FDZ-METHODENREPORT

Methodological aspects of labour market data

05|2022 EN Revision and new data quality concept due to deviant interviewer behavior in the IAB Job Vacancy Survey

Mario Bossler, Nicole Gürtzgen, Alexander Kubis, Benjamin Küfner, Lukas Olbrich, Silvia Schwanhäuser

Bundesagentur für Arbeit

Revision and new data quality concept due to deviant interviewer behavior in the IAB Job Vacancy Survey

Mario Bossler (IAB), Nicole Gürtzgen (IAB), Alexander Kubis (IAB), Benjamin Küfner (IAB), Lukas Olbrich (IAB), Silvia Schwanhäuser (IAB)

Contents

List of Figures

Zusammenfassung

Dieser Bericht dokumentiert einen konkreten Fall von Interviewerfälschung in der IAB-Stellenerhebung, die anschließende Revision der Daten und das neue Datenqualitätskonzept. Im konkreten Fall wurden zwei Interviewer mit gefälschten Interviews im Zeitraum zwischen Q1 2020 und Q2 2021 identifiziert. Alle Interviews der fälschenden Interviewer wurden rückwirkend aus den Mikrodaten gelöscht. Die Veröffentlichungen über die aggregierte Anzahl der offenen Stellen, die 44 Tage nach Quartalsende erfolgt sind, wurden revidiert. Angesichts der aufgedeckten Fälschungen von Interviews wurden zusätzliche Maßnahmen eingeführt, um solche Fälschungen frühzeitig zu erkennen und zu verhindern. Zu diesen Maßnahmen gehören die Begrenzung der maximalen Interviewzahl pro Interviewer der Interviewer, eine verbesserte Schulung und Überwachung der Interviewer, eine intensivere Auswertung der Audiodateien und vor allem die Einführung eines paradatenbasierten Dashboards zur systematischeren und regelmäßigen Überwachung der Interviewer.

Abstract

This report documents a specific case of interviewer falsification in the IAB job survey, the subsequent revision of the data and the new data quality concept. In the specific case, two interviewers with falsified interviews were identified in the period lasting from Q1 2020 to Q2 2021. All interviews of the falsifying interviewers were retrospectively deleted from the microdata. Publications on the aggregated number of vacancies that are scheduled 44 days after the end of the quarter were revised. In light of the discovered falsification of interviews, additional measures were introduced to detect and prevent such falsifications at an early stage. These measures include limiting the maximum number of interviews per interviewers, improved training and supervision of interviewers, more intensive evaluation of interviewers' audio files, and most importantly, the introduction of a paradata-based dashboard for more systematic and regular monitoring of interviewers.

Keywords

Danksagung

1 Introduction

In the context of survey data quality, interviewers play an important role as they impact response rates, help answer respondents' questions or ensure questionnaire completion. Nonetheless, interviewers have long been identified as a potential key error source in surveys. A particularly serious case of interviewer error is interviewer falsification, defined as the "intentional departure from the designed interviewer guidelines or instructions, unreported by the interviewer, which could result in the contamination of data" (AAPOR, 2003, p. 1). Such departures can take many forms, ranging from complete fabrications of interviews to systematically miscoding responses to filter questions to shorten the interview. While instances of interviewer falsification are frequently reported for face-to-face surveys with interviewers working alone in the field (e.g., Beste et al., 2021; Finn & Ranchhod, 2017; Kosyakova et al., 2019; Schräpler & Wagner, 2005), such reports are rare for telephone interviews (see Porras & English, 2004 for an exception). The main argument for the lower prevalence of falsifications in telephone interviews is that monitoring measures such as live listening to interviews can be implemented more easily and at lower cost (e.g., AAPOR, 2003; Thissen & Myers, 2016). Nonetheless, telephone surveys are not entirely immune to falsifications, especially when interviewers take sophisticated strategies to avoid detection.

A recent example of deviant interviewer behavior in a telephone survey was reported in the IAB-Job Vacancy Survey (IAB-JVS). The Institute of Employment Research (IAB) was informed in August 2021 by the survey institute that one of the telephone studio's interviewers was suspected of manipulating responses to filter questions in order to reduce the length of the interview. After this suspicion had been confirmed, subsequent analyses of the paradata revealed another interviewer who had also deviated from the quality standards in 2020. Both interviewers were discovered prior to the release of the microdata, which are usually published with a time lag of two years at the Research Data Center (FDZ) of the Federal Employment Agency at the Institute for Employment Research (Bossler et al., 2022). Hence, no revision of the microdata was necessary. Unlike the microdata, the aggregate data that are part of the European Vacancy Statistics had already been released for the year 2020. The revision of these data followed in autumn 2021.

The aim of this report is to document the recent case of deviant interviewer behavior in the IAB-JVS, the subsequent revision of the data and the new data quality concept. Section 2 provides a brief description of the IAB-JVS and gives a short overview of the chronology of the recent case of deviant interviewer behavior. Section 3 describes the subsequent revision of the data. Finally, Section 4 describe the measures that were taken by the Institute of Employment Research to improve its processes to prevent and identify potential deviant interviewer behavior in the IAB-JVS.

2 Deviant Interviewer Behavior in the IAB Job Vacancy Survey

The IAB-JVS is a repeated cross-sectional survey of German establishments. The survey collects data on the number and structure of vacancies along with further information on recruitment processes. The IAB-JVS is part of the official European Job Vacancy Statistics, to which it delivers quarterly data on vacancies and employment. The survey's waves start in the fourth quarter (Q4) of each year with a paper & pencil (PAPI) questionnaire combined with an online completion option. In the subsequent three quarters, a subset of the respondents from the initial survey is asked to update parts of the survey information with follow-up interviews using a computer assisted telephone mode (CATI). For the 2019/2020 wave, respondents were reinterviewed by phone as an exception in two additional quarters (in Q4 2020 and Q1 2021) to establish a special COVID panel. To maintain a high response rate, all CATI interviews focus only on a few key questions on vacancies and employment. An important filter question in the phone survey concerns the existence of any vacancy. Reporting zero vacancies considerably reduces the number of follow-up questions and therefore the length of the interview. In what follows, these "zero vacancy" interviews will be referred to as "short" interviews.

