

Plümpe, Verena; Stegmaier, Jens

Article

Micro data on robots from the IAB Establishment Panel

Journal of Economics and Statistics

Provided in Cooperation with:

De Gruyter Brill

Suggested Citation: Plümpe, Verena; Stegmaier, Jens (2023) : Micro data on robots from the IAB Establishment Panel, Journal of Economics and Statistics, ISSN 2366-049X, De Gruyter Oldenbourg, Berlin, Vol. 243, Iss. 3/4, pp. 397-413,
<https://doi.org/10.1515/jbnst-2022-0045>

This Version is available at:

<https://hdl.handle.net/10419/333253>

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

Verena Plümpe and Jens Stegmaier*

Micro Data on Robots from the IAB Establishment Panel

<https://doi.org/10.1515/jbnst-2022-0045>

Received July 15, 2022; accepted July 17, 2022

Abstract: Micro-data on robots have been very sparse in Germany so far. Consequently, a dedicated section has been introduced in the IAB Establishment Panel 2019 that includes questions on the number and type of robots used. This article describes the background and development of the survey questions, provides information on the quality of the data, possible checks and steps of data preparation. The resulting data is aggregated on industry level and compared with the frequently used robot data by the International Federation of Robotics (IFR) which contains robot supplier information on aggregate robot stocks and deliveries.

Keywords: robots, plant-level, micro-data, data quality

JEL Classification: O14, O33, C81

1 Introduction

It is a widely discussed topic in politics and academia how increasing automation, besides machine learning and artificial intelligence, will impact the future of work. Especially the rise of industrial robots and subsequent effects on labor market outcomes and productivity have caught much attention in the economic literature (see Acemoglu et al. 2020; Brynjolfsson et al. 2018; Dauth et al. 2021; Graetz and Michaels 2018). Existing empirical evidence is largely based on robot supplier data provided by the International Federation of Robotics (IFR), which captures *initial sales of robots* by country and industry but cannot account

***Corresponding author: Jens Stegmaier**, IAB Nuremberg, Nuremberg, Germany,
E-mail: Jens.Stegmaier@iab.de

Verena Plümpe, Halle Institute for Economic Research, Halle (Saale), Germany,
E-mail: verena.pluempe@iwh-halle.de

for re-sales of robots.¹ As the share of robots sold to wholesale importers or robot integrators can be quite substantial (Bonfiglioli et al. 2020, Humlum 2019), micro-level data is needed that captures the actual *robot use in production*.

Within the last few years micro-level robot data have been compiled for several countries. Many of them rely on import data that can neither account for re-sales nor capture domestic purchases of robots.² However, the latter is crucial for major robot producing countries such as Germany. In fact, no micro-level data on robot stocks had been available for Germany, one of the most robot-intensive economies worldwide and one of the strongest economies in Europe.

Therefore, we³ introduced a section dedicated to *robot use in production* in German plants in the 2019 wave of the IAB Establishment Panel Survey, which is a high-quality long-standing panel survey that includes about 15,500 plants sampled annually from around 2 million German employers. It is nationally representative as a whole but also at the sector level, for firm-size classes, and across German federal states.⁴ These newly introduced survey questions capture the micro-level demand side of robot use. They address not only the extensive margin of robot use but also its intensive margin, i.e., the number of installed robots over a time span of five years. Furthermore, the data shed light on the heterogeneity of robot use by collecting additional information on robot types. With this plant-level data set it is possible to gain a deeper understanding of robot adoption and its economic and societal consequences.

In the following we introduce the survey methodology of the robot related questions in the IAB Establishment Panel Survey 2019 and outline data processing steps to guide other researchers in their data preparation process. We examine resulting data quality and compare aggregate robot stocks and deliveries on industry-level with the commonly used IFR robot data. Finally, we highlight the research potential of this particular robot data set.

1 The source of information can be installations or shipments depending on the sales channel. However, if robots are re-sold across industries or countries, there is uncertainty regarding timing of installation and final location. For further methodological information, see IFR (2021).

2 For example, Barth et al. (2020) use import data for Norway, Dixon et al. (2021) for Canada, and Cheng et al. (2019) for China. In some cases for sub-samples re-sellers can be excluded, e.g. for Danish data (Humlum 2019) or French data (Acemoglu et al. 2020; Bonfiglioli et al. 2020). Only Koch et al. (2021) use a direct measure of robot use for Spanish firms.

3 The development of the survey questions on robots was a joint initiative of Steffen Müller (Halle Institute for Economic Research and University Magdeburg) and Jens Stegmaier (IAB Nuremberg). Liuchun Deng (Yale-NUS Singapore) and Verena Plümpe (Halle Institute for Economic Research) completed the project team.

4 We use the IAB Establishment Panel, waves 2014–2019. DOI: 10.5164/IAB.IABBP9318.de.en.v1. For more information on the IAB Establishment Panel, see Bechmann et al. (2019).

