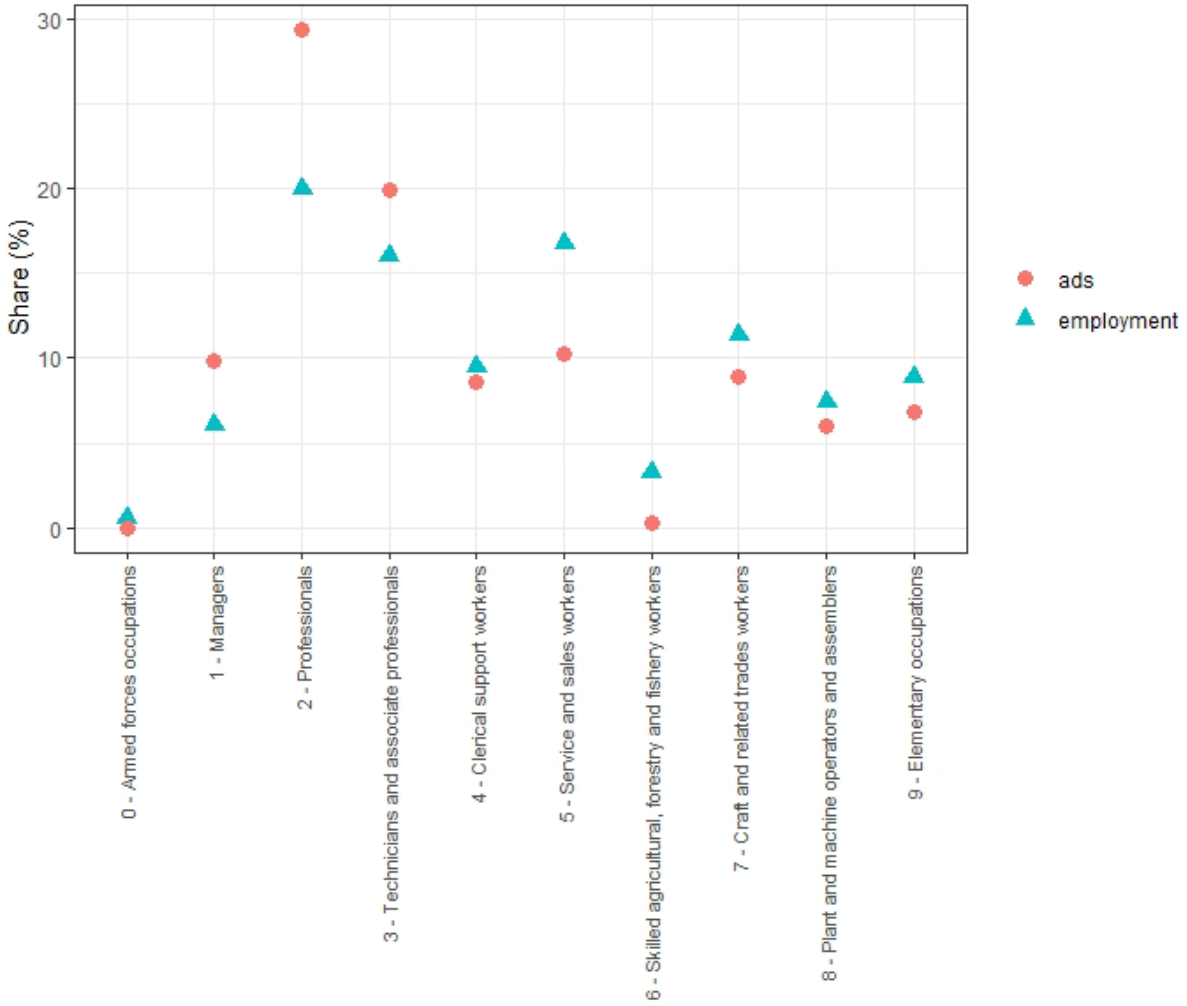


Source: Own elaboration based on OJA database and Eurostat (cult_emp_edu, lfsa_egaed).

3.3 Occupation-level analysis

If we look at how the database is distributed by occupations, and compare it with actual employment figures (Figure 6), over-represented occupations indeed coincide with those requiring the highest educational levels, namely: professionals, managers and technicians and associate professionals. The remaining occupations are all under-represented, especially service and sales workers.

