KOF
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

UNIL | Université de Lausanne
HEC Lausanne

77 NRP
Digital Transformation
National Research Programme

Swiss Job Tracker: New Weekly Indices of Online Job Postings in Switzerland

Methodology

Dr. Matthias Bannert, KOF Swiss Economic Institute, ETH Zürich
Jeremias Klaeui, Department of Economics, University of Lausanne
Dr. Daniel Kopp, KOF Swiss Economic Institute, ETH Zürich
Dr. Michael Siegenthaler, KOF Swiss Economic Institute, ETH Zürich,
Severin Thöni, KOF Swiss Economic Institute, ETH Zürich
Joseph Winkelmann, Harvard Business School, Harvard University

Corresponding author: Michael Siegenthaler, siegenthaler@kof.ethz.ch
The associated open-source programs and codes are published here:
https://github.com/swissjobtracker

November 2022

Content

1. Introduction

Swiss Job Tracker is a new labor market dashboard that publishes time series aggregated from large-scale online data. The dashboard contains indices on the development of unique online job postings in Switzerland. The underlying data represent the near-universe of job postings in Switzerland published on job boards and firm websites. These data are collected by the private company x28 AG and cover at least 72% of all vacancies and at least 90% of all posted vacancies.

We provide these new job indices to the interested public on an interactive dashboard, *www.swissjobtracker.ch*. The time series are published weekly in the form of indices. Hence, in contrast to extant job indices and indicators of the state of the job market, the indices allow monitoring the state of the Swiss job market at a very high temporal frequency. The job indices can additionally be broken down by industry, broad occupation, and canton.

The indices complement the "Job Radar" by x28 by focusing on high-frequency changes over time. The "Job Radar" shows the stocks and composition of the vacancies collected by x28 on a quarterly basis.

The purpose of this methodological note is to document how we process the raw vacancy data and construct the indices of online job postings published on Swiss Job Tracker. We complement this methodological document with the source code of the calculation and the dashboard.

Swiss Job Tracker was created as part of the project "What Workers Want: Determinants and Implications of Job Search Strategies on an Online Job Platform" within the Swiss National Science Foundation's (SNSF) National Research Program on "Digital Transformation" (NRP-77). The dashboard resulted from a collaboration between members of the KOF Swiss Economic Institute of ETH Zurich, the University of Lausanne, and the private company x28 AG.

The code repository for calculating the job indices and of the dashboard www.swissjobtracker.ch is published on an open-source repository: **https://github.com/swissjobtracker**

2. Data and coverage

The raw data that we use to calculate indices of online job postings comes from x28. x28 is specialized in collecting and synthesizing data from online job openings posted in Switzerland. x28 continuously crawls job postings from all major online job boards in Switzerland and from company websites. x28 also flags duplicate job postings and systematizes each job posting by assigning it to official industry and occupation classifications. These data form the raw material for the indices of unique job postings published on Swiss Job Tracker. We access the data via the advanced programming interface (API) of x28 each week.

By design of the data source, the indices reflect the evolution of job openings that are either posted on online job boards or on firm websites. The new indices thus miss out job openings that are not published online. However, most vacancies are nowadays posted online. According to a 2017 company survey, firm's own internet site is the most widespread channel among all channels to post job openings, covering around two-thirds of all vacancies.[1] According to the same survey, nearly 60 percent of all vacancies are posted (sometimes additionally) on online job boards. According to Swiss firm data covering the current sample period of our indices (2017-2022), the share of vacancies posted online either via the firm website or a job board is 80% of all vacancies.[2] This is a very high number, which reflects that the "hidden" labor market - vacancies that are filled via informal channels such as the personal or social network - is rather small in Switzerland.

[1] Buchs, H., & Buchmann, M. (2018). Verdeckter Arbeitsmarkt in der Schweiz ist eher klein. *Die Volkswirtschaft*, *11*, 39-41.