2 Survey Questions on Robot Usage

2.1 Background and Development

The IAB Establishment Panel of the Institute for Employment Research (IAB) of the Federal Employment Agency is an annual survey of approximately 15,500 establishments representing all sectors and establishment size classes of the German economy. Data collection is conducted by Kantar. It is based on a stratified random sample of establishments that employ at least one employee subject to social security contributions on June 30 each year. The IAB Establishment Panel has repeatedly surveyed establishments in West (East) Germany since 1993 (1996). The data are collected in personal interviews mostly with establishment owners or managers. In order to correct for panel mortality and to reflect the formation of new establishments, the sample is updated every year. The response rate of establishments that have been surveyed repeatedly is over 80 percent. The IAB Establishment Panel focuses on employment-relevant questions and covers a broad range of topics. It is therefore a natural home for data on robots with a labour market policy perspective.⁵

In the following, the focus is on the questions on robots in the 2019 wave of the IAB Establishment Panel. The development of these questions is the result of a joint effort by the IAB and the Halle Institute for Economic Research (IWH). Prior to the development of the questions, preparatory consultations were held with various experts, including the VDMA Robotics + Automation Association as well as with companies in which robots are used. On this basis, a version of the instrument was developed, which was thoroughly tested in the pretest of the IAB Establishment Panel.⁶ The information asked in the pretest about the number of robots used was sometimes difficult and in some cases impossible to answer for some of the respondents. Thus, in the final wording of the survey questions, estimations were also allowed if the actual number of robots was unknown to the respondent.

After the data collection of the main survey, the information provided by establishments on the number of robots used was examined in detail. This revealed a small number of establishments that were either relatively large and indicated to use robots without providing information on the number of robots or, in a few cases, provided contradictory information.⁷ Kantar therefore was instructed to contact selected

⁵ More information on methodological and technical characteristics can be found in Bechmann et al. (2019) or Ellguth et al. (2014).

⁶ In the cognitive pretest about 100 establishments complete the questionnaire. In a second step, the respondents and the interviewers discuss the questionnaire before specific topics - comparable to a qualitative interview - are addressed. See Ellguth et al. (2014) for more information.

⁷ We discuss patterns of missing values in the final data in Section 3.1.

establishments in a second round of data collection. The result of that second field phase was mixed, as such establishments are typically harder to survey. We checked answers of 16 plants which reported a very high number of robots or inconsistently the number of robots over time. Except for one case, there was always a positive confirmation of significant robot use or explanation for small inconsistencies. We further tried to collect data on robot use from 36 larger plants that reported missing robot numbers in the first round. Among those plants, 9 provided robot number at least for 2018 (two of them with more than 100 robots), 4 changed their answer to “no” to whether they used robots between 2014 and 2018, and the remaining 23 plants could not provide any robot numbers or no person capable of providing information could be contacted during the short follow-up survey phase. This follow-up survey was not carried out with too much vigour so as not to compromise the willingness of panel establishments to be re-surveyed for the next wave of the survey.

2.2 Survey Questions

The section of the questionnaire on the use of robots begins by asking the respondents whether robots have been used in the company’s production process in the last five years.

Q77a robot usage

Have you used robots over the last 5 years for operational performance or production?

A robot is any automated machine with multiple axis or directions of movement, programmed to perform specific tasks (partially) without human intervention. This includes industrial robots but also service robots. This excludes machine tools, e.g. CNC-machines.

Possible answers: Yes/No.

When phrasing the question,⁸ we kept the definition of a robot close to the ISO standard⁹ without becoming too complex at the same time. Precisely, in the survey a robot is defined as “an automated machine with multiple axis or directions of

⁸ Original German wording can be found in the Questionnaires https://fdz.iab.de/de/FDZ_Establishment_Data/IAB_Establishment_Panel/IABBP_9320.aspx.

⁹ The ISO 8373:2021 definition of a robot: An *automatically controlled, reprogrammable multi-purpose manipulator, programmable in three or more axes, which can be either fixed in place or fixed*

movement, programmed to perform specific tasks (partially) without human intervention”. Since it is difficult to differentiate between robots and other high-tech machines, e.g., CNC machines, the latter were explicitly excluded. However, the question should also cover service robots.

The subsequent question targets the intensive margin of robot use (Q77b), which is a novelty regarding large micro-level data on robots for Germany.

Q77b number of robots

[if Q77a = yes]

How many robots have you used in total over the last five years? An estimation will suffice. If more robots are used in one robot cell, please count them individually. As already mentioned an estimation will suffice.

[Interviewer: If “none” enter “0”. Please enter “XXXX” if there is no information available for single years.]

Possible answers: Number of robots used in each year from 2014 to 2018.

The respondents are asked to indicate the number of robots per year for the period 2014 to 2018. The pretest showed that some interviewees could not state the exact number of robots per year. Thus, question Q77b asks for an estimated number of robots if the actual number is unknown. Based on the retrospective survey, a quasi-panel data set can be built due to the high willingness to be re-interviewed in the IAB Establishment Panel. For example, 11,268 of the establishments already participated in wave 2018 and 9815 had been interviewed in 2017.¹⁰ However, it should be noted that this constitutes not a real panel data set, where plants report the question on robot use each year. Thus, the intensive margin is not necessarily representative for earlier years.