The CATI interviews were conducted by a subcontractor of Economix, the responsible survey institute. The survey institute and telephone studio were committed to a regular and close supervision of all involved interviewers. The Institute of Employment Research (IAB) was informed in August 2021 by the survey institute that one of the telephone studio interviewers exhibited a disproportionately large fraction of short interviews in Q2 2021. The interviewer conducted 2,534 out of 9,039 interviews during the period from April to June 2021. While the overall fraction of short interviews was 75 percent, the fraction of short interviews conducted by the interviewer was 97 percent. As noted above, such a strikingly large fraction of "short" interviews may be viewed as an indicator for interviewer fraud, given that interviewers may manipulate responses to filter questions to reduce the length of the interview (Eckman et al., 2014; Hood & Bushery, 1997; Josten & Trappmann, 2016; Kosyakova et al., 2015).

The large fraction of short interviews of this interviewer were discovered within standard plausibility checks during the extrapolation procedure. The large fraction of short interviews gave rise to an implausibly low number of extrapolated vacancies. The plausibility was assessed based on different quality checks, most notably by a comparison of the extrapolated number of registered vacancies with those reported by the official statistics of the Federal Employment Agency. Afterwards, the IAB and the survey institute performed an extensive check of the suspected interviewer's audio files. The amount of evaluated audio files was considerably larger than that of previous controls, which used to be regularly performed at random. The evaluation of the audio files has eventually substantiated the suspicion of deviating (unscientific) interviewer behavior since the identified interviewers deviated from the instructions in a significant number of interviews. After the detection of the deviant interviewer (henceforth referred to as interviewer 1), her/his interviews were removed from the data by the survey institute and the extrapolation of the data was repeated. As the large fraction of affected interviews raised selectivity concerns, the survey insti-

tute provided a detailed analysis of whether the removal resulted in a biased distribution with respect to the stratification variables (see Section 3). Up to this point in time, neither the aggregate data for the second quarter had been delivered to Eurostat nor had the microdata been released yet. As a result, the data for the second quarter was immediately revised prior to any release to the public.

Using the paradata of the phone survey, the IAB then conducted further retrospective analyses of the quarterly data. As the responsible telephone studio had been in charge since Q1 2020, this affected the interviews that were conducted from Q1 2020 onwards. The paradata were available only upon request to the IAB, since up to this point in time the provision of these data was not part of the contractual arrangement between the IAB and the survey institute (as well as its subcontractor). In addition to analyzing the fraction of short interviews, this subsequent analysis also explored potential irregularities in the response behavior for other filter questions. The results of this analysis identified another potentially fraudulent interviewer (henceforth interviewer 2) who also conducted a disproportionately large fraction of short interviews in Q1 2020 (87 percent as compared to the overall fraction of about 70 percent).

After the detection of the fraudulent interviewers, the IAB's quality standards require that all interviews that had been conducted by the two interviewers be removed from the data. For interviewer 1 showing deviant behavior in Q2 2021, this additionally required the retroactive removal of all interviews that were conducted by her/him in the past. While interviewer 1 was not involved in the IAB-JVS in Q1 2021, (s)he conducted more than a quarter of all interviews in 2020. Interviewer 2 conducted 16.4 percent of all interviews in Q1 2020. Similar to the removal of data for Q2 2021, the large fraction of affected interviews required a selectivity analysis. The results of this analysis are documented in Section 3. The revision of the data from the previous quarters was conducted in a similar manner as described above, namely via the exclusion of the affected interviews and a re-weighting of the remaining data after the exclusion. While this procedure resulted in an official revision of the aggregate statistics, it did not affect the release of the microdata, as the latter are published with a time gap of two years.

3 Dataset revision

After interviewer 1 and interviewer 2were identified as falsifiers, it was decided that the IAB-JVS data would be retroactively revised. As mentioned in Section 2, all observations from all past quarterly telephone follow-ups conducted by the excluded interviewers were thus retroactively removed from the data. The sample adjusted for these interviews then formed the new revised net sample on which a new weighting was based. The new weights affect the weighting factors in the microdata and thus also the extrapolation of the data to the aggregate number of vacancies in Germany.

Interviewer 1 was not only employed in Q2 2021, but in the telephone follow-ups since Q1 2020. Therefore, all affected data since the first quarter of 2020 have been revised. Interviewer 2 conducted interviews only in Q1 2020. Both interviewers were not involved in the data collection prior to 2020. As mentioned above, in Q4 2020 and Q1 2021, two samples were interviewed in parallel as part of the IAB-JVS. First, the establishments from the third quarter were surveyed further to gain insights about the effects of the Covid pandemic on establishments by surveying the establishments for two additional quarters (Q4 2020, Q1 2021). We refer to this sample as the COVID Panel or the 2019/2020 wave. Both additional telephone follow-ups were affected by interviewer falsification and had to be revised. Second, a new sample was drawn by default in Q4 2020 to conduct the IAB-JVS in its usual design (2020/2021 wave). These establishments were then also interviewed again by default in the first quarter. This wave from Q4 2020, which was also used to extrapolate Eurostat's vacancy data, was not affected by the falsifying interviewers in Q4 2020 and Q1 2021.

Figure 1 shows the observation numbers for the two partially overlapping survey waves. The 2019/2020 wave was affected by interviewer falsification from Q1 2020 and had to be revised. Thus, the net interviewed sample of this survey wave reduced in all quarters from Q1 2020 to Q1 2021. The underlying case numbers can be seen in the last row of Table 1. Thus, in Q1 2020, the number of cases reduced from 8549 establishments surveyed to a revised sample of 5150 establishments. In Q2 2020, the number of cases reduced from 8296 to 6227. In Q3 2020, the number of cases reduced from 8552 to 6569. In Q4 2020, i.e. the first COVID-related additional survey, the number of cases reduced from 5510 to 3460. In Q1 2021, i.e. the second COVID-related additional survey, the number of cases reduced from 5513 to 3176.

