1601 2026

SOEP Survey Papers
Series C - Data Documentations (Datendokumentationen)

IAB-SOEP-Migration – 2024: Sampling, Nonresponse, and Weighting of the Sample M10

Hans Walter Steinhauer, Rainer Siegers, Felix Süttmann, Reiner Gilberg, Folkert Aust, Martin Kleudgen, and Christian Hunkler

DIW BERLIN

Running since 1984, the German Socio-Economic Panel (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin.

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing.

The SOEP Survey Papers is comprised of the following series:
Series A – Survey Instruments (Erhebungsinstrumente)
Series B – Survey Reports (Methodenberichte)
Series C – Data Documentation (Datendokumentationen)
Series D – Variable Descriptions and Coding
Series E – SOEPmonitors
Series F – SOEP Newsletters
Series G – General Issues and Teaching Materials

The SOEP Survey Papers are available at **http://www.diw.de/soepsurveypapers**

Please cite this paper as follows:

Hans Walter Steinhauer, Rainer Siegers, Felix Süttmann, Reiner Gilberg, Folkert Aust, Martin Kleudgen, and Christian Hunkler. 2026. IAB-SOEP-Migration – 2024: Sampling, Nonresponse, and Weighting of the Sample M10. SOEP Survey Papers 1601. Series C. Berlin: DIW/SOEP

ISSN: 2193-5580 (online)

DIW Berlin
German Socio-Economic Panel (SOEP)
Anton-Wilhelm-Amo-Str. 58
10117 Berlin
Germany

soeppapers@diw.de

IAB-SOEP-Migration – 2024: Sampling, Nonresponse, and Weighting of the Sample M10

Hans Walter Steinhauer[1], Rainer Siegers[1], Felix Süttmann[1], Reiner Gilberg[2], Folkert Aust[2], Martin Kleudgen[2], and Christian Hunkler[1]

[1]German Institute for Economic Research (DIW Berlin)
[2]infas Institute for Applied Social Science

March 17, 2026

Abstract
This paper provides details on the sampling design, the fieldwork, as well as nonresponse and population adjustments for the 2024 sample M10 of the IAB-SOEP-Migration survey. The survey is conducted in cooperation between the Institute for Employment Research (IAB) and the Socio-Economic Panel (SOEP). Sample M10 refreshes the previous IAB-SOEP-Migration samples adding 2,128 households with at least one dual citizen having a nationality of one of the member states of the European Union (EU), except Germany.

1 Introduction

Panel studies are widely recognized as valuable tools in social and behavioral research, offering insights into dynamic patterns and long-term trends within a given population. These studies rely heavily on the participation of individuals who are willing to contribute their time and experiences over an extended period of time. However, maintaining panel participation and ensuring that the sample remains "representative" of its population is challenging. A crucial aspect of panel studies is the refreshment and augmentation of the sample. Adding new participants into an ongoing panel compensates for attrition while also capturing changes in the population over time. This process is essential for maintaining the "representativeness" of the panel and for preserving the integrity of the findings.

The Socio-Economic Panel (SOEP) is one of the longest-running panel studies in the field of economic and social sciences. The SOEP provides data on various socio-economic aspects of individuals and households over an extended period since 1984. By including participants from different subgroups of the German population, the panel ensures diversity. Specifically, the IAB-SOEP-Migration samples M1, M2, M7, M8a, M8b, and M8c focus on migration to Germany motivated by job opportunities. More details on the IAB-SOEP-Migration samples are provided by Brücker et al. (2014). The SOEP samples M8a, M8b, and M8c cover a specific subgroup of the population living in Germany, namely skilled labor immigration from non-EU countries to Germany; for details see Steinhauer, Trübswetter, and Zinn (2022), Steinhauer et al. (2025a) and Steinhauer et al. (2025b). These samples provide a basis to evaluate the Skilled Labor Immigration Act (Fachkräfteeinwanderungsgesetz), which initially came into effect on March 1, 2020, with the goal of facilitating the immigration of skilled workers from non-EU countries to Germany. Sample M10 focuses on dual citizens living in Germany with at least one nationality of the member states of the EU.

This paper details the approach applied to sample and weight the new sample M10. In Section 2, we provide information on the target population, the sampling frame, and details on the sampling design. The fieldwork process and its results are described in Section 3. Section 4 details the different steps of weighting. Characteristics of the final weights are displayed in Section 5, while Section 6 summarizes.

2 Population and Sampling Design

The target population for the refreshment sample M10 covers foreigners with an EU nationality aged 18 years and older living in private households. We sample from this population using the population register as a sampling frame. The German population register ("Einwohnermelderegister" or "Melderegister") is the administrative database that records demographic information of residents living in Germany. It serves as the primary source of official population data and plays a crucial role in selecting samples from the general population in Germany. The population register is maintained by local registration offices ("Einwohnermeldeämter") in each municipality. It is a decentralized system, with each registration office responsible for keeping accurate and up-to-date records of individuals residing within the respective municipality. Residents are required by law to report any changes, such as change of address or significant life events like births, deaths,

marriages, and divorces, to their local registration office. The information contained in the population register typically includes the following details:

- personal information including full name, date of birth, gender, and nationality;
- the current residential address;
- information on marital status, such as whether a person is single, married, divorced, or widowed; and
- details about family relationships, including spouses, children, and other dependents residing at the same address.