These extensive and intensive margin questions are followed by three questions that can be used to investigate heterogeneity in robot use.

to a mobile platform for use in automation applications in an industrial environment. Robots and robotic devices – Vocabulary; <https://www.iso.org/obp/ui/\#iso:std:iso:8373:ed-3:v1:en>.

¹⁰ This is about 73%, and respectively 64%, of the plants surveyed in 2019. Equivalently, 56% were surveyed in 2016, 50% in 2015, 44% in 2014.

Q78–80 heterogeneity of robot usage

[if Q77b>=1 in 2018]

Q78: How many of the robots used in 2018 had an acquisition price of less than €50,000? Please consider – if possible – only the purchase price of the robot itself, i.e., without further costs for tools or the integration of the robots into your production.

Q79: How many of the robots used in 2018 are separated from the workers by a protective device, e.g., cage, fence, separate room, light barrier or sensor mat, during regular operation?

Q80: And how many of the robots in use in 2018 have you purchased in 2018?

Possible answers: Number of respective robots for the year 2018.

Survey question Q78 relates to the cost structure and asks for the number of robots with a purchasing price of less than €50,000. Further, question Q79 concerns the number of robots in use that are separated from the personnel by protective devices during their operation. Both of these questions can indicate to some extent whether an establishment uses collaborative robots, which are usually less expensive than large industrial robots and not separated from employees with protection devices. The last question asks how many of the robots used in 2018 were also only purchased in the year in the same year.

3 Data Quality and Data Treatment

3.1 Item Non-response

Almost all surveyed establishments answered the first question on robot usage, less than one percent of respondents did not state if they used robots within the last 5 years. As observed in the pretest, the share of missing values is higher for the intensive margin question (Q77b), where 60 of the 616 robot using establishments (Q77a = 1) did not report the number of robots in 2018 (Table 1). The number of missing values rises the further the observation dates back in time. The share of non-response increases from around 10 percent in 2018 to almost 22 percent in 2014. For the remaining questions (Q78–Q80) the proportion of missing values – even though these questions all refer to the year 2018 – ranges between around 13 and 16 percent. This relatively high share presumably results from the fact that these questions require specific information that may not be readily available to all respondents, such as the purchase price of robots.

Also, these shares of missing values are smaller than for other crucial variables in the IAB Establishment Panel, such as questions on business volumes or total wages. To examine whether there is a structural pattern of missing values with respect to certain establishment characteristics we run a set of cross-sectional OLS regressions. Table 2 depicts results regarding the distribution of missing values for Q77b, where for each year a dummy for missing value is regressed on plant size, regions, and industries. We conclude that the structure of missing values does not significantly differ across German Federal States or industries.¹¹ But there is a significant positive association between the proportion of missing values in robot stocks and establishment size for all years. For example, compared to the smallest plant size class (1–5 employees) the likelihood of a missing value in robot stock in 2018 is 11 percentage points higher for a plant with 200–499 employees. Moreover, we find no significant correlation with other plant characteristics related to education of employees, innovation and technology, business structure and performance.¹²

Table 1: Item non-response of robot questions.

Question	^a Non-missing values	^a Missing values	Missing values as % of valid
Q77a	15,312	127	0.8
Q77b 2018	556	60	9.7
Q77b 2017	524	92	14.9
Q77b 2016	506	110	17.9
Q77b 2015	493	123	20.0
Q77b 2014	482	134	21.8
Q78	515	101	16.4
Q79	532	84	13.6
Q80	535	81	13.1

This table depicts for each of the robot related survey questions the number of non-missing values, the number of missing values, and the share of missing as percent of valid answers. ^aValues are counted as missing if the respondent did not answer the respective question in the IAB Establishment Panel Survey 2019, otherwise as non-missing.

11 Only for 3 non-manufacturing sectors with few observed robot users (<5 each) we find a weak positive association.

12 We regress the number of missing values in robot stock 2018 (Q77b) on the share of employees out of all employees that (i) perform simple tasks (ii) finished an apprenticeship (iii) hold a university degree, log business volume, log export volume, log investment, log wage per employee, dummy variables for foreign ownership, works council, collective wage agreement, 4 innovation measures, and technology status (equal to 1 if equipment is relatively up to date and 0 if rated as average or outdated).

Table 2: Distribution of missing values in 2018 robot stock across plant size, industries, regions.