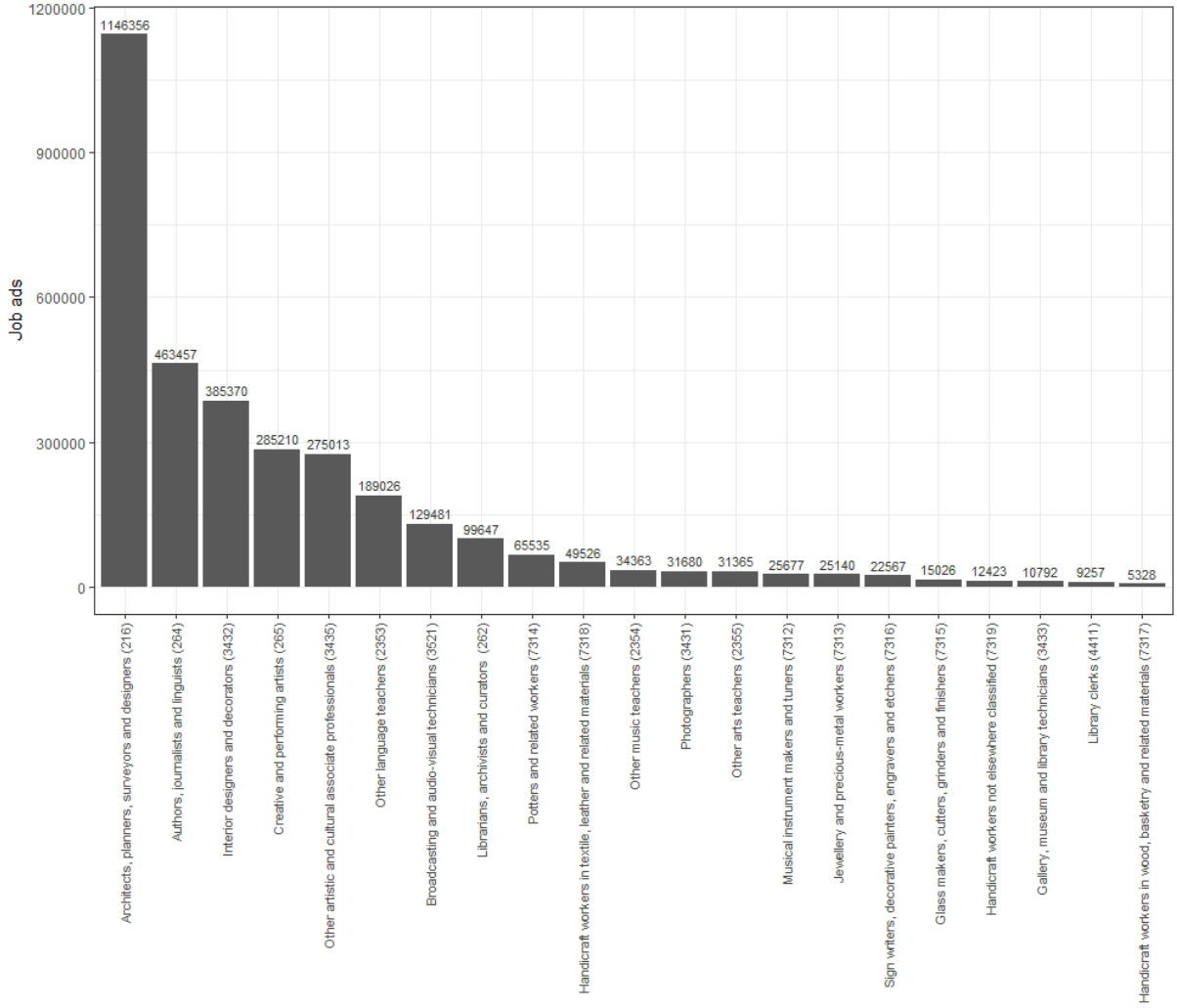
Figure 6. Share of job ads and actual employment (Eurostat) per occupation (ISCO 1 digit), EU27 + UK.



Source: Own elaboration based on OJA database and Eurostat (lfsq_egais).

Focusing on cultural occupations, and at a more detailed level, we see in Figure 7 that the group with by far the largest presence in the database is architects, planners, surveyors and designers (ISCO 216), with more than a million observations. This is followed by authors, journalists and linguists (ISCO 264); interior designers and decorators (ISCO 3432); and creative and performing artists (ISCO 265). Other occupations, e.g. Handicraft workers in wood, basketry and related materials (ISCO 7317), or library clerks (ISCO 9257), are barely present. Therefore, for analyses focused on certain occupations, online ads may be a good resource for occupations that are more likely to be advertised online; while for others, the potential of this tool will be more limited. This finding is in line with a research carried out by Nesta to study the skills required in creative occupations using online job ads data from Burning Glass Ltd. Due to the low number of job adverts for craft occupations workers, which the authors attribute to high levels of self-employment and the close nature of the professional community where recruitment is mainly conducted, these occupations were actually excluded from the analysis (Nesta, 2017).