As mentioned above, the deviant interviewer 2 was detected in Q2 2021 during the processing of the data, such that her/his interviews in Q2 2021 were removed prior to any release to the public. Thus, no retrospective revision of the 2020/2021 wave was necessary.

Figure 1: Number of job vacancies and observation; new and old extrapolation, Germany
Q4 2019 – Q1 2021
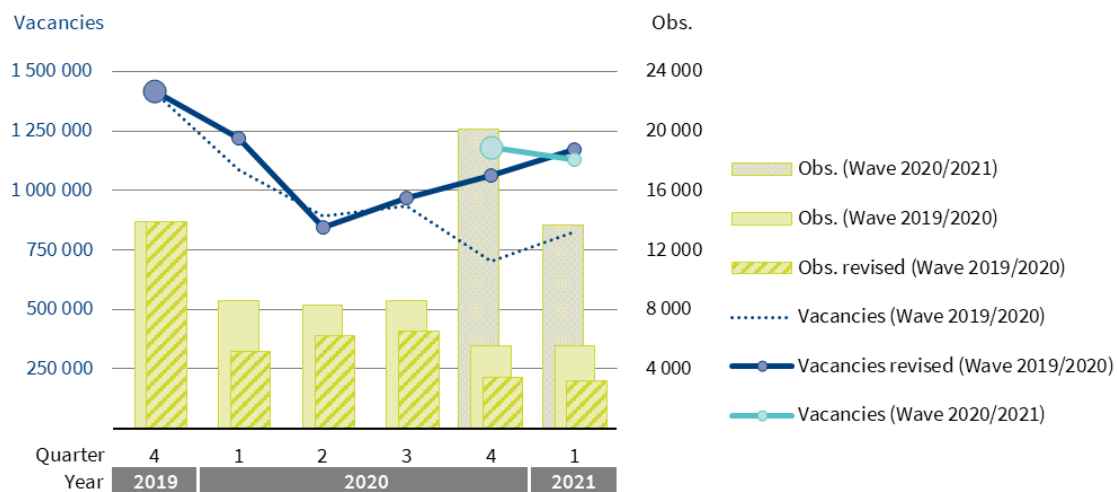
Note: Bars indicate the numbers of observations (right axis) before and after the revision. Light green bars are the original sample sized of the 2019/2020 wage, which includes interviewer falsifications. The striped green bars show the revised sample sizes of the 2019/2020 wave. The grey bars show the sample sizes of the 2020/2021 wave, which we did not need to revise. The dotted blue line represents the (initial) number of vacancies with interviewer falsification of the 2019/2020 wave. The dark blue line represents the revised number of vacancies of the 2019/2020 wave and the light blue line the number of the 2020/2021 wave, which does not need revision.
Source: IAB Job Vacancy Survey. © IAB

In the revision of the data from Q1 2020 to Q1 2021, which was collected as part of the 2019/2020 survey wave, the net sample was significantly reduced. Based on the revised sample, a new weighting of the data was carried out. Tables 1-6 show how the original sample and the revised sample are distributed across industries, establishment size classes and eastern and western Germany, respectively. These three dimensions (industry, establishment size, and region) are important because they determine the stratification of the gross sample and thus also form the basis of the weighting. The gross sample is drawn such that a representative number of establishments are selected from each industry, establishment size class and the two regions. This ensures, among other things, that a representative extrapolation of vacancies can be made for all industries in Germany, as reported to Eurostat.

The numbers of observations in Tables 1, 3 and 5 show that the sample reduced across all industries, size classes and in both regions. However, a sufficient number of observations still remained in the sample after the sample was revised. Therefore, it was possible to calculate new weights without making any adjustments to the general weighting methodology. However, we would like to point out that the nonresponse factor, which is a crucial part of the weighting, not only considers the conventional unit nonresponse of the establishments, but also has to compensate for the non-response due to excluded interviews in the respective telephone follow-ups. Given that the stratification cells still include a large number of observations, projections with the revised sample are still possible for all dimensions of stratification. In any case, researchers should be cautious about calculating weighted statistics for the establishment size category that includes all establishments with at least 1000 employees, since the corresponding number of observations is particularly small after the revision. In addition to the absolute numbers, Tables 2,

4, and 6 present the relative importance of industries, establishment sizes and regions before and after the revision. From a visual inspection the revision hardly affected the distribution of relative frequencies.

The blue lines in Figure 1 show the developments of the weighted absolute number of vacancies in Germany before and after the revision, which is the most important statistic of the IAB-JVS, as it also forms the basis for the mandatory data delivery to Eurostat. In telephone follow-ups Q1 to Q3 2020, the differences before and after the revision the absolute number of vacancies does not change. However, we observe that the pandemic-related drop in vacancies in Q1 2020 was not quite as pronounced as originally thought. It should be noted, however, that this quarter was affected by the pandemic only in the last weeks of March (Bossler et al. 2020). In Q2 2020, on the other hand, the reduction in vacancies is somewhat more pronounced after the revision of the data, which underlines that the pandemic-related slump in labor demand in Germany was severe. In Q3 2020, there are hardly any differences in the absolute number of job vacancies before and after the revision.

The falsification of the interviews was evident with regard to the number of vacancies in Q4 2020 and Q1 2021 of the 2019/2020 wave.[1] Here, the blue lines in figure 1 show that the interviewer falsification significantly underestimated the number of vacancies. This is plausible insofar as the interviewer falsification was noticed by an abbreviation of the questionnaire in that the interviewers intentionally chose or directed the interview towards the filter without vacancies. Therefore, after the exclusion of the affected interviews, the weighted absolute number of vacancies significantly increased. An interesting plausibility check is provided by the 2020/2021 data wave, which is the data that was used to calculate the official vacancy statistic that was unaffected by interviewer falsification. The numbers of vacancies of the unaffected wave are illustrated in the lighter blue line in Figure 1. It shows a similar level as the revised data from the same quarters that needed to be revised as part of the 2019/2020 wave. This check demonstrates the importance to revise the data retrospectively, and it suggests, that after revision we obtain a plausible level of vacancies from the revised 2019/2020 wave, as it is corroborated from two independent data sources.