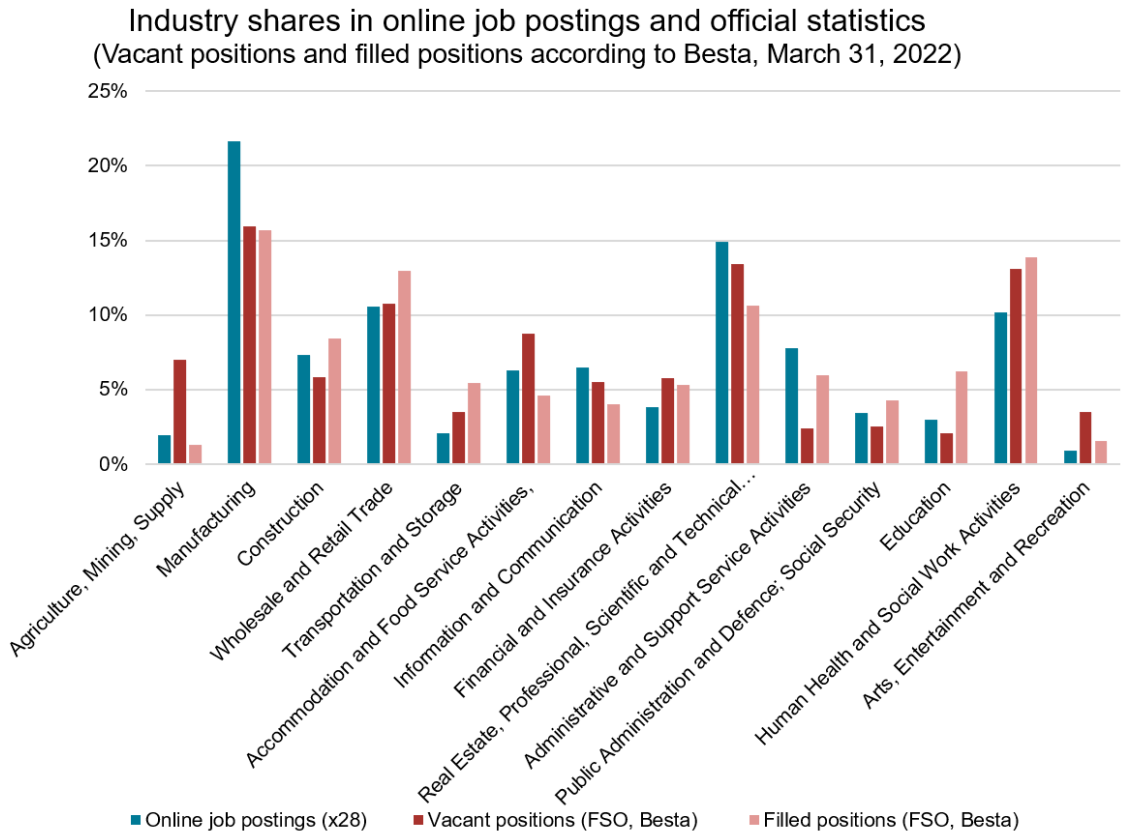
[2] These number are taken from the documentation of the Swiss labor market monitor, see Buchmann, Marlis, Helen Buchs, Eva Bühlmann, Felix Busch, Ann-Sophie Gnehm, Debra Hevenstone, Yanik Kipfer, Urs Klarer, Jan Müller, Marianne Müller, Stefan Sacchi, Alexander Salvisberg & Anna von Ow (2022). Stellenmarkt-Monitor Schweiz 1950 – 2021 [Dataset]. FORS, Lausanne.

What is the share of the vacancies posted online that are covered by the x28 data? An answer to this question is provided by a benchmarking report conducted by the Stellenmarkt-Monitor team on behalf of x28.[3] The report analyzed the data coverage in the year 2017 (Stellenmarkt-Monitor Schweiz, 2018). The report finds that 91.5% of all vacancies posted online are also recovered by x28. Overall, we expect that the data covers at least 90% of all vacancies *posted online* and approximately 90%*80% = 72% of all vacancies in Switzerland.

