INSTITUTE FOR EMPLOYMENT RESEARCH
The Research Institute of the Federal Employment Agency

IAB-DISCUSSION PAPER

Articles on labour market issues

1|2021 Matching for three: big data evidence on search activity of workers, firms, and employment service

Tobias Hartl, Christian Hutter, Enzo Weber

ISSN 2195-2663

Matching for three: big data evidence on search activity of workers, firms, and employment service

Tobias Hartl (IAB),
Christian Hutter (IAB),
Enzo Weber (IAB, University of Regensburg, Germany)

Mit der Reihe „IAB-Discussion Paper“ will das Forschungsinstitut der Bundesagentur für Arbeit den Dialog mit der externen Wissenschaft intensivieren. Durch die rasche Verbreitung von Forschungsergebnissen über das Internet soll noch vor Drucklegung Kritik angeregt und Qualität gesichert werden.

The “IAB-Discussion Paper” is published by the research institute of the German Federal Employment Agency in order to intensify the dialogue with the scientific community. The prompt publication of the latest research results via the internet intends to stimulate criticism and to ensure research quality at an early stage before printing.

Contents

List of Figures

List of Tables

Abstract

We generate measures for search intensity of employers and job seekers and – as a novel feature – for placement intensity of employment agencies. For this purpose, we tap big data on online activity from the job exchange of the German Federal Employment Agency and its internal placement-software. We use these data to estimate an enhanced matching function where the efficiency parameter varies with the search and placement intensities. The results show that the intensity measures significantly contribute to the variation in job findings.

Zusammenfassung

Wir generieren Maße für die Suchintensität von Arbeitgebern und Arbeitssuchenden und - zum ersten Mal - für die Vermittlungsintensität von Arbeitsagenturen. Zu diesem Zweck greifen wir auf Big Data zu Online-Aktivitäten aus der Online Jobbörse der Bundesagentur für Arbeit und ihrer internen Vermittlungssoftware zurück. Wir verwenden diese Daten, um eine erweiterte Matchingfunktion zu schätzen, bei der der Effizienzparameter mit den Such- und Vermittlungsintensitäten variiert. Die Ergebnisse zeigen, dass alle drei Intensitätsmaße erheblich zur Erklärung der Job-findings-Variation beitragen.

JEL

C55, C78, J64

Keywords

big data, labour market search intensity, matching function, online activity

Acknowledgements

We thank Ortwin Herbst, Maren Müller, Michael Stops, Clemens Usbeck, and participants of the annual meeting of the FEA’s statistics department, the workshop “Genesis of information from online job advertisements”, the IWH/IAB workshop, and the virtual COVID-19 conference at Aarhus University for valuable support and helpful suggestions.

1 Introduction

The key task of a labour market is bringing together job seekers and open positions. This matching naturally depends on the number of unemployed and the number of vacancies, as formalized in matching theory (e.g. Pissarides, 2000). The matching function represents a key ingredient of many macroeconomic models. In standard specifications, however, the sole instruments available to vary hirings are for firms to change the number of vacancies and for workers to enter or leave unemployment. This neglects a substantial variation of hirings, reflecting the need for modeling time-varying matching efficiency. Understanding the sources of this variation in matching efficiency is important since it is a crucial driver of fluctuations of employment and unemployment. Beyond the stock variables, mismatch (Sahin et al. (2014), Hutter/Weber (2017)), information and institutions, and the *behaviour* of the relevant agents play a role. Whether people get into work, for example, depends crucially on how intensively unemployed look for jobs and how much effort employers make when trying to fill an open position. This behaviour can be described as search intensity (e.g. Pissarides (2000)).

Beyond job seekers and employers on the two market sides, a further agent is present in the labour market: the employment service. The activities of the intermediary between the two sides can have an impact on whether unemployed and vacancies come together. As a counterpart to search intensity, we introduce the notion of placement activity of the labour market intermediary.