	(1) MV in 2014 b/se	(2) MV in 2015 b/se	(3) MV in 2016 b/se	(4) MV in 2017 b/se	(5) MV in 2018 b/se
5–9 employees	0.0067 (0.1076)	–0.0382 (0.1036)	–0.0837 (0.0994)	–0.0565 (0.0915)	–0.0124 (0.0754)
10–19 employees	0.0671 (0.0960)	0.0573 (0.0925)	0.0270 (0.0887)	0.0573 (0.0817)	–0.0139 (0.0673)
20–49 employees	0.0222 (0.0819)	–0.0140 (0.0789)	–0.0213 (0.0757)	–0.0014 (0.0697)	–0.0034 (0.0574)
50–99 employees	0.0726 (0.0871)	0.0379 (0.0840)	0.0101 (0.0805)	0.0158 (0.0741)	0.0484 (0.0611)
100–199 employees	0.0613 (0.0836)	0.0212 (0.0805)	–0.0044 (0.0773)	0.0158 (0.0711)	0.0642 (0.0586)
200–499 employees	0.1399 ^c (0.0809)	0.1125 (0.0779)	0.0818 (0.0748)	0.1087 (0.0688)	0.1145 ^b (0.0567)
500–999 employees	0.1235 (0.0983)	0.0293 (0.0947)	0.0383 (0.0908)	0.0830 (0.0836)	0.1161 ^c (0.0689)
1000–4999 employees	0.2100 ^b (0.1042)	0.1920 ^c (0.1003)	0.1293 (0.0963)	0.1462 ^c (0.0886)	0.1519 ^b (0.0730)
5000 + employees	0.3524 ^b (0.1515)	0.3347 ^b (0.1459)	0.3392 ^b (0.1400)	0.3799 ^a (0.1289)	0.4771 ^a (0.1062)
Constant	0.1129 (0.0888)	0.1395 (0.0855)	0.1405 ^c (0.0820)	0.0796 (0.0755)	0.0189 (0.0622)
<i>N</i>	616	616	616	616	616

Cross-sectional OLS regressions, where a dummy for missing value (MV) in robot stock per year (Q77b) is regressed on 10 plant size classes, regional dummies for German Federal States (bula) and sector dummies (br19fb). The latter two are not displayed here. There is no significant correlation of the dependent variable with regions. For 3 non-manufacturing sectors there is a significant correlation with the dependent variable but they each contain less than 5 robot users. ^a0.01, ^bp<0.05, ^cp<0.1.

3.2 Consistency Checks

Further we test whether responses to the robot questions are internally consistent within the IAB Establishment Panel wave 2019. First of all, there are no contradictions when combining the extensive margin (Q77a) and intensive margin (Q77b) survey questions. None of the plants that answered the first question with “yes” reported zero robots for all years on the intensive margin. Also, none of the plants that answered Q77a with “no” reported a positive number of robots on the intensive margin.

Secondly, we check how the number of robot deliveries (Q80) relates to the reported robot stock numbers (Q77b). The reported number of robot purchases in

2018 (Q80) never exceeds the reported robot stock in 2018. This is reassuring and suggests that newly purchased robots are quickly installed and not kept in storage. However, we find small inconsistencies when adding up the robot stock in 2017 with new purchases and compare it to the 2018 robot stock. We would expect that the latter to be lower (in case of depreciation) or equal to the sum of the 2017 stock plus new deliveries. But for 8 plants the reported stock in 2018 is higher. Potential explanations for this inconsistency are other sources of robots besides purchase, e.g., re-use of robots within a multi-establishment firm or merge of establishments. But none of the plants merged within the sample period and more than half of them were single plants and not part of a multi-establishment firm. Thus, we cannot rule out that the inconsistencies are likely to arise due to inaccurate information provided by the respondent.

Also the survey answers regarding the robot purchasing price (Q87) are consistent with the intensive margin (Q77b). The number of robots that costed less than 50,000 Euro in 2018 never exceed the total number of robots in 2018. For the second question on heterogeneity of robots (Q79) we find a minor inconsistency for one plant,¹³ but overall the data is of high internal consistency.

3.3 Data Preparation and Cleaning

Before comparing these micro-level robot numbers with the industry-level IFR data, we perform several data cleaning steps. Firstly, we recode the value -9 to missing values for Q77a–Q80. To ensure internal consistency we set all sub-questions (Q77b, Q78, Q79, Q80) from missing to zero if the extensive margin of robot use (Q77a) is reported as zero.

Additionally we make use of a similar survey question in the IAB Establishment Panel wave 2017 to replace some of the missing information in the intensive margin question (Q77b). In a survey section on digitization a question was posed regarding program-controlled means of production, including CNC machines or industrial robots.¹⁴ The overlap of the survey questions (Q77a, Q77b, and the 2017 question on CNC/robot use) is about 60% due to sample attrition. Considering only the available information, about 80% of the plants that reported robot use in the 2019 wave (Q77a) also stated use of CNC machines or robots in the 2017 wave.¹⁵ The other 20% of robot users (Q77a) state no CNC/robot use in 2017. If we consider more

¹³ For one plant the number of robots separated from workers by a protective device exceeds the total robot stock in that year by 1. We fix the Q79 robot number in our cleaning routine.

¹⁴ These broader measures of automation were previously examined by Zator (2019).

¹⁵ We could also consider missing values for these relative numbers. It would switch to 40% missing, 47% both yes, 13% contradictions.

specifically the robot stock in 2017 (Q77b) it shows that still 51 plants misreport in a way. This could be related to different respondent knowledge or the timing of the survey. For non robot using plants (Q77a = 0) there are no inconsistencies.