We now benchmark the industry and regional composition of job openings in the x28 data against official labor market data from the Federal Statistical Office of Switzerland (SFSO). We use the number of open positions (vacancies) and the number of filled positions (jobs) according to the most recent jobs statistics (Besta) as the benchmark. The Besta numbers are calculated from a large-scale, representative survey of firms conducted at the end of each quarter by the SFSO. Due to the oversampling of large firms, the survey covers approximately 45% of total employment in Switzerland. Conforming with the reference date the Besta survey, we compute the stock of online job openings in the x28 data on March 31, 2022.

Figure 1 shows the share of an industry's job postings in the total number of job postings in the x28 data and likewise for the number of open and filled positions according to Besta. The figure reveals a close correspondence between the industrial composition in the x28 vacancy data and the industry composition in official labor market data. The shares of the industries in the total economy are highly correlated. For instance, the correlation between the shares of online job openings and the share of filled positions is 0.89.

Figure 1: Industry coverage of x28 vacancy data

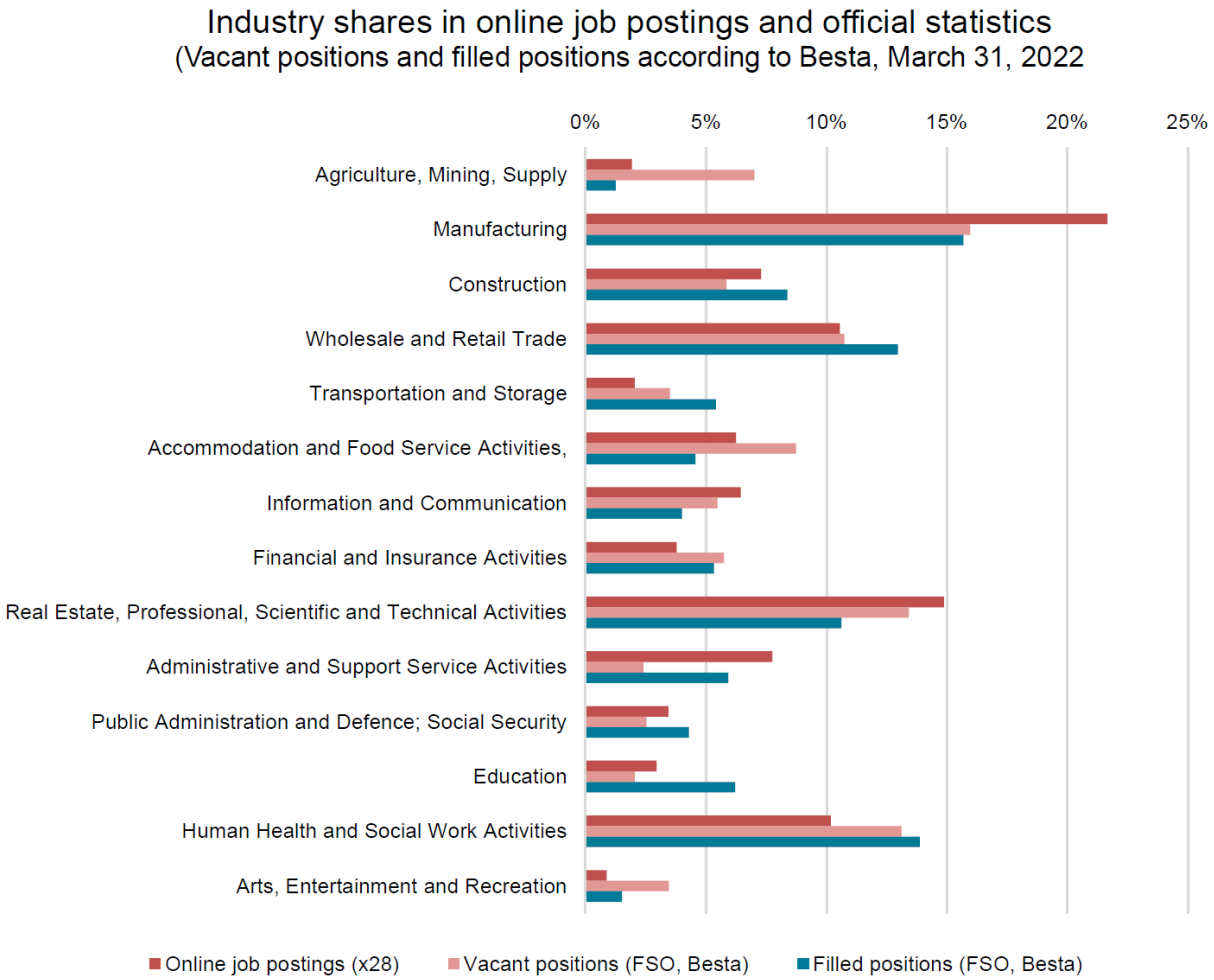


[3] Stellenmarkt-Monitor Schweiz (2018). Stellenausschreibungen auf den Webseiten von Schweizer Unternehmen. Abschätzen der Ausschöpfung durch die x28-Erhebungs-Technologie, Universität Zürich, Zürich.

Figure 2 shows the regional composition of online job postings according to the x28 data and vacant filled positions according to the Besta in March 2022. In the x28 data, the same vacancy may be assigned to several NUTS-II regions. Again, the figure reveals a close correspondence between the x28 data and the official labor market data. The data may undercover the Italian-speaking and, less clearly, the French-speaking region of Switzerland.

The results of this benchmarking exercise can also be compared to ascientific report on the coverage of the x28 data conducted by the Stellenmarkt-Monitor team (Stellenmarkt-Monitor Schweiz, 2018, cited above). This report also suggests that the coverage of the x28 is very high both in general and in all major industries, occupations, and regions. The report suggests that the x28 data may have some holes regarding vacancies from small firms (25 employees or less).

Figure 2: Regional coverage of x28 vacancy data



Importantly, the fact that x28 may not always capture all vacancies from all firms and regions is not an issue for our indices per se. The indices focus on the change in the number of vacancies over time. It is possible that the over-time changes in job openings are well captured by the x28 data although the stock may be somewhat too small.

