

Lochner, Benjamin

Article

IABSE-ADIAB: IAB job vacancy survey data linked to administrative data

Journal of Economics and Statistics

Provided in Cooperation with:

De Gruyter Brill

Suggested Citation: Lochner, Benjamin (2024) : IABSE-ADIAB: IAB job vacancy survey data linked to administrative data, Journal of Economics and Statistics, ISSN 2366-049X, De Gruyter Oldenbourg, Berlin, Vol. 244, Iss. 1/2, pp. 159-165,
<https://doi.org/10.1515/jbnst-2023-0004>

This Version is available at:

<https://hdl.handle.net/10419/333279>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Data Observer

Benjamin Lochner*

IABSE-ADIAB – IAB Job Vacancy Survey Data Linked to Administrative Data

<https://doi.org/10.1515/jbnst-2023-0004>

Received January 13, 2023; accepted January 16, 2023

Abstract: The IABSE-ADIAB dataset provides the linkage between the most recent case of a hire from the IAB Job Vacancy Survey and the administrative data of the hired person. The data comprise the entire working history of persons who were identified as IAB Job Vacancy Survey hires from 2010 to 2020. The data are processed in a similar manner to the Sample of Integrated Employment Biographies of the IAB and can be linked to the survey data using the provided unique key variable.

Keywords: IAB job vacancy survey, German administrative microdata, labour market data, data linkage

JEL-Classification: J23, J63, E24

1 Introduction

In many search and matching models, employers post vacancies to attract job seekers. However, details of the search and matching process are often hidden behind the concept of an aggregated matching function. To test the model predictions, micro-data that helps open up the black box of the matching function is crucial.

The IAB Vacancy Survey-ADIAB (IABSE-ADIAB) is a newly processed dataset that builds the linkage between the most recent case of a hire of the IAB Vacancy Survey (IAB-JVS) and the administrative (process-generated) data of the hired persons. The former contains details on many aspects on the hiring practices of German establishments. The latter contains the entire employment history of the hired persons. The data are processed in a similar manner to the IAB's Sample of Integrated Employment Biographies (SIAB, see Frodermann et al. 2021) and can be linked to the survey data from 2010 to 2020. The linkage is based on the matching approach

***Corresponding author: Benjamin Lochner**, Institut für Arbeitsmarkt-und Berufsforschung und Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberg, Germany,
E-mail: benjamin.lochner@iab.de

proposed by Lochner (2019). Overall, this technique finds up to 72 percent of all reported hires.¹ The combination of data from the firm-side and worker-side is unique and provides an ideal laboratory for answering many research questions that relate to the search and matching process. To the best of my knowledge, there is no other dataset with these features worldwide.

This article describes how this new dataset was generated, how it is structured, and for which research topics the data might be relevant. It is organized as follows. Section 2 explains the data sources of the IABSE-ADIAB. Section 3 explains details with respect to linkage and the identification of IAB-JVS hires. Section 4 explains the data structure and how to access them. Section 5 summarizes and gives an overview of potential research questions.

2 Data Sources

The IABSE-ADIAB collects all employment spells from the Integrated Employment Biographies (IEB) for all identified hires reported in the IAB-JVS.

2.1 Administrative Data of the Integrated Employment Biographies

The individual data originate from the IEB which are processed by the IAB. The data have been prepared in the same manner as the SIAB. They unite data from different data sources, each of which may contain information from different administrative procedures. The source of the individual employment spells is the Employee History (BeH). This data source is the integrated notification procedure for health, pension and unemployment insurance. The BeH covers all employees who are subject to social security at least one day in a given year. Since 1999, the data also cover employees in marginal part-time employment and unpaid family workers. These spells are supplemented by establishment information from the Establishment History Panel (“Betriebs-Historik-Panel”, see Ganzer et al. (2022), which aggregates establishment-level variables (with June 30 each year as record day). Another source is the Benefit Recipient History (“Leistungsempfänger-Historik”), which contains episodes during which individuals receive any form of earnings replacement benefits from the Federal Employment Agency (such as unemployment benefits, unemployment assistance, maintenance allowance, and other contributions paid by the

¹ This number relates to reported hires with no missing in the key variables. Between 2010 and 2020, the matching algorithm identifies 57,582 hires.