Figure 7. Amount of job ads per cultural occupation, EU27 + UK.



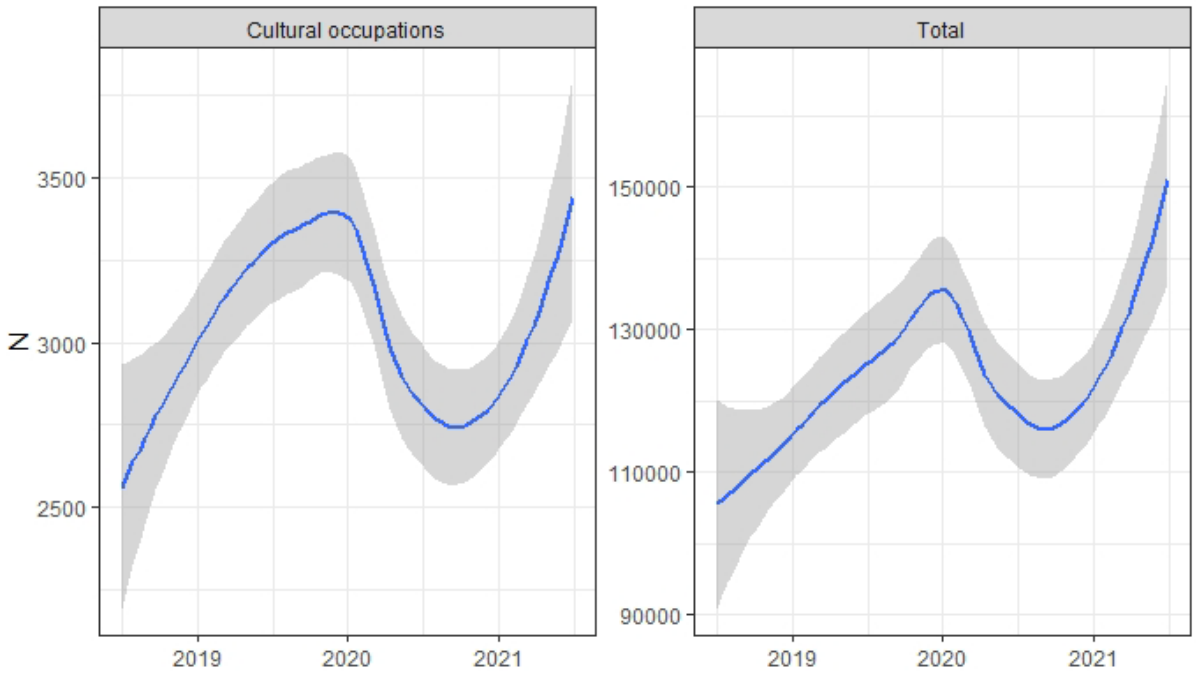
Source: Own elaboration based on OJA database.

3.4 Over time analysis

As regards the analysis of job ads over time, before going to the results, it is important to point out to one important feature that characterises these data. The grab date variable indeed records the day when the advert was collected (i.e. when it was added to the database), not the day it was posted, except for a few cases for which the posting date is present in the advert. This implies that no immediate conclusions should be drawn about the evolution of the publications because months might have passed before the grab date. However, some general trends can be observed, as will be described below.

Figure 8 shows that both total ads and those for cultural occupations in general follow a similar pattern, with very strong growth up to 2020 followed by a very steep decline from early 2020, right before the COVID-19 pandemic outbreak.