In sum, we expect that the indices of job openings published on Swiss Job Tracker paint a reliable and informative picture of longer-term trends and short-term movements of the online job market in Switzerland. We also expect that the new indices provide a reliable insights on the newest developments on the Swiss job market at large. After all, the indices capture a sizeable share - likely more than 72% - of all unique vacancies in the Swiss labor market. Moreover, it is quite likely that job openings that are published online and the ones published exlusively on other channels follow common movements. Indeed, below we document a close correspondence between our new indices of online job postings and the official indices of job postings from Besta. The advantage of our data

relative to the official statistics is that we publish these results in real time and hence several months before the publication of the Besta data.

3. Index construction

3.1 Sample selection

Before computing the indices, we conduct two cleaning steps.

1. We remove duplicates in the data flagged by x28. The duplicates are detected by x28 based on technical "fingerprints" of job postings such as identifier numbers.
2. We remove all vacancy postings from recruitment agencies (using a flag set by x28). Vacancies by recruitment agencies represent a disproportionately large share of all online job postings (up to half of all job openings in certain quarters), largely overrepresenting the share of workers employed at temporary employment agencies in total employment (2-3% in the last quarters). One explanation for this discrepancy is that recruitment agencies recruit for another firm (but we typically do not know for which one).[4] More generally, recruitment agencies hire differently than companies that post their vacancies directly. Indeed, vacancies posted by recruitment agencies follow quite erratic time series patterns. We thus delete vacancies by recruitment agencies to avoid double-counting of vacancies and to ensure that our indices are not biased by recruiting trends in such agencies. Removing these ads also circuments the problem that we often cannot assign vacancies from recruitment agencies to the industries and regions in which the recruited workers would work. Including them would generate a (large) discrepancy between the aggregate indicator and the subindices.

Apart from that, we include all vacancies where the place of work is within Switzerland. Hence, we incorporate all vacancies that were not deleted until the reference day. This reflects our approach to compute an index of all job openings currently posted online. One concern with this approach is that certain job ads may remain online very long. In some cases, firms may not even try to fill these job openings anymore. However, in the x28 data, the share of vacancies that remain online for a very long time is very low. For instance, the share of job postings older than one year that has not yet been deleted is only 0.2% of all job postings.