Federal Employment Agency). Episodes of unemployment benefit II receipt are included through the Unemployment Benefit II Recipient History (“Leistungshistorik Grundsicherung”). These spells are available from 2005 onward. The data is additionally supplemented by the data about jobseekers and information on active labour market policy participation from the Participants-in-Measures History Files (“Maßnahmeteilnahmehistoriken”).

2.2 IAB Vacancy Survey

The IAB-JVS comprises data on search and recruitment processes at German establishments. It is the official data provider for Eurostat’s vacancy statistics. Among other things, it reports the number of vacancies on the German labor market, including those vacancies that are not reported to the Federal Employment Agency. The main questionnaire is conducted in every fourth quarter of a given year. It collects information on the number and structure of vacancies, future labor demand, the current economic situation, and the expected economic development of participating establishments. The survey additionally collects information on the most recent hire of an establishment. In this part of the survey, establishments report details about their last filled vacant position the filled during the last 12 months. They are asked about certain job characteristics, such as the exact job requirements, the hiring channel, the search duration, and the exact hiring date.

A detailed description of the survey data, their structure and information on the sample and the survey in general can be found in the data manual of the IAB JVS (Bossler et al. 2021), which also lists generated variables and details on weighting. More information can be found on the official survey website <https://www.iab.de/en/befragungen/stellenangebot.aspx>.

3 Data Linkage

To identify the reported most recent hires from the IAB-JVS in the IEB, a deterministic matching algorithm is executed (see Lochner 2019 for methodological details). The algorithm uses overlapping information that is both present in the IAB-JVS and the administrative data. Once the hiring spell is identified, the entire working history of the hired person is drawn from the IEB.

Specifically, the method uses

- (i) the system-free, internal IAB establishment identifier,
- (ii) the hiring dates,
- (iii) the gender of hired person,

- (iv) the age of hired person, and
- (v) the occupational code.

Additional plausibility checks with respect to contract hours and the provided wage information are performed.

The algorithm has six main steps and up to three sub-steps. Each step differs in the acceptance tolerance with respect to deviations in the overlapping information in the two data sets. Each step tries to find combinations in the two data sets that match according to certain step-specific definitions. A step ends whenever there are no more valid and unique matches to find. A match is valid and unique if and only if, in a given time period and in a given establishment, the step-specific conditions are met and there is no other hire in the administrative data that has the same attributes.

The following steps are considered:

Step 1: Within a given establishment, all four main variables (hiring date, gender, age and occupation code) must be equal without any deviations between the two data sets.

Step 2: This step allows for deviations in the hiring date of up to 31 days between the two data sets. All other attributes must match exactly.

Step 3: This step allows for constellations where, either in the administrative data or in the survey data, the occupation code is missing. All other attributes must match exactly.

Step 4: This step allows for absolute deviations in age up to 2 years. All other attributes must match exactly.

Step 5: This step combines Step 3 and 4 as it allows deviations in the age of up to two years as well as missing occupational information. All other attributes must match exactly.

Step 6: This step allows for implausible occupation code combinations as well as deviations in the hiring date up to another 31 days. All other attributes must match exactly.

4 Data Structure and Access

4.1 Data Structure

Table 1 gives an overview of the number of unique hires that were identified with each step. In total, the procedure can identify 57,582 of the most recent IAB-JVS hires in the admin data, originating from 57,384 persons. Their complete working history amounts to 2,069,117 IEB-spells (4,476,181 non-overlapping observations after episode splitting).

Table 1: Results of the matching procedure by algorithm step. Source: IABSE-ADIAB.

Step	N	Percentage	Cummulative
1	14,285	24.81	24.81
2	2,876	4.99	29.80
3	19,715	34.24	64.04
4	6,045	10.50	74.54
5	10,672	18.53	93.07
6	3,989	6.93	100.00
Total	57,582	100.00	

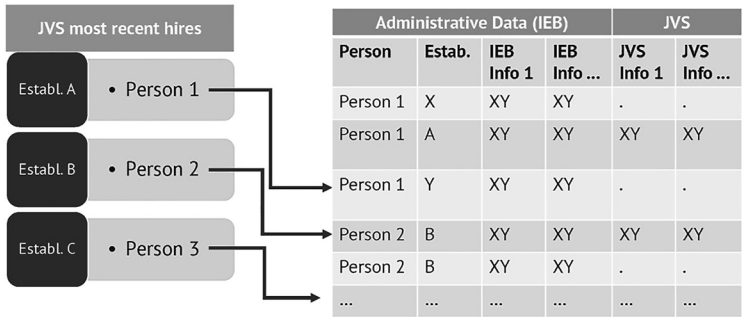


Figure 1: Illustration of data linkage. Source: Own representation.