---

[1] Note that an alternative explanation for larger differences in Q4 2020 and Q1 2021 could be the reduced sample size in these two waves. However, we have no indication that the sample size should bias the estimates in these two quarters.

Table 1: Observations (Obs.) by industry, before and after the revision

Survey Wave 2019/2020 (Q4 2019 – Q1 2021)

Industry	Obs. in Q4_ 2019	Obs. in Q1_2020	Obs. in Q2_2020	Obs. in Q3_2020	Obs. in Q4_2020	Obs. in Q1_2021	Obs. in Q4_2019 (revised)	Obs. in Q1_2020 (revised)	Obs. in Q2_2020 (revised)	Obs. in Q3_2020 (revised)	Obs. in Q4_2020 (revised)	Obs. in Q1_2021 (revised)
A Agriculture, forestry	497	335	305	311	216	224	497	194	218	245	139	123
B Mining, ores and earths	218	144	142	131	104	92	218	88	112	107	56	49
C1 Nutrition, textiles, clothing, furniture etc.	565	366	359	347	222	212	565	199	263	270	137	131
C2 Wood, paper, printing	641	443	428	451	294	298	641	257	325	340	178	173
C3 Chemistry, plastics, glass, construction materials	676	446	419	454	294	303	676	251	321	343	196	174
C4 Metals, metal production	750	500	475	486	314	323	750	265	365	372	206	202
C5 Machines, electronics, vehicles	765	513	471	496	328	327	765	285	361	380	214	185
D Energy utilities	317	208	209	214	148	137	317	114	146	168	93	91
E Water, waste management	519	333	334	334	245	238	519	187	258	249	150	140
F Construction	699	424	377	373	231	235	699	241	278	280	135	129
G Trade, retail, repairs	629	392	364	366	219	226	629	235	246	284	141	128
H Transport, warehouses	566	315	317	314	195	183	566	175	237	241	112	108
I Hospitality	576	318	289	291	173	165	576	181	212	228	104	87
J Information and communication	460	276	267	281	179	169	460	174	203	208	119	93
K Financial services, insurance	535	328	320	324	203	213	535	197	250	261	137	127
L Real estate	660	399	391	399	254	271	660	204	301	311	162	165
M Liberal professions, scientific and technical services	593	356	337	358	226	231	593	211	240	284	149	121
N1 Other commercial services, without temporary employment agencies	581	352	326	345	224	213	581	198	250	245	133	107
N2 Temporary employment agencies	421	248	221	231	138	133	421	146	168	173	80	70
O Public administration	694	462	496	509	338	361	694	345	390	405	229	214
P Education, child care	566	323	346	380	246	242	566	248	268	292	158	139
Q Health and social services	735	368	406	398	234	247	735	277	294	299	137	141
R Art, entertainment, recreation	551	317	291	319	204	186	551	179	216	232	114	105
S Other services	692	383	406	440	281	284	692	299	305	352	181	174
Total	**13906**	**8549**	**8296**	**8552**	**5510**	**5513**	**13906**	**5150**	**6227**	**6569**	**3460**	**3176**

Note: Number of observations by industry for each quarter of the 2019/2020 wave. The first part displays the numbers of observations before revision. The second part displays the numbers of observations after revision.

Source: IAB Job Vacancy Survey. © IAB.

Table 2: Relative (Rel.) number of observations by industry, before and after the revision

Survey Wave 2019/2020 (Q4 2019 – Q1 2021)

Industry	Rel. in Q4_ 2019	Rel. in Q1_2020	Rel. in Q2_2020	Rel. in Q3_2020	Rel. in Q4_2020	Rel. in Q1_2021	Rel. in Q4_2019 (revised)	Rel. in Q1_2020 (revised)	Rel. in Q2_2020 (revised)	Rel. in Q3_2020 (revised)	Rel. in Q4_2020 (revised)	Rel. in Q1_2021 (revised)
A Agriculture, forestry	3,6	3,9	3,7	3,6	3,9	4,1	3,6	3,8	3,5	3,7	4,0	3,9
B Mining, ores and earths	1,6	1,7	1,7	1,5	1,9	1,7	1,6	1,7	1,8	1,6	1,6	1,5
C1 Nutrition, textiles, clothing, furniture etc.	4,1	4,3	4,3	4,1	4,0	3,8	4,1	3,9	4,2	4,1	4,0	4,1
C2 Wood, paper, printing	4,6	5,2	5,2	5,3	5,3	5,4	4,6	5,0	5,2	5,2	5,1	5,4
C3 Chemistry, plastics, glass, construction materials	4,9	5,2	5,1	5,3	5,3	5,5	4,9	4,9	5,2	5,2	5,7	5,5
C4 Metals, metal production	5,4	5,8	5,7	5,7	5,7	5,9	5,4	5,1	5,9	5,7	6,0	6,4
C5 Machines, electronics, vehicles	5,5	6,0	5,7	5,8	6,0	5,9	5,5	5,5	5,8	5,8	6,2	5,8
D Energy utilities	2,3	2,4	2,5	2,5	2,7	2,5	2,3	2,2	2,3	2,6	2,7	2,9
E Water, waste management	3,7	3,9	4,0	3,9	4,4	4,3	3,7	3,6	4,1	3,8	4,3	4,4
F Construction	5,0	5,0	4,5	4,4	4,2	4,3	5,0	4,7	4,5	4,3	3,9	4,1
G Trade, retail, repairs	4,5	4,6	4,4	4,3	4,0	4,1	4,5	4,6	4,0	4,3	4,1	4,0
H Transport, warehouses	4,1	3,7	3,8	3,7	3,5	3,3	4,1	3,4	3,8	3,7	3,2	3,4
I Hospitality	4,1	3,7	3,5	3,4	3,1	3,0	4,1	3,5	3,4	3,5	3,0	2,7
J Information and communication	3,3	3,2	3,2	3,3	3,2	3,1	3,3	3,4	3,3	3,2	3,4	2,9
K Financial services, insurance	3,8	3,8	3,9	3,8	3,7	3,9	3,8	3,8	4,0	4,0	4,0	4,0
L Real estate	4,7	4,7	4,7	4,7	4,6	4,9	4,7	4,0	4,8	4,7	4,7	5,2
M Liberal professions, scientific and technical services	4,3	4,2	4,1	4,2	4,1	4,2	4,3	4,1	3,9	4,3	4,3	3,8
N1 Other commercial services, without temporary employment agencies	4,2	4,1	3,9	4,0	4,1	3,9	4,2	3,8	4,0	3,7	3,8	3,4
N2 Temporary employment agencies	3,0	2,9	2,7	2,7	2,5	2,4	3,0	2,8	2,7	2,6	2,3	2,2
O Public administration	5,0	5,4	6,0	6,0	6,1	6,5	5,0	6,7	6,3	6,2	6,6	6,7
P Education, child care	4,1	3,8	4,2	4,4	4,5	4,4	4,1	4,8	4,3	4,4	4,6	4,4
Q Health and social services	5,3	4,3	4,9	4,7	4,2	4,5	5,3	5,4	4,7	4,6	4,0	4,4
R Art, entertainment, recreation	4,0	3,7	3,5	3,7	3,7	3,4	4,0	3,5	3,5	3,5	3,3	3,3
S Other services	5,0	4,5	4,9	5,1	5,1	5,2	5,0	5,8	4,9	5,4	5,2	5,5
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Note: Relative number of observations by industry for each quarter of the 2019/2020 wave. The first part displays the numbers of observations before revision. The second part displays the numbers of observations after revision.