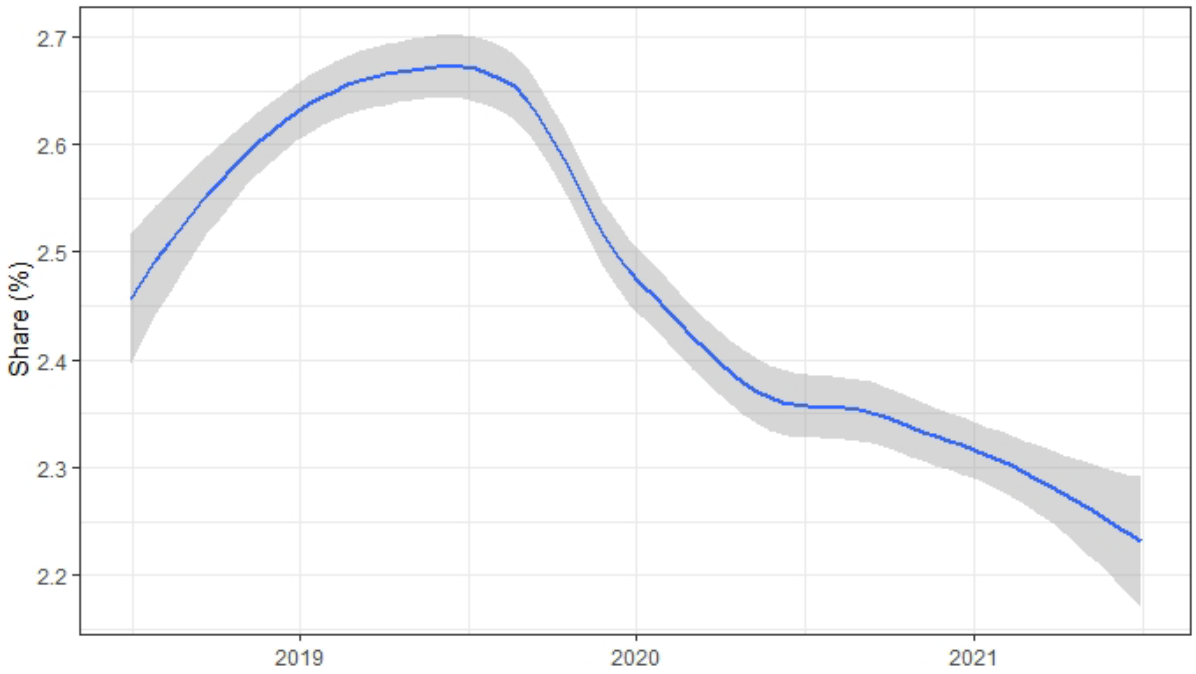
Figure 8. Smoothed trend in the number of job ads, EU27 + UK.



Source: Own elaboration based on OJA database. *Note:* The trend has been smoothed by applying locally estimated scatterplot smoothing (LOESS). The shaded area represents a 95% confidence interval.

The beginning of the year 2020, and even a bit earlier, also coincides with a decrease in the share of cultural occupations in the total number of job ads, which does not recover after Covid-19 and is still about half a point lower at the end, as can be seen in Figure 9.

Figure 9. Share of ads for cultural occupations across time (smoothed), EU27 + UK.