Conventional measures of search intensity are often based on survey data, e.g. from time-use surveys (Mukoyama/Patterson/Sahin (2014), Gomme/Lkhagvasuren (2015), Krueger/Mueller (2011)). With the digitalisation of labour markets, online data such as the presence of individuals in online job search (Kuhn/Skuterud (2004), Faberman/Kudlyak (2019)), job seekers' search behaviour at search engines (Baker/Fradkin (2017)), or applications to job postings got into the focus.

We contribute to the literature by measuring search intensity using a source of big data that directly captures online activity: We evaluate how often the job exchange website of the German Federal Employment Agency (FEA) and its placement platform have been accessed for search or placement activities. Beyond search intensity of job seekers and firms, our approach enables – for the first time - measuring placement activity of employment agencies.

We enhance a standard matching function by search and placement behaviour. This can provide more flexibility in explaining the dynamics of empirical data (compare Gomme/Lkhagvasuren (2015)) and might also cure a missing-variables problem. We find that both search intensities and placement activity significantly explain job findings. Together, they capture a considerable part of hirings variation ignored by the standard matching function.

The paper is structured as follows: The next section presents the data. Section 3 discusses the role of search intensity and placement activity for the matching precess. In Section 4, we estimate an enhanced matching function and show the results. The final section concludes.

2 Data

Germany is a typical example of labour markets with a strongly institutionalised public employment agency. Workers who know that they will be unemployed are obliged to report the job-seeking status to the FEA immediately in order to be granted the full amount of unemployment benefits later. Hence, the FEA is the central intermediary for the unemployed. It runs a job exchange website[1] where job seekers (JS) can apply for open positions or offer their workforce, and firms (F) can find workers or place job offers. Once the job exchange is accessed, server log files are stored in anonymous form. These server log files provide valuable information, e.g. about which part of the job exchange website the user has visited. Thus, they allow distinguishing whether the job seekers' or employers' area of the job exchange was accessed and hence measuring the respective search intensities (I^{JS}, I^{F}). For instance, if the log file implies that the visitor wanted to look over her job openings or to find suitable job candidates, the exchange website was most likely accessed by an employer. On the other hand, if, for instance, the visitor searched for suitable job openings, it can be assumed that a job seeker accessed the website.

Throughout the paper, we use "activated visits", i.e. only online activities where a visitor was active on the website beyond merely opening it are counted. Since activated visits involve more than one page view, it can be assumed that the visitor is interested in the content and took a closer look at it. Thus, activated visits represent the qualified traffic on the online job exchange platform. Furthermore, this helps exclude unwanted online traffic, e.g. by bots, from the data.

With the FEA's placement software *VerBIS*, employment agents (EA) screen the labour supply and demand sides to identify potential positions for job seekers or suggest candidates for an open position. Since administrative tasks are also carried out in VerBIS, the information in the log files serve to identify genuine placement activities (e.g. generating a placement proposal) to measure the placement intensity I^{EA}. To our knowledge, the placement activities of employment services have not yet been investigated.

[1] See https://jobboerse.arbeitsagentur.de.

Monthly hirings are measured as accumulated flow from unemployment into employment between two counting days of the FEA's statistics.[2] To ensure consistency, the same counting days apply for the online activities. In order to capture search and placement *intensities* instead of mere accumulated *activities*, we divide the accumulated activities by either the stock of unemployed (in case of I^{JS}), the stock of vacancies (in case of I^F), or by the sum of unemployed and vacancies (in case of I^{EA}), at the end of the previous counting period, respectively.[3] Also the data on the number of unemployed and vacancies stem from the FEA.

All variables are calendar-adjusted, i.e. divided by the number of working days between two counting days, and seasonally adjusted. Occasionally, there are missing data due to changes in the platforms. In the estimations of Section 4, they are neutralised by impulse dummies, while shift dummies capture potential level shifts after the breaks.

Figure 1 depicts the monthly intensities that show relevant variation during the sample. In the COVID-19 crisis, they experience a dramatic drop, reflecting the firms' reluctance to hire and difficulties for employment agents to pursue their placement tasks under corona conditions. Job seekers' search intensity decreased, too, due to a strongly increasing denominator (unemployment). Also the hirings plummeted.