Given the overall high internal consistency we use the 2017 question on CNC/robot use to make assumptions about previous non-robot use and replace missing values in the intensive margin question (Q77b). Accordingly, for 17 robot using plants (Q77a) we replace the missing robot stock in 2017 with zero if these plants reported zero high tech machines in 2017. Further, we assume that the 2017 question also refers to previous years and equivalently set missing robot stock numbers for these plants in 2014, 2015, and 2016 to zero.

4 Comparison of IAB Robot Data with IFR Data

In the following, cleaned robot data from the IAB Establishment Panel wave 2019 is compared to the most commonly used robot data set by the International Federation of Robotics (IFR). We map IAB sectors to IFR industries and apply cross sectional survey weights (*hrf_quer*) to obtain aggregate estimates on country and industry level.¹⁶

Despite the different nature of the two data sets, the robot numbers are quite similar in relative terms, as shown in Table 4. According to IFR data 85% of all robots are installed in the manufacturing sector, compared to 75% in IAB, which corresponds to the general focus of IFR data on manufacturing and it's limited coverage of non-manufacturing industries. Further, the relative distribution of robots across industries is highly correlated between the two data sets as depicted in Figure 1, which plots sector-level robot densities (number of robots per 10,000 employees) in 2018. The associated correlation coefficient across all industries is 0.85, with 1 being perfect correlation. If only manufacturing sectors are considered the correlation coefficient is 0.98 for robot density.

The picture is more nuanced when we compare both data sets in absolute terms with respect to (i) aggregate robot stocks, (ii) robot installations in 2018, and (iii) robot distribution across industries. In 2018 the total number of robots according to IAB is with about 108,700 robots only 50% of the total robot stock in IFR, as depicted in Table 4. Partly this deviation can be explained with an upward bias in IFR data. Firstly, in IFR the total robot stock is calculated based on the annually reported deliveries and the assumption that a life time span of a robot is 12 years. This assumption has not only been challenged in the literature (Graetz and

¹⁶ A cross-walk is provided in Table 3. Following Graetz and Michaels (2018) we apply basic data cleaning procedures to IFR to fix minor inconsistencies in the German robot stockss.

Table 3: IAB IFR industry classification crosswalk.

IAB sector (br19fb)	IFR industry	Label
1	A–B	Agriculture/forestry
2	C	Mining
3	E	Energy
4	10–12	Food/luxury
5	13–15	Textiles/clothing
6	16, 17–18	Paper/print/wood
7, 8	19–22	Chemicals/plastics
9	23	Non-metallic mineral
10, 11, 14	24–28	Metal/machinery
12, 13	26–27	Electronics
15	29–30	Automotive
16, 17	91	All other manufacturing
18, 19	F	Construction
20–43	P, 90	All other non-manufacturing
1–43	000	IFR all industries (000)
4–18	D	IFR manufacturing (D)
.	99	IFR unspecified (99)

Own crosswalk. For the comparison of the two data sets in Section 4 the sectors *energy* and *mining* are included in the residual category of *all other non-manufacturing* due to small observation numbers of robot using plants. Equivalently, the *textiles/clothing* industry is integrated into the residual category of *all other manufacturing*.

Michaels 2015), but also acknowledged by IFR itself (Jurkat et al. 2022). Further, the sharp increase in robotic innovations within the last decade (Tilley 2017) provided additional incentives for firms to replace their robot stock faster than just every 12 years. Secondly, the IFR stock is based on robot supplier data, which brings some uncertainty about the final location of a robot if it is sold to system integrators (IFR 2021). Thus, the reported robot sales are overestimated, if some robots are purchased by German headquarters of a firm but then re-exported (Jurkat et al. 2022).

In contrast, the IAB data has the advantage to capture the demand side and measure robot use directly at plant-level. Thus, it’s robot stock does not depend on an assumption regarding the duration of use. Additionally, during the second data collection round¹⁷ comments were obtained that in the automotive sector the life time span of a robot is rather between 3 and 5 years than the assumed 12 years in IFR data. However, the deviation in robot stocks between the two data sets can also partly be explained with an underestimation that arises when the IAB micro-level data is aggregated. In fact, 10% of the interviewed robot using plants did not report the