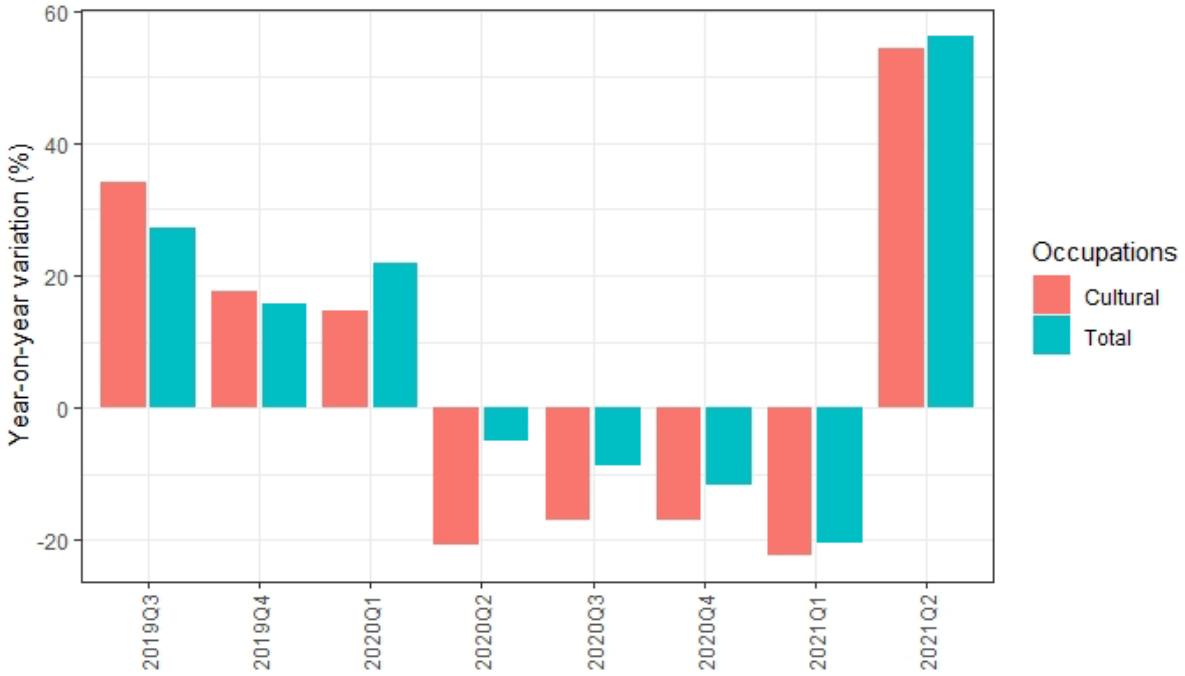


Source: Own elaboration based on OJA database. *Note:* The trend has been smoothed by applying locally estimated scatterplot smoothing (LOESS). The shaded area represents a 95% confidence interval.

These patterns become clearer if we look at them in terms of year-on-year growth by quarter. We see, in Figure 10, that the growth rates are positive, both in cultural occupations and in the total, until it changes in the second quarter of 2020 with the irruption of Covid-19. Then the rates remain negative until the first quarter of 2021, especially in cultural occupations (which is consistent with the greater impact of the pandemic on cultural employment), and finally there is a rapid recovery in the second quarter of 2021.

One aspect the reader may have noticed is that growth rates are high. For instance, in the third quarter of 2019, before the pandemic, annual growth rates were around 30%. Rather than reflecting actual growth rates, this is likely to reflect an increasing use of online platforms by companies to publicise job ads.

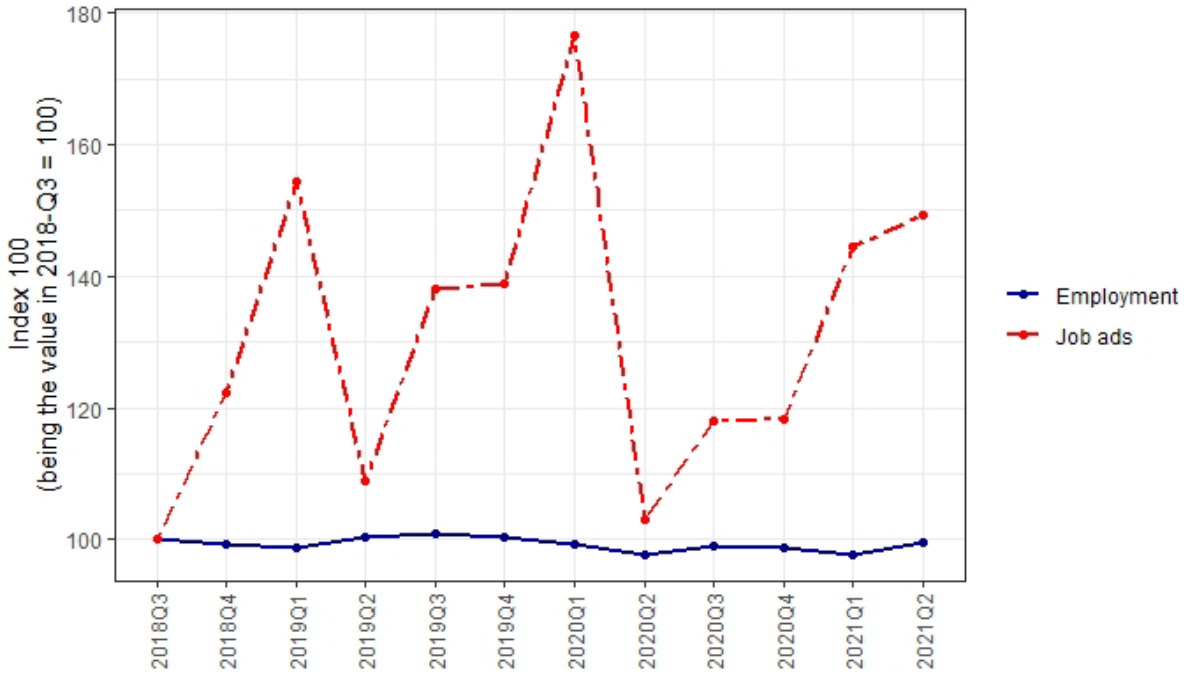
Figure 10. Year-on-year change in job ads by quarter, EU27 + UK.



Source: Own elaboration based on OJA database.