From a data quality perspective, our intensity measures have several advantages. They are particularly well suited for disentangling the search activities of job seekers, firms and placement agents. They are based on big data directly capturing online activity. Thus, they can build on large samples and do not have to rely on survey data or on counting actual applications. Finally, they can be accessed without any publication lag.

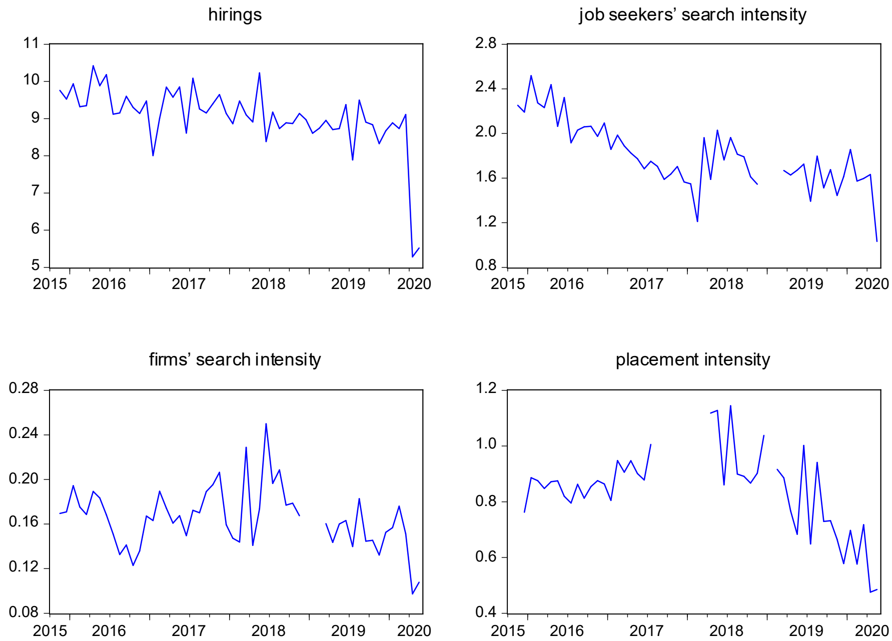
3 Enhancing the matching function

Search and matching theory states that vacancies (V) and unemployed (U) form matches (H for hirings) through a Cobb-Douglas production function. After log-linearisation, the matching function reads

$$ln(H_t) = \mu + \alpha ln(V_{t-1}) + (1-\alpha) ln(U_{t-1}), \tag{3.1}$$

[2] A counting day typically is around the middle of a month.
[3] See below for robustness checks on the timing of the denominators.

Figure 1: Hirings and search and placement intensities



Notes: hirings: outflow from unemployment into employment (in thousand). Source: FEA statistics. Intensities: activated visits by job seekers / employers / employment agency per working day normalised by unemployed / vacancies / unemployed+vacancies. Structural breaks after periods of missing data are eliminated by level shift dummies in ARMA models. Seasonally adjusted data.
Source: Netmind, own calculations.

where μ denotes the efficiency parameter and α and $(1-\alpha)$ are the elasticities of new matches with respect to vacancies and unemployed, respectively, under the assumption of constant returns to scale. While the efficiency parameter is constant in the standard version, time variation in matching efficiency can be substantial (e.g. Klinger/Weber (2016), Sedlacek (2014)).

This is why we explicitly allow μ to have (beyond a time-invariant part μ^*) a time-varying part that depends on the search intensity of job seekers and firms (I_t^{JS}, I_t^F), the placement intensity of the employment agencies (I_t^{EA}), and a normally-distributed error term ϵ_t.

$$\mu_t = \mu^* + \beta_1 ln(I_t^{JS}) + \beta_2 ln(I_t^F) + \beta_3 ln(I_t^{EA}) + \epsilon_t \tag{3.2}$$

In addition to mitigating the missing-variables-problem, this can provide more flexibility in explaining the dynamics of empirical data (compare Gomme/Lkhagvasuren (2015)). The new specification extends Hornstein/Kudlyak (2017) and Davis/Faberman/Haltiwanger (2013), who analyse explanatory power of job seekers' search intensity and recruiting intensity, respectively.