Source: IAB Job Vacancy Survey. © IAB.

Table 3: Observations (Obs.) by size class, before and after the revision

Survey Wave 2019/2020 (Q4 2019 – Q1 2021)

Establishments with … employees	Obs. in Q4_ 2019	Obs. in Q1_2020	Obs. in Q2_2020	Obs. in Q3_2020	Obs. in Q4_2020	Obs. in Q1_2021	Obs. in Q4_2019 (revised)	Obs. in Q1_2020 (revised)	Obs. in Q2_2020 (revised)	Obs. in Q3_2020 (revised)	Obs. in Q4_2020 (revised)	Obs. in Q1_2021 (revised)
< 10	3680	2083	1945	1952	1245	1233	3680	1282	1421	1482	754	710
10-19	3629	2284	2250	2286	1497	1505	3629	1368	1679	1736	942	853
20-49	3538	2251	2197	2304	1502	1496	3538	1344	1683	1750	959	864
50-249	2117	1353	1336	1425	909	913	2117	799	1009	1134	596	551
250-499	552	355	347	349	215	221	552	217	258	277	123	123
500-999	247	146	151	160	96	104	247	90	122	127	58	59
>= 1000	143	77	70	76	46	41	143	50	55	63	28	16
Total	**13906**	**8549**	**8296**	**8552**	**5510**	**5513**	**13906**	**5150**	**6227**	**6569**	**3460**	**3176**

Note: Number of observations by establishment size for each quarter of the 2019/2020 wave. The first part displays the numbers of observations before revision. The second part displays the numbers of observations after revision.

Source: IAB Job Vacancy Survey. © IAB.

Table 4: Relative (Rel.) number of observations by size class, before and after the revision

Survey Wave 2019/2020 (Q4 2019 – Q1 2021)

Establishments with … employees	Rel. in Q4_ 2019	Rel. in Q1_2020	Rel. in Q2_2020	Rel. in Q3_2020	Rel. in Q4_2020	Rel. in Q1_2021	Rel. in Q4_2019 (revised)	Rel. in Q1_2020 (revised)	Rel. in Q2_2020 (revised)	Rel. in Q3_2020 (revised)	Rel. in Q4_2020 (revised)	Rel. in Q1_2021 (revised)
< 10	26,5	24,4	23,4	22,8	22,6	22,4	26,5	24,9	22,8	22,6	21,8	22,4
10-19	26,1	26,7	27,1	26,7	27,2	27,3	26,1	26,6	27,0	26,4	27,2	26,9
20-49	25,4	26,3	26,5	26,9	27,3	27,1	25,4	26,1	27,0	26,6	27,7	27,2
50-249	15,2	15,8	16,1	16,7	16,5	16,6	15,2	15,5	16,2	17,3	17,2	17,3
250-499	4,0	4,2	4,2	4,1	3,9	4,0	4,0	4,2	4,1	4,2	3,6	3,9
500-999	1,8	1,7	1,8	1,9	1,7	1,9	1,8	1,7	2,0	1,9	1,7	1,9
>= 1000	1,0	0,9	0,8	0,9	0,8	0,7	1,0	1,0	0,9	1,0	0,8	0,5
Total	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**

Note: Relative number of observations by establishment size for each quarter of the 2019/2020 wave. The first part displays the numbers of observations before revision. The second part displays the numbers of observations after revision. Source: IAB Job Vacancy Survey. © IAB.

Table 5:   Observations (Obs.) by region, before and after the revision

Survey Wave 2019/2020 (Q4 2019 – Q1 2021)

Region	Obs. in Q4_ 2019	Obs. in Q1_2020	Obs. in Q2_2020	Obs. in Q3_2020	Obs. in Q4_2020	Obs. in Q1_2021	Obs. in Q4_2019 (revised)	Obs. in Q1_2020 (revised)	Obs. in Q2_2020 (revised)	Obs. in Q3_2020 (revised)	Obs. in Q4_2020 (revised)	Obs. in Q1_2021 (revised)
Western Germany	9692	5945	5757	5961	3826	3809	9692	3546	4319	4592	2441	2197
Eastern Germany	4214	2604	2539	2591	1684	1704	4214	1604	1908	1977	1019	979
Total	**13906**	**8549**	**8296**	**8552**	**5510**	**5513**	**13906**	**5150**	**6227**	**6569**	**3460**	**3176**

Note: Number of observations by region (Eastern Germany and Western Germany) for each quarter of the 2019/2020 wave. The first part displays the numbers of observations before revision. The second part displays the numbers of observations after revision.