We can also compare the quarterly evolution of the number of job ads with that of actual employment, taking index numbers with reference to the first quarter of the database (Figure 11). It is easy to see that, compared to the level of employment, the evolution of job ads is much more unstable, with extraordinarily high growth rates, but also very notable falls. This is to be expected, as advertisements are related to job flows and therefore increase or decrease much faster than employment (stock). There has also been a growth, albeit unstable, in the number of announcements since the initial quarter.

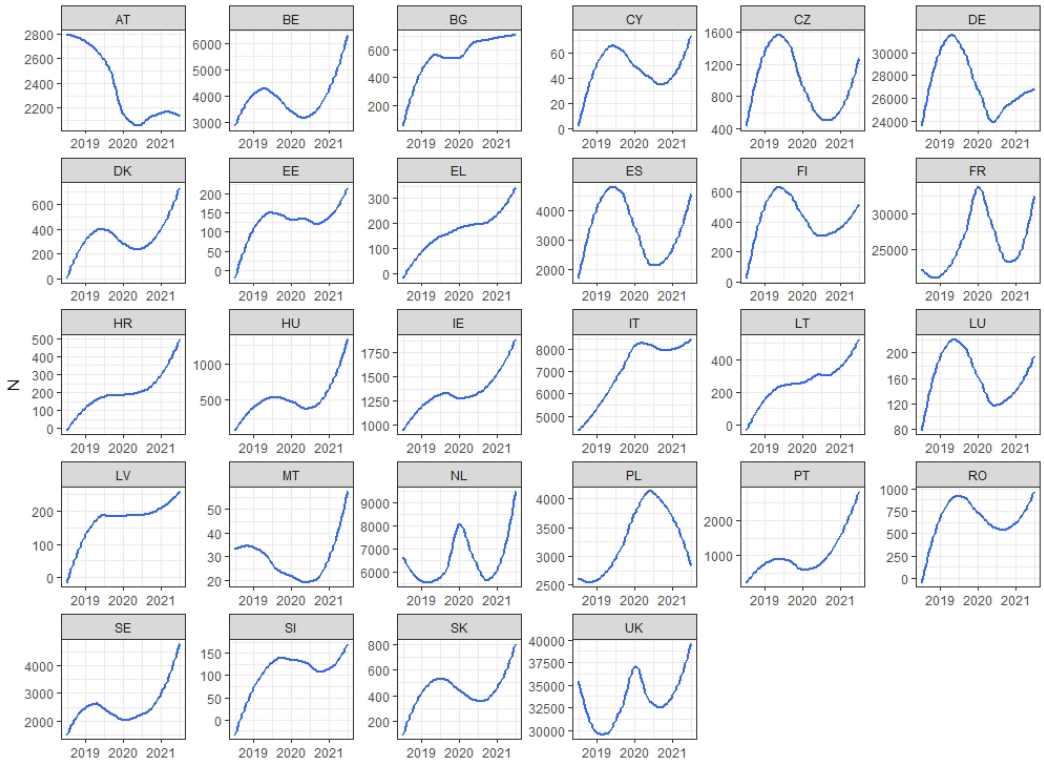
Figure 11. Evolution, in index numbers, of the number of total job ads and the actual level of employment (from Eurostat) by quarter, EU27.



Source: Own elaboration based on OJA database and Eurostat (lfsi_emp_q).