3.2 Algorithm to choose the sources included

Firms post vacancies either on their own firm webpage or on specialised third-party sites, so-called job portals. We include all postings which are posted directly on a company's webpage. Our index also includes vacancy postings from job portals. One issue with the vacancies over longer time horizons is that job portals may enter and exit the datasource, in which case the total number of job openings crawled by x28 may increase or decrease strongly from week to week. Moreover, the data from job portals can generate issues in real time. For instance, if a portal unanticipatedly implements changes in the frontend, the crawler may be temporarily unable to extract the portal's job postings. For the same reason, the stock of vacancies on certain job portals is quite unstable on a weekly basis in the historical data.

We develop the following algorithm to account for portal exit and entry and the fact that the number of job openings on certain portals is unstable in the historical data.

We qualify as a portal every website which hosts postings of more than 100 different companies. This threshold serves to distinguish companies that centrally post vacancies for different subsidiaries from actual job portals. Next, we define the daily vacancy stock on each portal website separately and all non-portal websites jointly.

[4] Another explanation is that certain vacancies of recruitment agencies do not represent "true" open positions but instead the their efforts to recruit workers for their databases ("phantom vacancies").

We then exclude portals if an extreme number of vacancies is created or deleted on a single day. The assumption is that in those cases, the fluctuations stem from technical issues rather than labour market factors. We apply a Hampel filter to detect such outliers in the daily series of deleted and created job openings. For every day and job portal, we compute the distribution of vacancies added and vacancies deleted over the last 365 days. We flag a time series as having an outlier if, on one or more days in a week, a portal has more newly created vacancies than the median plus 4 median absolute deviations of the rolling 365 days distribution of vacancies created on that portal. We apply the same criterion to the time series of deleted vacancies per day. If a portal is flagged, we exclude the growth rates from that portal for one month (5 weeks) from the computation of the indices.[5]

The following table shows that over the period from January 2018 to June 2022, we delete on average a bit more than 25% of all job postings in the raw data due to this procedure. 69% of the included vacancies stem from company domains.

Table 1: Summary statistics for vacancy statistics Jan 2018 - Jun 2022

Average number of job postings per week	104’871
Average number on company domains (included in index computation)	54’039
Average number on company domains and stable portals (included in index computation)	78’129

3.3 Index computation

We then compute the index based on the sample of vacancies defined in the previous subsections. More specifically, it is constructed as a weighted average of the growth rates of vacancies on the non-excluded portals and the sum of vacancies on company domains. The weight of every source in this aggregation is its share in the previous week’s stock of vacancies among the included sources.

We recompute the entire index every week with the newest version of the entire vacancy database. Each week, we thus incoporate all updates to the database that occurred in the meantime, including possible updates to job postings older than one week. Therefore, it is possible that there will be ex-post revisions of the index values older than one week. We save all vintages of the main series in our database. It will therefore be possible to reconstruct these revisions or to generate a real-time database.

Figure 3 provides a comparison of an index of all vacancies (excluding recruitment agencies) in the database (the raw data, red line) and our cleaned aggregate job index (blue). The figure reveals that our cleaned series is remarkably smoother than the raw series. Although we remove only a fourth of all vacancies on average (Table 1), the cleaning steps remove a few large week-to-week jumps in the raw data. These jumps are typically due to portal entry or exit. One also notes that the cleaned series has a seasonal pattern. For instance, there are always less online job postings in the first week of January.

3.4 Subindices

We compute the subindices based on the same set of job openings (job portals and firm websites) as the aggregate index. The computation is identical to the computation of the aggregate index but the sample is restricted to the pool of vacancies falling into the specific subcategory.

[5] The two parameters of the filter, 4 (the factor on the median absolute deviation in the filter) and 5 (the number of weeks a portal is excluded) were chosen in a data-driven way to best match the average growth rate of the vacant positions according to Besta in the 2018-2021 period.

3.4.1 By industry (NOGA/NACE)

To compute indices by industry, we assign each job posting to a NOGA/NACE 1-digit code (Letters) using a mapping between the company posting the vacancy and the industries provided by x28. If a company is active in more than one industry the job posting is included in all the subindices for industries in which the company is active.