Source: IAB Job Vacancy Survey. © IAB.

Table 6:   Relative (Rel.) number of observations by region, before and after the revision

Survey Wave 2019/2020 (Q4 2019 – Q1 2021)

Region	Rel. in Q4_ 2019	Rel. in Q1_2020	Rel. in Q2_2020	Rel. in Q3_2020	Rel. in Q4_2020	Rel. in Q1_2021	Rel. in Q4_2019 (revised)	Rel. in Q1_2020 (revised)	Rel. in Q2_2020 (revised)	Rel. in Q3_2020 (revised)	Rel. in Q4_2020 (revised)	Rel. in Q1_2021 (revised)
Western Germany	69,7	69,5	69,4	69,7	69,4	69,1	69,7	68,9	69,4	69,9	70,5	69,2
Eastern Germany	30,3	30,5	30,6	30,3	30,6	30,9	30,3	31,1	30,6	30,1	29,5	30,8
Total	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**

Note: Relative number of observations by region (Eastern Germany and Western Germany) for each quarter of the 2019/2020 wave. The first part displays the numbers of observations before revision. The second part displays the numbers of observations after revision.

Source: IAB Job Vacancy Survey. © IAB.

4 New data quality concept

To closely monitor the field process and prevent deviant interviewer behavior in future waves of the IAB-JVS, we developed a new data quality assurance concept that is based on three pillars that are described in more detail below: Increased quality standards for the survey institute and telephone studio, a monitoring dashboard and an IAB evaluation of audio recordings.

Overall, the survey data is subjected to independent quality controls by various parties: the telephone studio, the survey institute and the IAB. All three institutes conduct their own independent data quality assurance measures and, hence, contribute to the efficacy of the quality assurance concept.

4.1 Increased quality standards for the survey institute and telephone studio

The key actors for ensuring high data quality are the survey institute and the telephone studio, which are responsible for conducting the interviews. As an immediate consequence of the detected interviewer fraud, the IAB and the survey institute stopped working with the previous telephone studio and started a cooperation with another telephone studio. The new telephone studio and the survey institute had to fulfill high requirements with respect to quality standards set by the IAB. These requirements are listed below:

1. *Experience with scientific surveys*
An important requirement for the new telephone studio is that it can prove experience with scientific surveys. This experience is important as it ensures that the interviewers and the administrative staff understand the needs and goals of standardized interviews applicable to scientific surveys. Moreover, interviewers are trained in scientific quality standards and, thus, are able to conduct high quality interviews.

2. *Limit the maximum interviewer workload*
To minimize the risk of bias due to deviant interviewer behavior and interviewer effects, each interviewer's workload must be restricted. This ensures limited impact of single interviewers on the data and data quality in general. Further, a subsequent detection of deviant interviewer behavior after the field period and, thus, the removal of those cases will have a lesser effect on results. For the IAB-JVS, the maximum number of interviews per month and interviewer is set to 100 interviews and the maximum number of interviews per quarter and interviewer is set to 300 interviews.

3. *Interviewer training*
Another important aspect of the new data assurance concept is enhanced interviewer training. Each interviewer participates in an extended training on the entire project and principles of interviewing for scientific surveys. Moreover, prior to each quarterly data collection each interviewer receives an update on quarter specific information such as changes in the

questionnaire and learnings from previous quarters. In each training, one member of the IAB researcher team participates to monitor and increase the quality of the training by highlighting the importance of standardized data collection. In addition, all training materials must be approved by the IAB. Moreover, retraining of interviewers is based on results of the supervision process and evaluation of audio recordings.

4.   Supervision

An important quality standard for standardized interviews is supervision. Whenever interviews are conducted, at least one supervisor is present to monitor the ongoing interviews. The supervisor can follow the interviewer-respondent-interaction by listening to their conversation and observing the interviewer's screen. Hence, the supervisor is able to closely monitor the entire interview process. To ensure quality and increase transparency, all supervisions are documented in a supervision protocol. In addition, the IAB has the opportunity to live supervise interviews. Based on all supervisions, deviant behavior of interviewers can be detected in real time and rapid feedback is provided to interviewers immediately after the interview.

5.   Provision and evaluation of audio recordings

To obtain insights into the data collection process, the telephone studio is required to provide audio recordings of the interviews. Further, the survey institute uses these audio recordings to supplement supervision and live monitoring. Evaluation results of audio recordings are used to provide feedback to interviewers and improve interviewer training (for more details see section 4.3.).

6.   Frequent provision of field reports

On a weekly basis, IAB receives a field report about the development of the field. This field report contains important outcome rates such as response rates, refusal, and contact rates (based on AAPOR-standards), information about the distribution of interviews on different establishment characteristics, as well as interviewer-specific analyses.

7.   Provision of survey data and paradata

Similar to the field report, IAB is provided with survey data as well as paradata in a standardized format by the survey institute on a weekly basis. This allows for regular data quality controls conducted by the IAB. These data are primarily fed into the dashboard described below. In addition, these data can be used for in-depth analyses of suspicious cases.

4.2   Development of a dashboard to monitor interviewers

As mentioned above, the IAB receives data for quality control every week. To provide fast feedback to the survey institute on potentially problematic interviewers, the data is processed and visualized in a dashboard developed in R Shiny (Chang et al., 2021). The dashboard is updated automatically as new data arrives.

We use survey data and paradata to build a variety of quality and performance indicators that are based on previous literature (e.g., Schwanhäuser et al., 2022) and experiences with further IAB-based surveys (Beste et al., 2021; Kosyakova et al., 2019). The dashboard contains descriptive figures on general field progress and interviewer-level indicators.

Field progress is depicted using standard AAPOR rates (e.g., response rate, contact rate, (hard/soft) refusal rates) (AAPOR, 2016). These rates allow for close monitoring of the overall progress of the survey. In later IAB-JVS waves, data from previous quarters are used to check for potential deviations in the field progress and enable timely measures. Another general indicator is based on call record data. Here, the dashboard shows the distribution of the number of contact attempts before successful interviews and its development over time. As before, these data are particularly useful when compared between quarters and years.