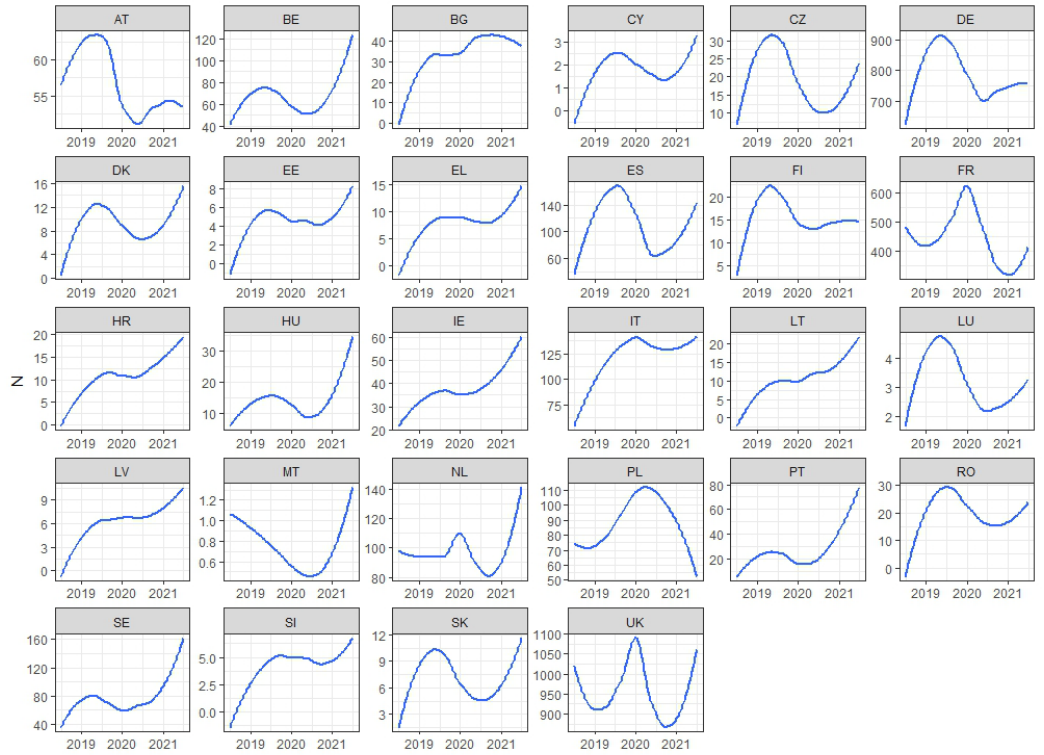
Finally, when analysing trends by country (where the sample is logically smaller), it is important to be very cautious. It should be borne in mind that what is being measured is the date on which the advertisements are collected, and not when they are published. This is probably the reason why, in some cases, significant peaks or trajectories might seem unlikely, both for all job ads (Figure 12) and for cultural occupations (Figure 13). However, generally speaking, upward trajectories are observed, confirming the general trend mentioned above. Stagnation (e.g. Italy or Slovenia) or significant falls (e.g. France or the UK) are also observed in several countries in the early 2020. This seems to confirm the remarkable impact that Covid-19 had on the European labour market, especially during the months of lockdown.

Figure 12. Smoothed trend in the number of total job ads by country.



Source: Own elaboration based on OJA database.

Figure 13. Smoothed trend in the number of job ads for cultural occupations by country.



Source: Own elaboration based on OJA database.

4 Conclusions

This report illustrates the main results of the analysis that we have carried out on CEDEFOP's job adverts database, with a focus on cultural occupations vacancies.

First, the data has been cleaned avoiding a very high number of duplicate records. Next, a first exploration of the data has revealed that the exploitation of some potentially interesting variables to be analysed could be constrained due to low coverage rates (e.g. the salary, which is only available for about a quarter of the records).

We then looked at the number of observations available for each available EU country and the UK and found that there are significant differences. Particularly notable are the western and northern countries, for which a considerably higher number of observations is available. This is likely due to the more advanced state of their digital economies and a greater use of online advertisements in their respective labour markets.

Next, a greater presence in the database of certain occupations (e.g. professionals, managers or technicians) than others has been observed, with architects, planners, surveyors and designers (ISCO 216) standing out in the case of cultural occupations. On the other hand, occupations that by their nature are hardly advertised online (e.g. those where self-employment is predominant, such as those related to crafts), are poorly represented.

Jobs requiring high levels of education are also relatively more represented compared to those requing less educated workers. CEDEFOP data could thus relatively well reflect the breakdown of cultural workers by educational level, considering that cultural occupations are generally occupied by more educated workers.

Finally, with regard to the time dimension, a general upward trend in the online job advertisements, as well as the sinking effect of Covid-19 since the beginning of 2020 can be observed. However, special care must be taken in drawing conclusions about changes of employment –or even job vacancies- over time, as the grab date of the adverts does not generally correspond the date of their online publication and it is not always possible to know both dates for each job ad.

As a whole, cultural occupation data from job ads are currently very different from official statistics on cultural employment. This finding is actually in line with previous research that has compared employment statistics and job ads for the general economy (Cammeraat and Squicciarini, 2021). One explanation lies in the fact that job ads data do not refer directly to employment as a stock, but to job vacancies as a flow. There are important differences between these concepts, ranging from the frequency of job switching to the level of unemployment (which can vary greatly both across countries and across occupations and sectors). Therefore, it is not surprising to find different patterns, and in particular a higher volatility of job vacancies. For instance, in times of economic uncertainty, new job vacancies can be expected to disappear faster than existing employment.

Nevertheless, the CEDEFOP data remain promising and their potential as a supplement to official statistics vast. To further test their potential applications for policy-relevant work, a number of possible research venues have been identified.