17 The second round of data collection by Kantar is explained in Section 2.1.

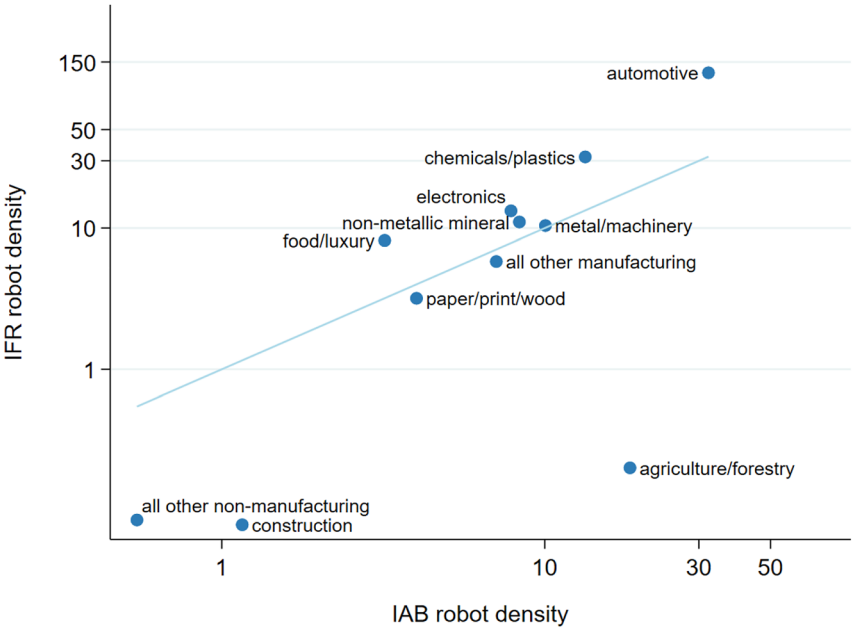


Figure 1: Comparison of robot density across sectors in IAB and IFR data in 2018. (1) Robot density is defined as the number of robots per 1,000 employees. (2) Robot counts in 2018 are aggregated at industry level and divided by the aggregate number of employees per industry to obtain IAB robot density. As the IAB Establishment Panel is representative on industry level, we use the same employment count to derive IFR robot density. (3) The correlation coefficient of robot density across industries between the two data sets is 0.85. If only the manufacturing industries are considered, the correlation coefficient is 0.98.

Table 4: Comparison of robot data from IAB and IFR in 2018 (aggregated level).

Source	IAB	IFR
Stock	108,666	215,795
[95% CI]	[83,178; 134,152]	
Deliveries	22,322	26,723
[95% CI]	[16,018; 28,625]	
Share of stock in manufacturing (in %)	75	85
Share of deliveries to manufacturing (in %)	62	93

This table compares aggregated robot data in 2018 in the IAB Establishment Panel Survey with IFR data. The crosswalk of industry classifications is depicted in Table 3. Cross-sectional survey weights (hrf_quer) are applied to estimate aggregate robot numbers in IAB. They are reported with 95% confidence intervals (CI). Total robot stocks and deliveries are compared in absolute terms for all industries (IFR industry 000) and in relative terms (IFR industry D) as share of robots in manufacturing.

number of robots in 2018, which leads to a downward bias in the aggregate. Secondly, as shown in Deng et al. (2020) robot use in Germany is very rare, with less than 2% of German plants using robots in 2018 and robot stocks highly concentrated among few plants. This relatively small N of robot users and large dispersion in robot stock lead to large confidence intervals for the estimated aggregate robot stock.¹⁸

Comparability of the data sets is higher when we consider the numbers on newly installed/delivered robots. In fact, numbers of newly delivered robots are more reliable than robot stocks in IFR because they do not depend on the robot life time span assumption. Also, the number of newly installed robots in IAB in 2018 is likely to be more precise, because it might be easier for the interviewed person to assess the number of new robots than the total number of installed robots. In fact, the IAB number on new installations in 2018 is about 22,300, while the IFR counts about 26,700 delivered robots, which is still within the confidence interval of the IAB number, as shown in Table 4. This supports the hypothesis of an overestimated IFR stock and overall positively cross-validates the IAB robot data in absolute terms.

We can further compare the distribution of robot stocks on a more disaggregated level across industries, as shown in Figures 2 and 3. For several large manufacturing sectors the robot stock is significantly smaller in IAB compared to IFR. Here the same explanations apply as on aggregate level regarding the robot life time span and missing values. Most prominently, the automotive sector in IFR has with about 108,000 a 4 times higher robot stock than IAB.¹⁹ Another example is the sector *chemicals, pharmaceuticals, and plastic products*, where IFR exhibits twice as many robots as IAB. But Figure 3 shows that in terms of delivered robots in 2018 these gaps are smaller and in the latter case the upper bound of the IAB confidence interval is very close to the IFR value of about 2,200 newly delivered robots. Further indications for an overestimation of IFR robot stock are the sectors *electronics and electrical equipment*, *manufacture of food products*, and *non-metallic mineral products*. Together they account for about 19,900 robots in IFR but only for about 10,700 robots in IAB, while robot deliveries across all these sectors are of similar magnitude in both data sets.

In contrast, IAB reports more robots than IFR for some smaller manufacturing industries and for non-manufacturing sectors. In the sector *manufacture of wood products* IAB exceeds IFR in terms of stock by 400 robots and robot deliveries are even

¹⁸ In fact, if we consider the upper bound of the confidence interval the IAB robot stock increases to about 134,100 robots in total.

¹⁹ Here 9 large robot using plants give a non-response to the intensive margin. For an analysis based mainly on the intensive margin question Q77b, a robustness check excluding the automotive sector is recommended.