The IAB-JVS comes in separate yearly data files. The relevant content belongs to the main survey questionnaire from the fourth quarter. The administrative data in the IABSE-ADIAB is structured identically to the SIAB. Both datasets can be uniquely connected by the identifier, which is present in both datasets.²

Figure 1 illustrates the mapping from the IAB-JVS to the IEB. The spell that displays the specific IAB-JVS hire is tracked in the IEB. Information from the IAB-JVS can be linked to these spells. Additionally, the IABSE-ADIAB shows the employment history of the hired person, that is, potential previous and future employment spells.

² The JVS data are in wide format, which means that variable names are time-specific. The administrative data are in long format. To connect both datasets, the user needs to first transfer the JVS data to long format. Example code can be provided by the Research Data Centre (FDZ).

4.2 Data Access

The IABSE-ADIAB data are weakly anonymised and can be accessed during guest visits to one of the locations of the Research Data Centre of the Federal Employment Agency at the Institute for Employment Research (IAB-RDC). Subsequently, remote data processing is possible.

Initially, an application for data use must be submitted to the IAB-RDC. After approval, a contract needs to be signed between the institution of the researcher and the IAB. Details on how to apply for the data set and the data processing options can be found on the website of the IAB-RDC (<https://fdz.iab.de/en/data-access/>). Access to the data can only be granted once the applicant has signed contracts with both research data centers. Testdata and further information is available on the website of the IAB-JVS (<https://fdz.iab.de/en/our-data-products/establishment-data/iab-job-vacancy-survey/>).

5 Summary and Outlook

The IABSE-ADIAB is a dataset that links survey data on hiring processes to high-quality administrative employment records on the person level. It builds on a deterministic linking algorithm that uses overlapping information that is present both in the IAB-JVS and the administrative data. The provided data contain all employment spells of identified IAB-JVS hires and can be supplemented by IAB-JVS data.

So far, the data have been used in the following articles: Gürtzgen et al. (2021) evaluate the effects of the high-speed internet expansion on the match quality of new hires. Carrillo-Tudela et al. (2022) investigate the extent to which search channels such as job postings, networks and the public employment agency have differential effects on labor market outcomes such as wages and employment. Lochner and Merkl (2022) explore the role of gender-specific application behavior as well the selection behavior of firms for the gender earnings gap in Germany. All of these articles have in common that they exploit data from the firm-side while controlling for important aspects from the worker-side, or vice versa.

As these examples show, the IABSE-ADIAB offers an ideal laboratory to assess the details of the search and matching process, as it allows observing important dimensions such as the characteristics of the hiring firm, the hired worker and the recruitment process itself. Furthermore, the data are well suited to study all issues related to labor demand, recruitment processes and their implications for matching outcomes in general.

References

- Bossler, M., Gartner, H., Kubis, A., and Küfner, B. (2021). *IAB- Stellenerhebung: Betriebsbefragung zu Stellenangebot und Besetzungsprozessen, Wellen 2000 bis 2018 mit Folgequartalen von 2006 bis 2019*. FDZ-Datenreport 09/2021 (de). Nürnberg.
- Carrillo-Tudela, C., Kaas, L., and Lochner, B. (2022). Matching through search channels. Unpublished manuscript.
- Frodermann, C., Schmucker, A., Seth, S., and Berge, P.V. (2021). *Stichprobe der integrierten arbeitsmarktbiografien (SIAB) 1975-2019*. FDZ-Datenreport 01/2021 (de). Nürnberg.
- Ganzer, A., Schmucker, A., Stegmaier, J., and Stüber, H. (2022). *Betriebs-historik-panel 1975-2020*. FDZ-Datenreport, 03/2022 (de). Nürnberg.
- Gürtzgen, N., Lochner, B., Pohlman, L., and van den Berg, G.J. (2021). Does online search improve the match quality of new hires? Lab. Econ. 70, <https://doi.org/10.1016/j.labeco.2021.101981>.
- Lochner, B. (2019). *A simple algorithm to link last hires from the job vacancy survey to administrative records*. FDZ-Methodenreport 06/2019 (en). Nürnberg.
- Lochner, B. and Merkl, C. (2022). *Gender-specific application behavior, matching, and the residual gender earnings gap*. IAB Discussion Paper 22/2022.