The main purpose of the dashboard are interviewer-level indicators that enable the assessment of interviewers in manifold ways. The interviewers' productivity is shown in terms of their overall workload and their daily productivity (e.g., number of (successful) calls, share of successful calls). Interviewers with suspiciously high workloads or implausibly high shares of successful calls could be detected by examining these indicators. To assess the interviewers' quality of work, we use multiple data quality indicators and item-level timestamp data. Data quality indicators include the share of "short" interviews, the average interview duration, the average share of triggered follow-up questions, and the average share of rounded responses. These indicators point to a variety of potentially deviant interviewer behaviors such as shortcutting the interview or reinforcing satisficing behaviors. Figure 2 shows a screenshot of the dashboard page on the indicators. The timestamp data provide questionnaire module-specific durations. In the dashboard, distributions of these durations are illustrated by interviewer and for the entire sample. Interviewers who speed through the interview or shorten questionnaire text should have particularly short duration and can be identified in the dashboard.

As illustrated in Figure 2, the dashboard is enriched by various options (e.g., selecting field weeks, selecting specific interviewers) that allow for examining suspicious interviewers or field weeks in greater detail.

4.3   Evaluating audio recordings of interviews

While the dashboard is a useful tool to obtain insights on the interviewers' work, it can only deliver indications for potentially deviant behavior. More detailed insights on interviewer behavior are provided by audio recordings of the interviews. Here, the interviewer's voice is recorded during the interview, so that deviations such as skipping questions, shortening questionnaire text, or inadequate probing techniques can be detected by listening to the recording. Of course, listening to all audio files is too time and work intensive, which is why we set the focus on interviewers flagged as rather suspicious with regard to the indicators used in the dashboard.

For the examination of audio recordings, at least three recordings of interviews conducted by the three or more most suspicious interviewers are selected each week and evaluated using a standardized coding scheme. In the evaluation process, each question and the introductory text are checked separately for possible deviations from the questionnaire. In addition, any other anomalies during the interview are noted. Furthermore, the codes from the coding scheme are merged with the interview data. This allows for classifying omitted questions as correctly omitted due to an active filter or as incorrectly omitted. As the dashboard is updated on a weekly basis and audio recordings are available shortly after the interview, deviant behavior can be detected early. In that case, we send feedback to the survey institute that takes further measures. Using this approach, we evaluate at least 120 interviews each quarter.

Figure 2: Screenshot of dashboard, simulated data.
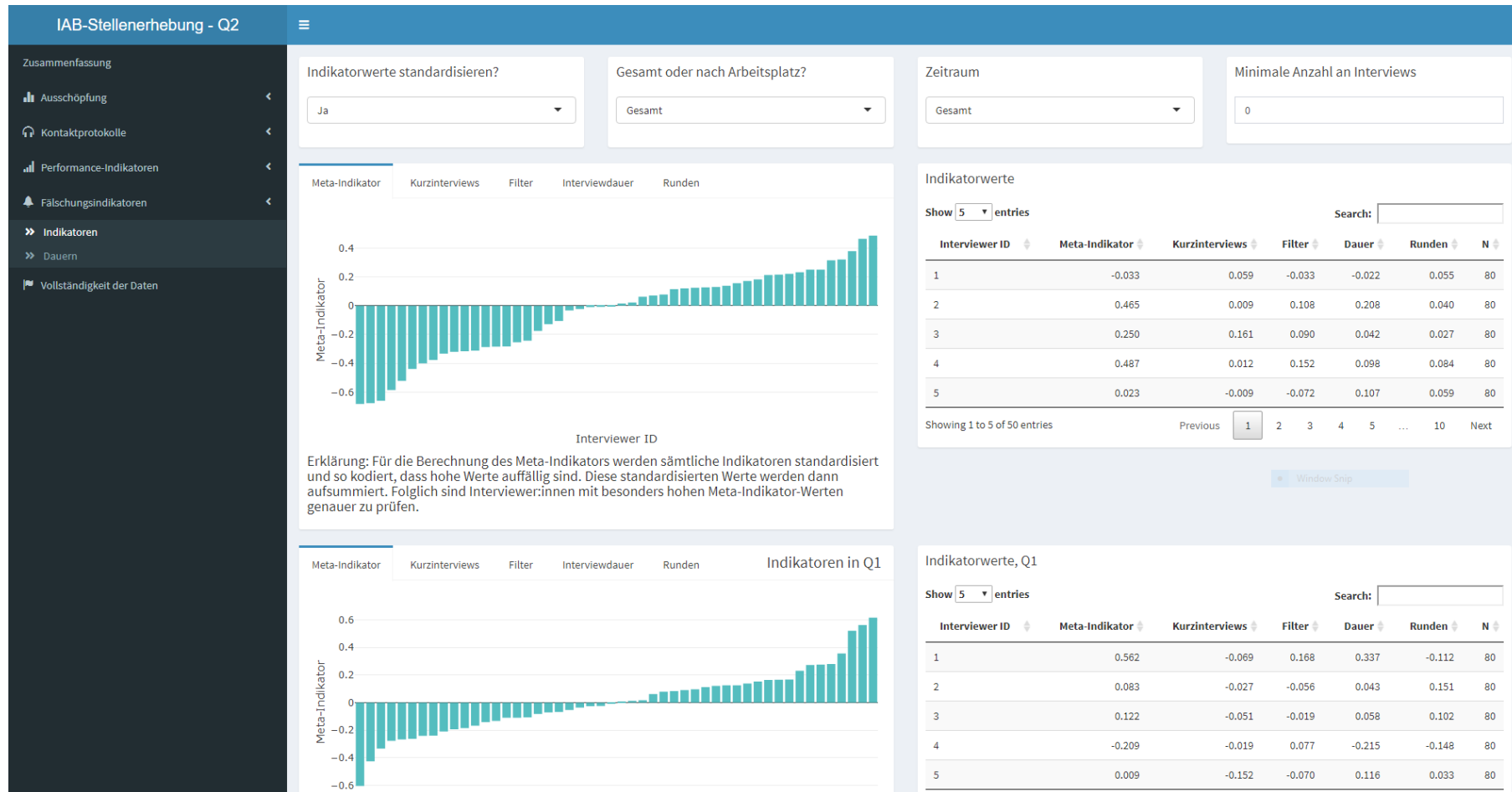
5 Conclusion