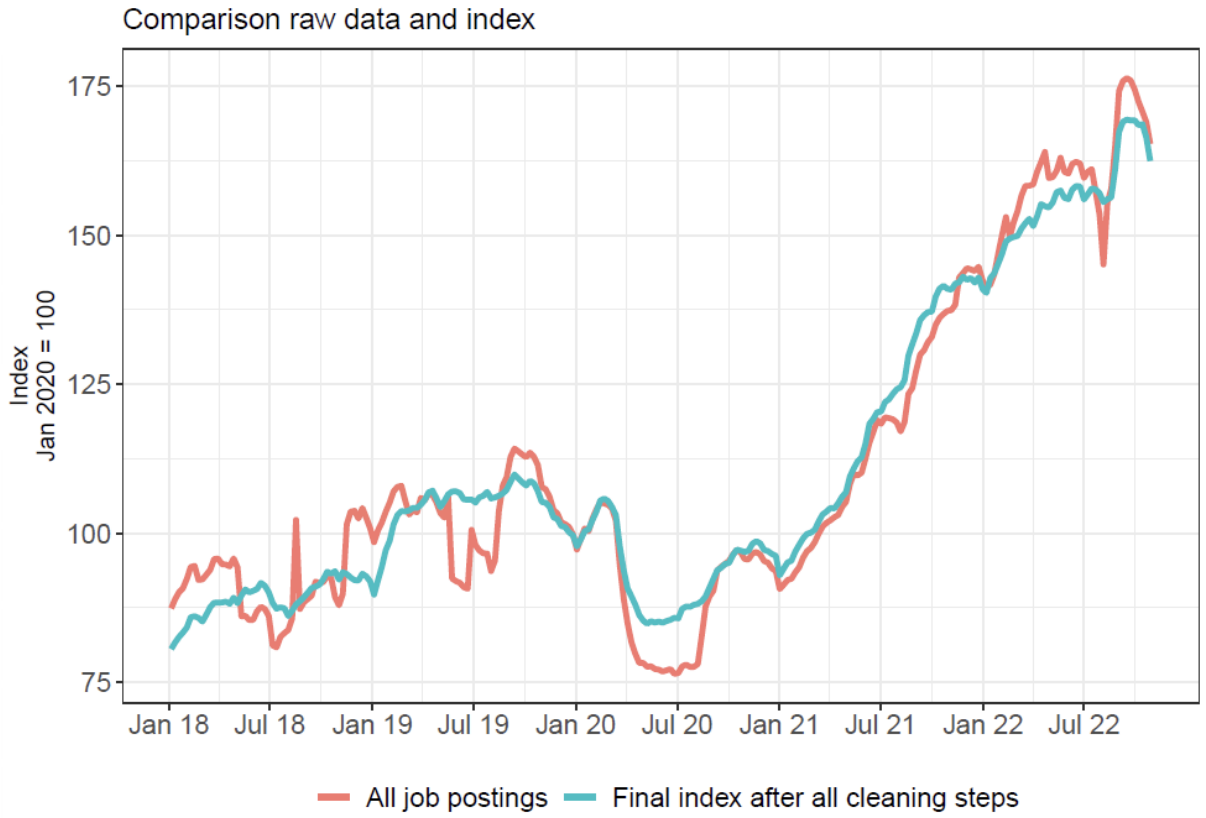
3.4.2 By broad occupation (ISCO)

To compute indices by broad occupation, we assign every job posting to a ISCO-08 one-digit-level occupation using a mapping between the jobtitle and the AVAM occupation provided by X28 and the mapping between the AVAM nomenclature and the ISCO provided by the State Secretariat for Economic Affairs. If a job title is assigned to more than one ISCO code, the posting is included in more than one subindices.

3.4.3 By canton

x28 assigns each job posting to a canton using text processing to extract the locality of the vacancy. Job postings can count towards the subindices of more than one canton if the wording in the job posting implies a bigger geographic area. For instance, if the locality is the "Basel area", we assign the job posting to Basel-Stadt and Basel-Landschaft. We use this information to compute canton-specific job indices.