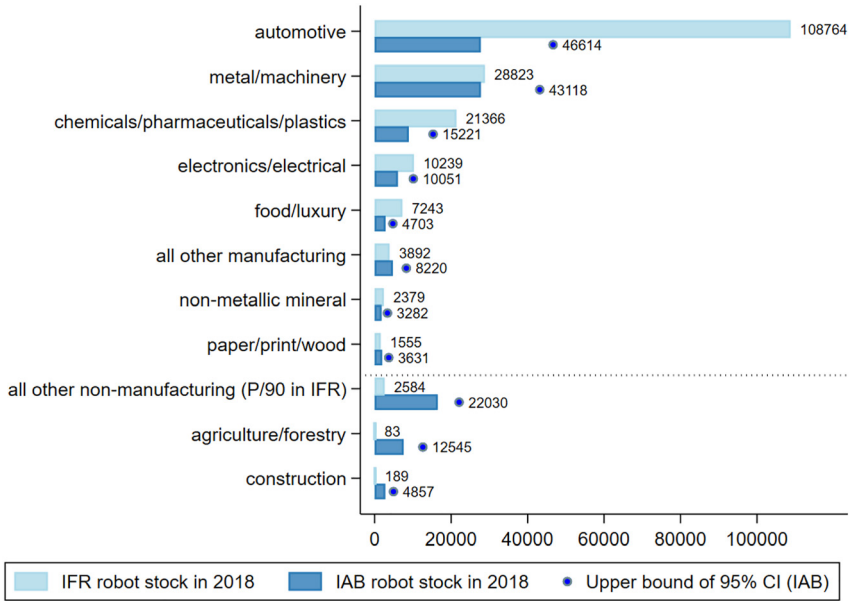


Figure 2: IAB and IFR robot stocks across sectors in 2018. The dashed line separates manufacturing and non-manufacturing sectors.

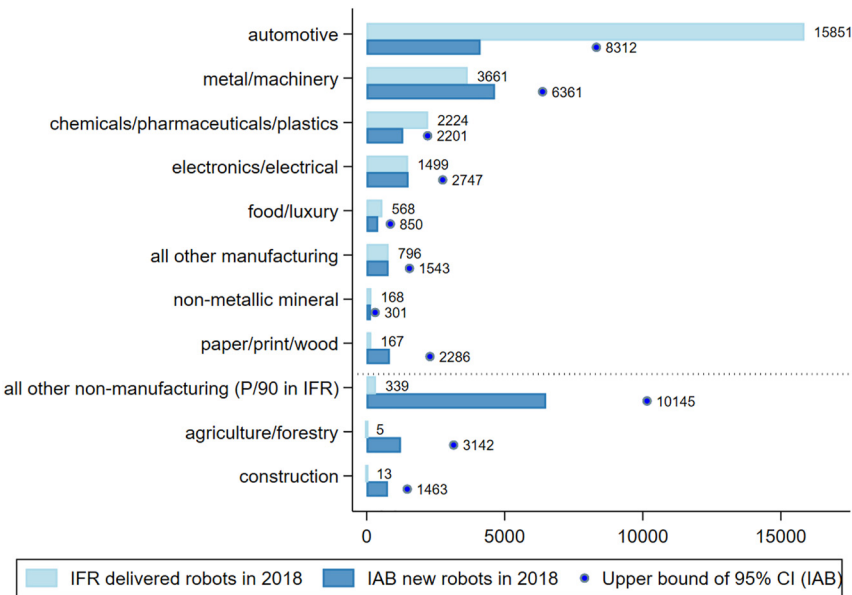


Figure 3: IFR robot deliveries and IAB robot installations in Germany in 2018. The dashed line separates manufacturing and non-manufacturing sectors.

5 times larger in IAB.²⁰ The agricultural sector accounts for 7% of the robot stock in IAB, compared to less than 1% in IFR.²¹ Also for robot deliveries the coverage in IAB is larger, where 1200 new installations are estimated compared to 5 new robots between 2015 and 2018 in IFR. Such small number is implausible, since several types of robots have been developed and implemented in the agricultural sector, such as milking robots, fruit-picking robots, smart farming robots (Shamshiri et al. 2018). But also the estimated IAB robot stock for the agricultural sector due has some statistical uncertainty, as only 16 robot using plants were observed in this sector. A similar pattern shows for the *construction* sector.²² In these cases likely explanations for the deviation are a lack of sector information by robot suppliers or IFR compliance rules, which require that at least four companies must be behind each data point since 2014 (Jurkat et al. 2022). Both aspects are captured in the IFR *unspecified category* (99) which contains about 28,600 robots in 2018.

Altogether, we conclude that given the different nature of the two data sets the robots stocks and deliveries in most manufacturing sectors in 2018 are of similar magnitude if upper bounds of the 95% confidence intervals are considered for sectors where IAB might be underestimated.

5 Analysis Potential and Data Access

Our section on robots in the IAB Establishment Panel resulted in the first data set that contains a *direct* measure of robot use and intensity at the plant level for Germany. The time dimension of the data allows to investigate questions not only on actual robot usage but also about robot adoption as in Deng et al. (2020). That paper shows not only capital deepening through robot purchases by incumbent users in Germany, but also robot technology diffusion as across all sectors many new plants adapt robots. Thus, micro-level data can provide additional relevant insights regarding factor (re-)allocation within sectors that cannot be captured by industry-level data. Compared to existing research the empirical patterns in the robot data of the IAB Establishment Panel confirm micro-level evidence from other European countries. For example, it can be shown that robots in Germany are overall similarly distributed across industries as in Spain, when compared to the results in Koch et al. (2021), but robot adoption seems to be less lumpy than in

20 Also, for the category *all other manufacturing* (furniture, jewellery, sports goods, medical instruments, repair and installation of machinery) the robot stock in IAB is 1,000 robots larger.