Using search activity from the current month adheres to the concept of the matching function that unemployed and vacancies present at the beginning of the counting period (i.e. $t-1$) form matches during the counting period. Nonetheless, it is conceivable that also earlier search activities may lead to matches in month t. Therefore, also lagged search activities were investigated as robustness check below.

4 Results

Inserting (3.2) into (3.1) yields

$$\begin{aligned} ln(H_t) &= \mu^* + \alpha ln(V_{t-1}) + (1-\alpha) ln(U_{t-1}) \\ &\quad + \beta_1 ln(I_t^{JS}) + \beta_2 ln(I_t^F) + \beta_3 ln(I_t^{EA}) + \epsilon_t. \end{aligned} \tag{4.1}$$

Since all variables appear in logs, β_1 to β_3 are the elasticities of hirings with respect to the three search intensities. We measure these intensities by using the online data described in Section 2 as proxies. Both the standard and the enhanced model are estimated via nonlinear least squares and with heteroscedasticity- and autocorrelation-robust standard errors.

Table 1 shows the results. All three intensities contribute positively to the hirings variable. The elasticity of hirings with respect to the job seekers' search intensity is highest with about 0.19, while the other two elasticities amount to approximately 0.12. The effects are statistically significant with p-values between 0.011 and 0.086.

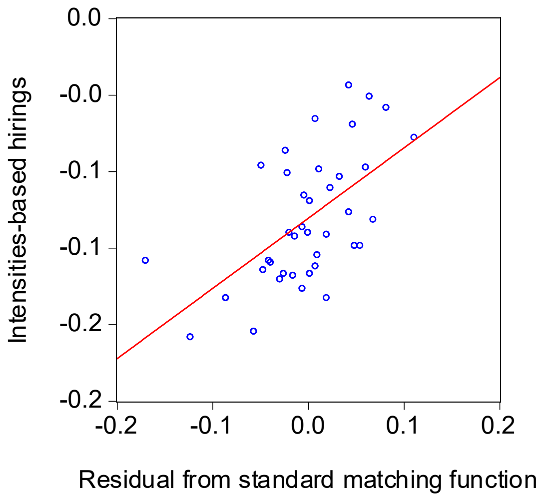
Table 1: Estimation results of standard and enhanced matching function

Parameter	coef.	std. dev.	p-value	coef.	std. dev.	p-value
	Standard matching function			Enhanced matching function		
μ, μ^*	-5.2593	0.0764	0.0000	-5.0768	0.1004	0.0000
α	0.2705	0.0574	0.0000	0.3329	0.0942	0.0012
β_1				0.1937	0.1093	0.0857
β_2				0.1216	0.0570	0.0405
β_3				0.1173	0.0438	0.0114
R^2		0.3737			0.5993	

Notes: Estimation of Equation (4.1). Sample: 2015:12 to 2020:3.
Source: Own calculations.

Beyond significance, the effects are also economically relevant: From the standard matching function, we can infer a "Solow residual", which then can be compared to the part of the fitted value of $ln(H_t)$ driven by the three intensities in the enhanced matching function (compare also Davis/Faberman/Haltiwanger (2013)). Figure 2 reveals a substantial correlation between the two measures (r=0.59). This emphasizes that the intensities are able to explain an important part of the otherwise neglected variation of matching efficiency and hence hirings.

Figure 2: Part of hirings explained by search intensities and residuals from standard matching function



Notes: The graph shows the part of $ln(H_t)$ that is explained by the three intensities in the enhanced matching function and the residual backed out from a standard matching function.
Source: Own calculations.

This finding is supported by the fact that the explained part of the variation in $ln(H_t)$ is only 37 percent in the standard matching function, whereas it increases by 22.6 percentage points (or 60%) to just under 60 percent in the enhanced version. Including the intensities increases the point estimate for α by 23 percent from 0.27 to 0.33, which points to a moderately-sized under-estimation of the matching parameter in case of the standard matching function.