First, while the structure of the data shows that reliable direct extrapolations on the performance of cultural employment are by no means possible, it would be worth exploring ways to overcome this problem through corrective factors, such as the application of weights related to the actual share of cultural jobs at national level, as suggested by Cammeraat and Squicciarini (2021). However, in order to find appropriate weights for cultural occupations, more granular employment data from the EU-LFS data survey should be collected by Eurostat.

Second, although the representativeness of occupations that are less likely to be advertised online is limited, studies about job vacancies could be carried out with a focus on those occupations that are more likely to be present in online labour markets (e.g. architects). Indicators can also be explored to compare online labour markets across countries, rather than general employment trends.

Third, variables concerning working hours (full-time or part-time) and type of contract (temporary, permanent, traineeship or self-employment) could also be exploited, observing differences in job vacancies between occupations and between countries. In particular, the sister database to the one used here, i.e. the skills database, can also provide valuable analysis[7]. For example, to compare the skills required in different occupations; or to assess whether the same occupations require different skills in different countries.

[7] For further details, see https://www.cedefop.europa.eu/en/projects/skills-online-job-advertisements

This analysis has enabled us to outline some insights about the CEDEFOP database and its usability for the analysis of cultural employment trends. Most of the considerations drawn through this analysis (for instance on time trends) could support any researcher interested in understanding the structure and potential uses of these data. While, at this stage, these data remain experimental, it should be borne in mind that online job adverts will definitely continue to grow. Therefore, this database, like other similar ones that may emerge, is very likely to gradually improve its coverage and representativeness of the labour market in the future. Although there are still considerable challenges to be addressed, its potential for research and application in the policy-making arena remains to be further explored as long as refinements are made.

References

Alexandru, C. and Necula, M., Methodological aspects related to CEDEFOP data used for OJA statistics V.1., 2020

Beręsewicz, M., & Pater, R., *Inferring job vacancies from online job advertisements – 2021 edition*, Luxembourg: Publications Office of the European Union, 2021, doi: 10.2785/963837.

Cammeraat, E. and Squicciarini, M., *Burning Glass Technologies' data use in policy-relevant analysis: An occupation-level assessment*, OECD Science, Technology and Industry Working Papers, 2021, doi: 10.1787/cd75c3e7-en

CEDEFOP, *The online job vacancy market in the EU: Driving forces and emerging trends*, Cedefop research paper No. 72, Luxembourg: Publications Office, 2019, doi: 10.2801/16675

European Commission, Green Paper: unlocking the potential of cultural and creative industries. Brussels, 2021.

European Commission, A New European Agenda for Culture, 2018.

Eurostat, *Guide to Eurostat culture statistics – 2018 edition*, Publications Office of the European Union, Luxembourg, 2018, doi: 10.2785/45136

Eurostat. *Culture statistics. 2019 edition,* Publications Office of the European Union, Luxembourg, 2019, doi: 10.2785/118217

EY, Rebuilding Europe. The cultural and creative economy before and after the COVID-19 crisis, 2021.

Giabelli, A., Malandri, L., Mercorio, F., Mezzanzanica, M., & Seveso, A. (2021). NEO: A System for Identifying New Emerging Occupation from Job Ads. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(18), 16035-16037. https://doi.org/10.1609/aaai.v35i18.18004

Giambona, F., Khalawi, A., Buzzigoli, L., Grassini, L., & Martelli, C. (2021, December). Big data analysis and labour market: an analysis of Italian online job vacancies data. In *ASA 2021 Statistics and Information Systems for Policy Evaluation: BOOK OF SHORT PAPERS of the on-site conference* (Vol. 132). Firenze University Press.

Marconi, G. & Vergolini, L., The Demand for Language Skills in the European Labour Market: Evidence From Online Job Ads (July 2, 2022). Available at SSRN: https://ssrn.com/abstract=4152463 or http://dx.doi.org/10.2139/ssrn.4152463

NESTA, *A closer look at Creatives. Using job adverts to identify the skill needs of creative talent*, 2017, https://data-viz.nesta.org.uk/creative-skills/index.html

OECD, *The Culture Fix. Creative People, Places and Industries*, Local Economic and Employment Development (LEED), OECD Publishing, Paris, 2022, doi: 10.1787/991bb520-en

Throsby, D., 'The concentric circles model of the cultural industries', *Cultural trends*, Vol. 17, No 3, 2008, pp. 147-164, doi: 10.1080/09548960802361951

List of figures

List of tables

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

— by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

— at the following standard number: +32 22999696,

— via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

Open data from the EU

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
joint-research-centre.ec.europa.eu

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

EU Science

Publications Office
of the European Union