Figure 3: Comparison of cleaned index with raw data



4. Index validation

In this section, we validate our new job indices by comparing the indices with official quarterly indices on the the number of vacant positions in Switzerland published by the SFSO. SFSO gathers information on the number of job vacancies during the Besta survey, a representative quarterly survey among 18,000 secondary and tertiary sector enterprises (65,000 businesses). The results can be broken down to by economic divisions and major regions (NUTS-II). In contrast to our index, the Besta is

available on a quarterly basis only and is published with a lag of several months. Moreover, it does not allow for a differentiation by single cantons or occupations. The data refers to the last day of each quarter.

Figure 4 shows how the number of job vacancies in the Swiss economy evolve according to the Besta data and according to our index. Panel A compares the evolution of the total index and Panel B compares the quarter-to-quarter changes of the index. The evolution of the two indices is very similar (correlation coefficient of 0.96). This is not self-evident given the different data sources.

The figure also illustrates the publication lead of the new weekly indices relative to the Besta data. At the time of writing this document (October 27, 2022), the most recent data point of the Besta refers to the second quarter, i.e. June 30, 2022. In contrast, the newest data point of our new job indices is from the third week in October 2022. The new job indices therefore reveal the evolution of job openings in July, August, September, and October 2022. The Swiss Job Tracker data suggests that the number of open vacancies continued to increase until September 2022 but may then have reached a turning point.

Figure 4: Evolution of open positions according to the index of the SFSO and according to the index from swissjobtracker.ch

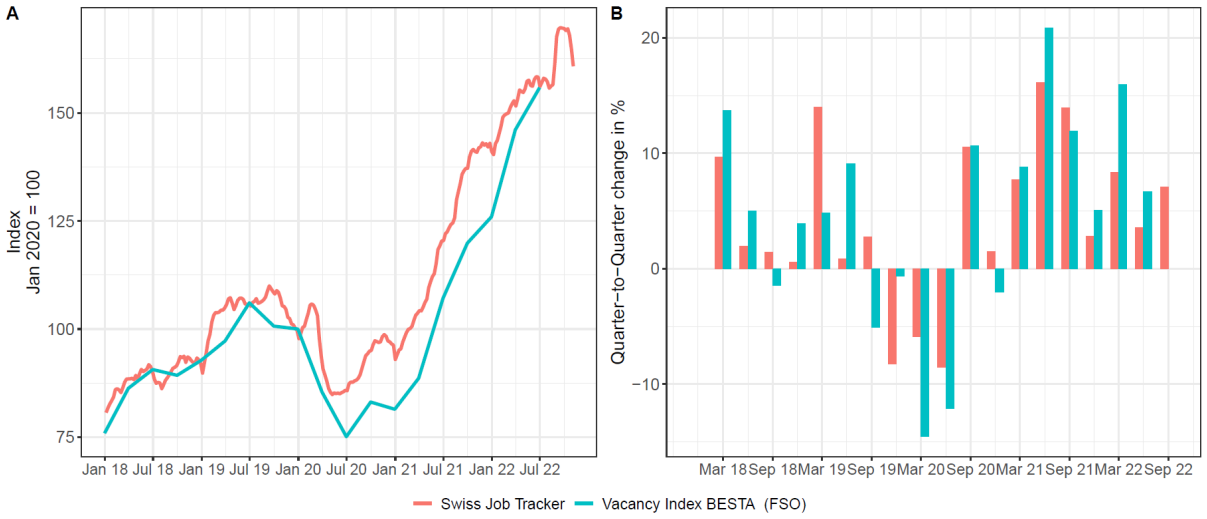


Figure 5 compares the evolution of job vacancies according to the Besta with the evolution of open positions according to the Swiss Job Tracker by industry. We show only those industries for which the level of aggregation is the same in the two data sources. For most industries, the two series are closely related. Note, that some differences were considerably smaller if we would change the base year (e.g. Information und Kommunikation).

Figure 5: Evolution of open positions according to the index of the SFSO and according to the index from swissjobtracker.ch by industry