Until now we excluded the COVID-19 crisis months 2020:4 and 2020:5 due to their extreme values. If we include them into the sample, the estimated elasticities become even more significant. The drop in vacancies by 10 percent can by far not explain the collapse of hires by 42 percent, but the drop in the search intensities captures the collapse quite well. Logically, the explanatory power of the enhanced compared to the standard matching function increases even more. Therefore, the estimation results presented above can be considered conservative.

Some endogeneity of the search intensity measures could result from the fact that search activity of unemployed and vacancies that already formed a match will disappear during the month. Ceteris paribus that would lead to less intensive search in months with many matches. In this sense, our positive elasticities of matches with regard to search intensity would be conservative, i.e. represent lower bounds. Still, the search effort that disappears during the month would be replaced by search effort of new unemployed and vacancies. In order to take this into account in a robustness check,[4] we relate search efforts to the mean of the stocks from the beginning and the end of the counting period, i.e. we define I_t^{JS} as accumulated activated visits divided by the average of U_{t-1} and U_t, and analogously for I_t^F and I_t^{EA}. In this case, the estimated elasticities do not change substantially, though. They slightly increase to 0.1954 (β_1), 0.1235 (β_2), and 0.1182 (β_3).

Beyond the number of unemployed and vacancies, and also beyond search and placement intensities, also compositional effects could play a role in the matching function (see e.g. Barnichon/Figura (2015) and Ravenna/Walsh (2012)). Therefore, as a robustness check, we allowed for a more general setting including control variables. A typical set of variables capturing relevant characteristics of job seekers is given by the shares of long term (>1 year unemployment duration), older (>55 years of age), younger (<25 years of age), female and foreign unemployed among total unemployment. The unemployment shares are taken from the FEA's statistics. Since aggregate unemployment itself enters the matching function with the first lag, we use the same lag for the control variables. Especially the share of long-term unemployed has a clear negative effect on matches with an elasticity of -0.6838 (compare also Barnichon/Figura (2015) who find a substantial role of the rate of long-term unemployed for matching efficiency). However, this does not impair the role of the search intensities: They amount to 0.2540 (β_1), 0.0987 (β_2), and 0.1412 (β_3).

Regarding the dynamic properties of the enhanced matching function, also earlier search activities could lead to matches in month t. Therefore, in a further robustness check, we include lagged search and placement intensities (i.e. I_{t-1}^{JS}, I_{t-1}^{F}, and I_{t-1}^{EA}) to Equation (4.1). However, the lagged terms turn out to be insignificant, and the estimated elasticities of the competing contemporaneous search variables do not change substantially.

5 Conclusion

This article introduces innovative big data allowing the instantaneous measurement of search and – for the first time – placement intensity in the labour market in form of online activity. We use these data to estimate an enhanced matching function where the efficiency parameter varies with the job seekers' and firms' search intensities and the placement intensity of the employment agencies. The results show that all intensity measures significantly explain the variation in job findings.

[4] Tables of the robustness checks are not included due to limited space. They are available upon request.

A key insight for macro-labour models is that the standard matching function neglects substantial variation of hirings and hence employment that can be explained by search intensity. The underlying study shows how the matching function can be enhanced by suitable measures and contributes pioneering work especially with respect to placement intensity.

In the COVID-19 crisis so far, the novel data reveal that all three search intensities dropped substantially. Our results demonstrate that this will have adverse impacts beyond the decline in vacancies. The most critical labor market effects of the crisis may arise not via the separation but via the hiring margin (Merkl/Weber (2020)).

References

Baker, Scott; Fradkin, Andrey (2017): The Impact of Unemployment Insurance on Job Search: Evidence from Google Search Data. In: The Review of Economics and Statistics, Vol. 99, No. 5, p. 756–768.

Barnichon, Regis; Figura, Andrew (2015): Labor Market Heterogeneity and the Aggregate Matching Function. In: American Economic Journal: Macroeconomics, Vol. 7, p. 222–249, URL https://ideas.repec.org/a/aea/aejmac/v7y2015i4p222-49.html.

