**Estimating French labour market tightness using OJAs**

Since the end of 2017, French NSI’s Labour Department (Dares) and French National Employment Agency (NEA - *Pôle emploi*) have been working on a new version of the French labour market tightness indicator. Among other changes, new tightness indicators will include online job advertisements (OJAs).

***Using online data to improve statistical outputs***

From 1998 until mid-2017, French quarterly statistics on labour market tightness were published [1]. The main indicator computed was the number of new job offers relayed by the French NEA over the number of new jobseekers registered in the French NEA, broken down by occupation. Historical labour market tightness statistics were then computed using only job offers relayed by NEA. A newer version of this indicator, including online job advertisements (OJAs), is to be published by the end of 2020. It will be an annual statistics provided by region (over 100 French “departements”) and occupation, at a lower level of granularity (186 occupations instead of 77).

The initial purpose of including OJAs was to provide a more robust tightness indicator with a better coverage of job offers. Indeed, as mentioned in the ESSnet Big Data no.1 Final Technical Report [2], “coverage of job ads by *Pôle emploi* may fluctuate due to changes [...] of online job ads providers.” For example, a website closure or change in its business model (eg. charge per ad) could lead to an increase in the number of job offers reported on the French NEAwebsite [3]. Moreover, a French survey carried out in 2016 [4] showed that NEA’s offers are not representative of job offers. Thus, diversifying sources is a strategy to obtain a more reliable labour demand picture.

***From raw data to statistical sources***

As explained in the ESSnet Big Data no.1 Final Technical Report[2],Dares collected online job offers through scraping from a sample of 14 websites. This includes specialized as well as generalist job boards and one aggregator. These chosen websites were scraped on a daily basis from late 2018 onwards[[1]](#footnote-1). In 2019, the number of online job offers collected by Dares amounted to more than 8.5 million.

Once collected, online job offers go through the following process:

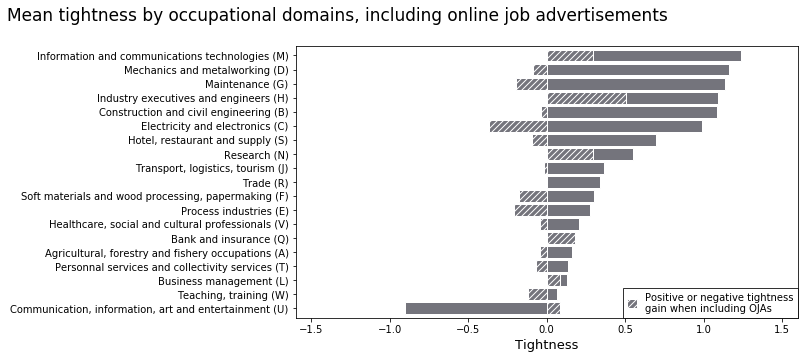
* fields harmonization across websites
* location recoding using external database of city names and codes
* job offers titles processing and cleaning
* occupation categorization (into the French professions’ nomenclature, ROME) with a SVM model trained on ~1,558,800 annotated job offers titles.
* duplicates removal

At the end of this process, resulting job offers are then added to the job offers relayed by NEA. However, due to remaining statistical sources mismatch, OJAs and NEA offers are not yet deduplicated between them. Finally, they are combined to other data sources (number of jobseekers, labour force needs...) to estimate labour tightness broken down by region and/or occupation.

***First results***

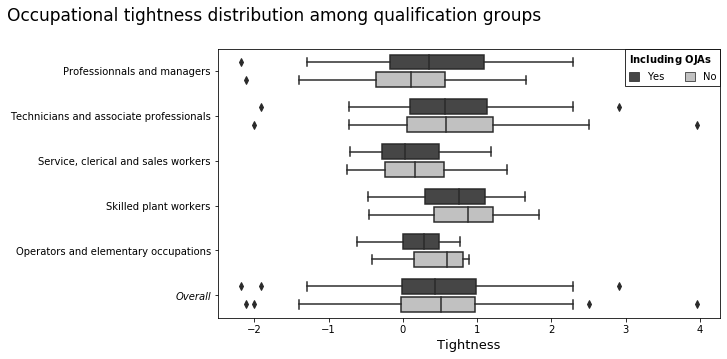
Although this new version of the French labour market tightness indicator is not finalised yet[[2]](#footnote-2), some preliminary results are shown below. While tightness is calculated by occupation at the most detailed level, we focus on aggregated results by occupation groups or qualification[[3]](#footnote-3).

When looking at the average tightness by occupation groups (according to French FAP nomenclature occupational classification system), half of them are slightly, or not impacted by the addition of OJAs (+/- 0.1, see Fig. 1). Only three occupation groups gain at least 0.3 points on the tightness scale: information and communication technologies (+0.30), research (+0.30) and industry executives and engineers (+0.51). A significant reduction of labour tightness (higher than 0.3 on the tightness scale) is noted for only one occupation group: electricity and electronics (-0.37).

**Figure 1: Mean tightness by occupation groups, including OJAs**

It appears that most impacted groups are either groups with highly qualified occupations or, on the contrary, relatively lowly qualified occupations. Occupational tightness distribution broken down by qualification corroborates this finding (Fig. 2). Overall, labour tightness in upper qualified occupations increases when adding OJAs (+0.28), whereas lower qualified occupations’ tightness decreases (-0.25). Thus, it seems that adding OJAs to NEA’s offers modifies the distribution of labour demand in favour of more qualified occupations, leading as a result to an intensification of labour tightness in these occupations. This result is consistent with that obtained by Dares [3] which shows that 20% of recruitments for executive jobs are advertised online when only 5% of vacancies in elementary occupations are. On the contrary, elementary occupations are over-represented in French national employment agency’s job offers.

***Figure 2: Tightness distribution among qualification groups***



***Possible insights***

So far, the main insight remains the consolidation of job offers’ handling process (localisation, classification, company recoding and deduplication). Apart from technical issues, comparison with administrative sources will eventually also give a better understanding on how online job offers reflects labour market demand and possibly helps adjusting OJAs weights in tightness computation.

Finally, although many improvements are to be made, this publication shows the technical feasibility of integrating data from the Internet into statistical indicators computation and paves the way for other experiments using big data and online sources at Dares.

***References***

[1] Bergeat M. (2017), “Les tensions sur le marché du travail au 2e trimestre 2017”, *Dares indicateurs n°056/2017*, Dares.

[2] Swier N. and al. (2018), “Web scraping / Job vacancies Deliverable 2.2 Final Technical Report (SGA-2)”, Eurostat.

[3] Niang, M. (2020, forthcoming), “Comment mesurer les tensions sur le marché du travail en France ?”, *Document d’études*, Dares.

[4] Bergeat M., Rémy V. (2017), “Comment les employeurs recrutent-ils leurs salariés ?”*, Dares analyses n° 064/2017*, Dares.

1. For some websites, however, data collection by scraping only began in 2019. [↑](#footnote-ref-1)
2. The publication of this new indicator of labour market tightness is due on September 2020. [↑](#footnote-ref-2)
3. Due to this report format, detailed results are not presented but will be published by Dares in the incoming months. [↑](#footnote-ref-3)