21 This sector has an estimated stock of 7,600 robots in IAB compared to only 83 robots in IFR.

22 Based on 12 observations robot using plants the estimated value for the IAB robot stock is about 2700 compared to 189 robots in IFR.

Denmark when compared to findings by Humlum (2019).²³ Also, the German data reveal significant heterogeneity of robot use, e.g., high concentration of robots among few firms, confirming findings by Bonfiglioli et al. (2020).

Furthermore, the rich data of the IAB Establishment Panel allow to investigate specific research questions that cannot be directly examined with aggregate data. For example, Dauth et al. (2021) point out that the effects of robotization might also depend on the institutional structure or the bargaining power of workers, without being able to isolate the direct channels. The IAB Establishment Panel's information on works councils or collective agreements could be used to take initial steps in this regard.

Access to the data is possible via the Research Data Center (FDZ) at the IAB, with both on-site use and remote data access available. Researchers intending to utilize the weakly anonymized data are required to apply for data access. Because direct data access is only possible on-site, test data are also available at the FDZ for researchers to build and test their own programs. These test data match the structure of the actual data, but cannot be used for valid analyses as the variables' values are artificially generated.²⁴ Since the IAB's administrative worker level data are available with a time lag and many analyses require a certain period of time after treatment, such research can only be done since recently. Many interesting questions can thus be explored – the data are readily available.

References

- Acemoglu, D., Lelarge, C., and Pascual, R. (2020). Competing with robots: firm-level evidence from France. In: *AEA papers and proceedings*, Vol. 110, pp. 383–388.
- Barth, E., Roed, M., Schne, P., and Umblijs, J. (2020). How robots change within-firm wage inequality. In: *IZA discussion paper 13605*.
- Bechmann, S., Tschersich, N., Ellguth, P., Kohaut, S., and Baier, E. (2019). Technical report on the IAB establishment panel. In: *FDZ-Methoden report*.
- Bonfiglioli, A., Crino, R., Fadinger, H., and Gancia, G. (2020). Robots imports and firm-level outcomes. In: *CEPR discussion paper 14593*.
- Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? In: *AEA papers and proceedings*, Vol. 108, pp. 43–47.
- Cheng, H., Jia, R., Li, D., and Li, H. (2019). The rise of robots in China. *J. Econ. Perspect.* 33: 71–88.
- Dauth, W., Findeisen, S., Suedekum, J., and Woessner, N. (2021). The adjustment of labor markets to robots. *J. Eur. Econ. Assoc.* 19: 3104–3153.

²³ Within our sample more than $\frac{2}{3}$ of robot using plants increase their robot stock by more than 10% between 2014 and 2018.

²⁴ More information on data access can be found on the Website of the FDZ https://fdz.iab.de/en/FDZ_Establishment_Data/IAB_Establishment_Panel.aspx.

- Deng, L., Plmpe, V., and Stegmaier, J. (2020). Robot adoption at German plants. In: *IWH discussion papers 19/2020*.
- Dixon, J., Hong, B., and Wu, L. (2021). The robot revolution: managerial and employment consequences for firms. *Manag. Sci.* 679: 5586–5605.
- Ellguth, P., Kohaut, S., and Miller, I. (2014). The IAB Establishment Panel—methodological essentials and data quality. *J. Labour Mark. Res.* 47: 27–41.
- Graetz, G. and Michaels, G. (2015). Robots at work. In: *IZA discussion papers 8938*.
- Graetz, G. and Michaels, G. (2018). Robots at work. *Rev. Econ. Stat.* 1005: 753–768.
- Humlum, A. (2019). Robot adoption and labor market dynamics. In: *Working paper*. Princeton University.
- IFR, International Federation of Robotics (2021). *World robotics 2021: industrial robots*. VDMA Services GmbH.
- Jurkat, A., Klump, R., and Schneider, F. (2022). Tracking the rise of robots: the IFR database. *Jahrbücher für Nationalökonomie und Statistik*. 242: 669–689.
- Koch, M., Manuylov, I., and Smolka, M. (2021). Robots and firms. *Econ. J.* 131: 2553–2584.
- Shamshiri, R., Weltzien, C., Hameed, I., Yule, I., Grift, T., Balasundram, S., Pitonakova, L., Ahmad, D., and Chowdhary, G. (2018). Research and development in agricultural robotics: a perspective of digital farming. *Int. J. Agric. Biol. Eng.* 11: 1–11.
- Tilley, J. (2017). *Automation, robotics, and the factory of the future*. McKinsey, Available at: <https://www.mckinsey.com/business-functions/operations/our-insights/automation-robotics-and-the-factory-of-the-future>.
- Zator, M. (2019). Digitization and automation: firm investment and labor outcomes. In: *Working paper*. Available at: SSRN 3444966.