Davis, Steven J.; Faberman, R. Jason; Haltiwanger, John C. (2013): The Establishment-Level Behavior of Vacancies and Hiring . In: The Quarterly Journal of Economics, Vol. 128, No. 2, p. 581–622, URL https://doi.org/10.1093/qje/qjt002.

Faberman, R. Jason; Kudlyak, Marianna (2019): The Intensity of Job Search and Search Duration. In: American Economic Journal: Macroeconomics, Vol. 11, p. 327–357, URL https://ideas.repec.org/a/aea/aejmac/v11y2019i3p327-57.html.

Gomme, Paul; Lkhagvasuren, Damba (2015): Worker search effort as an amplification mechanism. In: Journal of Monetary Economics, Vol. 75, p. 106–122, URL https://ideas.repec.org/a/eee/moneco/v75y2015icp106-122.html.

Hornstein, Andreas; Kudlyak, Marianna (2017): Estimating Matching Efficiency with Variable Search Effort. 2017 meeting papers, Society for Economic Dynamics, URL https://ideas.repec.org/p/red/sed017/881.html.

Hutter, Christian; Weber, Enzo (2017): Mismatch and the Forecasting Performance of Matching Functions. In: Oxford Bulletin of Economics and Statistics, Vol. 79, No. 1, p. 101–123.

Klinger, Sabine; Weber, Enzo (2016): Decomposing Beveridge Curve Dynamics by Correlated Unobserved Components. In: Oxford Bulletin of Economics and Statistics, Vol. 78, p. 877–894.

Krueger, Alan B.; Mueller, Andreas (2011): Job Search, Emotional Well-Being and Job Finding in a Period of Mass Unemployment: Evidence from High-Frequency Longitudinal Data. In: Brookings Papers on Economic Activity, Vol. 42, p. 1–81, URL https://ideas.repec.org/a/bin/bpeajo/v42y2011i2011-01p1-81.html.

Kuhn, Peter; Skuterud, Mikal (2004): Internet Job Search and Unemployment Durations. In: American Economic Review, Vol. 94, No. 1, p. 218–232, URL https://www.aeaweb.org/articles?id=10.1257/000282804322970779.

Merkl, Christian; Weber, Enzo (2020): Rescuing the labour market in times of COVID-19: Don't forget new hires. In: VoxEU 07 April.

Mukoyama, Toshihiko; Patterson, Christina; Sahin, Aysegul (2014): Job search behavior over the business cycle. Staff Reports 689, Federal Reserve Bank of New York, URL https://ideas.repec.org/p/fip/fednsr/689.html.

Pissarides, Christopher (2000): Equilibrium Unemployment Theory, 2nd Edition, Vol. 1. The MIT Press, 1 ed., URL https://EconPapers.repec.org/RePEc:mtp:titles:0262161877.

Ravenna, Federico; Walsh, CARL E. (2012): Screening and Labor Market Flows in a Model with Heterogeneous Workers. In: Journal of Money, Credit and Banking, Vol. 44, p. 31–71, URL http://www.jstor.org/stable/23321956.

Sahin, AysegÃ¼l; Song, Joseph; Topa, Giorgio; Violante, Giovanni L. (2014): Mismatch Unemployment. In: American Economic Review, Vol. 104, No. 11, p. 3529–3564.

Sedlacek, Petr (2014): Match efficiency and firms' hiring standards. In: Journal of Monetary Economics, Vol. 62, p. 123–133.

Imprint

IAB-Discussion Paper 1|2021

Publication Date
14. Januar 2021

Publisher
Institute for Employment Research
of the Federal Employment Agency
Regensburger Straße 104
90478 Nürnberg
Germany

Download
http://doku.iab.de/discussionpapers/2021/dp0121.pdf

All publications in the series "IAB-Discusssion Paper" can be downloaded from
https://www.iab.de/en/publikationen/discussionpaper.aspx

Website
www.iab.de/en

Corresponding author
Christian Hutter
E-Mail christian.hutter@iab.de