The aim of this report was to document the recent case of deviant interviewer behavior in the IAB-JVS, the subsequent revision of the data and the new data quality concept. In this recent case, two falsifying interviewers were discovered in Q2 2021. Unlike the microdata, the aggregate data that are part of the European Vacancy Statistics had already been released for the year 2020. The falsifications therefore required a revised release of the aggregate data, which took place in autumn 2021. The results of the revision show that the pandemic-related drop in vacancies in Q1 2020 was not quite as pronounced as originally suggested by the falsified data. In Q2 2020, on the other hand, the reduction in vacancies is somewhat more pronounced after the revision of the data. In Q3 2020, there are hardly any differences in the absolute number of job vacancies before and after the revision.

The large number of interviews affected by the falsifications underlines the need to improve the processes to detect and prevent interviewer falsifications. Against this background, this report additionally documents the measures taken by the IAB to detect and prevent such falsifications in the future. These measures include a limit of the maximum interviewer workload, improved interviewer training and supervision, an intensified evaluation of interviewers' audio files and most notably, the implementation of a dashboard based on paradata to monitor interviewers in a more systematic and regular manner. The regular update of the dashboard and monitoring of the audio recordings shortly after the interviews shall ensure that potential deviant behavior can be detected early, in which case the survey institute will be immediately informed to take further measures.

6 References

AAPOR (2003): *Interviewer falsification in survey research: Current best methods for prevention, detection, and repair of its effects.*

AAPOR (2016): Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition.

Beste, Jonas; Olbrich, Lukas; Schwanhäuser, Silvia (2021): Interviewer:innenkontrolle im Panel Arbeitsmarkt und soziale Sicherung (PASS). (FDZ-Methodenreport, 04/2021 (de)), Nürnberg, 11 S.

Bossler, Mario; Gürtzgen, Nicole; Kubis, Alexander; Küfner, Benjamin; Popp, Martin (2022): The IAB Job Vacancy Survey: Establishment survey on labour demand and recruitment processes, waves 2000 to 2019 and subsequent quarters 2006 to 2021. (FDZ-Datenreport, 08/2022 (en)), Nürnberg, 23 S.

Buch, Tanja; Stöckmann, Andrea (2018): Entwicklung der Substituierbarkeitspotenziale auf dem Hamburger Arbeitsmarkt – Aktuelle Ergebnisse auf Basis einer Neubewertung der Substituierbarkeit von beruflichen Kerntätigkeiten. IAB-

Regional, Berichte und Analysen aus dem Regionalen Forschungsnetz, IAB Nord, 03/2018, Nürnberg.

Chang, Winston; Cheng, Joe; Allaire, JJ; Sievert, Carson; Schloerke, Barret; Xie, Yihui; Allen, Jeff; McPherson, Jonathan; Dipert, Alan; Borges, Barbara (2021): shiny: Web Application Framework for R. R package version 1.6.0. https://CRAN.R-project.org/package=shiny

Eckman, Stephanie; Kreuter, Frauke; Kirchner, Antje; Jäckle, Annette; Tourangeau, Roger; Presser, Stanley (2014): Assessing the mechanisms of misreporting to filter questions in surveys. Public Opinion Quarterly, 78(3), 721-733.

Finn, Arden; Ranchhod, Vimal (2017): Genuine Fakes: The Prevalence and Implications of Data Fabrication in a Large South African Survey. World Bank Economic Review, 31(1), 129-157.

Hood, Catherine; Bushery, John (1997): Getting more bang from the reinterviewer buck: Identifying 'at risk' interviewers. In Proceedings of the Survey Research Methods Section, American Statistical Association, 820–824.

Josten, Michael; Trappmann, Mark (2016): Interviewer Effects on a Network-Size Filter Question. Journal of Official Statistics, 32(2), 349-373.

Kosyakova, Yuliya; Olbrich, Lukas; Sakshaug, Joseph; Schwanhäuser, Silvia (2019): Identification of interviewer falsification in the IAB-BAMF-SOEP Survey of Refugees in Germany. (FDZ-Methodenreport, 02/2019 (en)), Nürnberg, 33 S.

Kosyakova, Yuliya; Skopek, Jan; Eckman, Stephanie (2015): Do interviewers manipulate responses to filter questions? Evidence from a multilevel approach. International journal of public opinion research, 27(3), 417-431.

Porras, Javier; English, Ned (2004): Data-driven approaches to identifying interviewer data falsification: The case of health surveys. In Proceedings of the Survey Research Methods Section, American Statistical Association, 4223-4228.

Schräpler, Jörg-Peter; Wagner, Gert. G. (2005): Characteristics and impact of faked interviews in surveys – An analysis of genuine fakes in the raw data of SOEP. Allgemeines Statistisches Archiv, 89(1), 7-20.

Schwanhäuser, Silvia; Sakshaug, Joseph W.; Kosyakova, Yuliya (2022): How to catch a falsifier: Comparison of statistical detection methods for interviewer falsification. Public opinion quarterly, 86(1), 51-81.

Thissen, M. Rita; Myers, Susan K. (2016): Systems and processes for detecting interviewer falsification and assuring data collection quality. Statistical Journal of the IAOS, 32(3), 339-347.

Walwei, Ulrich (2018): Trends in der Beschäftigung Älterer. Rahmenbedingungen für betriebliche Personalpolitik. In: WSI-Mitteilungen, Jg. 71, H. 1, S. 3–11.

Imprint

Corresponding author
Benjamin Küfner
Phone: +49 911 179-8183
Email: benjamin.kuefner@iab.de