Unfortunately, the population register does not contain any information on household compositions. Moreover, registration offices do not necessarily have to provide all the information listed above. Especially information on family relationships that would be useful in constructing households is not provided (see §46 BMG). Thus, sampling households directly from this register cannot be done without further steps. We account for this in the weighting procedure described in Section 4.

When drawing samples from the population register, the information listed above can be used to stratify the sample or focus on specific parts of the population. Further, when using the German population register as a sampling frame, a two-stage cluster sampling design is appropriate. Here, we select the municipalities hosting the population register as primary sampling units (PSU) at the first stage and at the second stage sample individuals listed in the registers within each municipality as secondary sampling units (SSU). Sampling individuals provides us with an address necessary to contact the household via an anchor person ("Ankerperson").

The sampling approach, applied by the infas Institute for Applied Social Science (hereafter "infas"), is a stratified two-stage cluster sampling design. The municipalities selected at the first stage are stratified by federal states municipality size and selected proportional to size. The measure of size for the probability proportional to size (PPS) sampling is defined as the number of foreigners with an EU nationality aged 18 years and older living in the municipality. At the second stage, within the municipality, we sample individuals with an EU nationality aged 18 years and older from the register. These individuals serve as anchor persons to obtain access to the households. Table 1 displays the distribution of the sample by federal states.

3 Fieldwork results and Response Rates

The 33,287 addresses provided by the registration offices were validated by infas. Prior to field work, sampled anchor persons were sent an invitation letter by mail outlining the data protection regulations and explicitly emphasizing that participation was voluntary. Before the start of the main field phase, the anchor persons were invited to a short pre-recruitment survey. In this survey we asked for their consent to participate or to refuse any further contact. Interviewers contacted the anchor persons between March 2024 and December 2024 for the main interview. Table 2 details the results of the fieldwork on the household-level. In total, there were 2,128 complete or partial interviews, resulting in a response rate on the household-level, calculated according to American Association for Public Opinion Research (2023), of $RR2 = \frac{2128}{25{,}595} = 0.083$. The refusal rate is $REF1 = \frac{3944}{25{,}595} = 0.154$.

Table 1: Number of target persons in the population by federal state.

	Population		Sample	
Federal state	Number	Proportion	Number	Proportion
Schleswig-Holstein	88,501	0.020	1038	0.031
Hamburg	81,564	0.019	1018	0.031
Lower Saxony	325,318	0.074	3482	0.105
Bremen	24,576	0.006	300	0.009
North Rhine-Westphalia	906,329	0.206	8673	0.261
Hesse	462,687	0.105	3004	0.090
Rhineland-Palatinate	240,385	0.055	1132	0.034
Baden-Württemberg	838,777	0.191	4241	0.127
Bavaria	850,902	0.194	4094	0.123
Saarland	56,786	0.013	449	0.013
Berlin	295,539	0.067	2251	0.068
Brandenburg	38,837	0.009	552	0.017
Mecklenburg-Western Pomerania	30,422	0.007	349	0.010
Saxony	69,888	0.016	1368	0.041
Saxony-Anhalt	42,387	0.010	808	0.024
Thuringia	41,829	0.010	528	0.016
Total	4,394,727	[a)]1.002	33,287	1.000

Note: [a)]Proportions do not sum up to one due to rounding.

Both rates are similar to previous refreshment samples in the IAB-SOEP-Migration survey. Some addresses were not deployed in the field because either interviewers did not contact them before the end of the field period or the number of desired interviews was already reached; see AAPOR code 3.11 in table 2. Some addresses were categorized as non-eligible because the household has moved abroad, was out of sample or screened out; see AAPOR codes 4.1 to 4.5 in Table 2.

4 Cross-Sectional Weighting

The derivation of weights typically involves three main steps (Brick & Kalton, 1996). First, design weights are computed as the inverse of the inclusion probability (see Section 2). In the second step they are adjusted to account for unit nonresponse, a process referred to as sample weighting adjustment by Kalton and Kasprzyk (1986). In the final step, weights are calibrated to ensure that the estimates align with known population parameters, such as totals, ratios, or specific distributions. This step is referred to as population weighting adjustment. Kroh, Siegers, and Kühne (2015) comprehensively outline the general weighting strategy employed by the SOEP and the incorporation of new samples.

Please note that the population register lists individuals without providing information about their household context. However, the IAB-SOEP-Migration Survey is a household panel survey in which all adults are interviewed. Therefore, a household with two or more individuals, for example, has a higher probability of selection compared to a single-person

Table 2: Fieldwork results on the household-level according to American Association for Public Opinion Research (2023).

Final Disposition Code	Number	Proportion
1. Interview		
(1.10) Complete	1085	0.033
(1.20) Partial	1043	0.031
2. Eligible, Non-Interview		
(2.10) Refusal	3567	0.107
(2.12) Break-off	377	0.011
(2.20) Non-contact	4569	0.137
(2.31) Dead	91	0.003
(2.32) Physically / mentally unable	630	0.019
(2.33) Language problems	1705	0.051
(2.36) Miscellaneous	9443	0.284
3. Unknown eligibility, non-interview		
(3.11) Not attempted or worked	3085	0.093
4. Not Eligible		
(4.10) Out of sample / screened out	7	0.000
(4.20) Moved abroad	474	0.014
(4.40) Untraceable	6072	0.182
(4.50) Non-residential building	1139	0.034
Total	33,287	1.000

Note: Proportions might not sum up to one because of errors due to rounding.

household. To determine a household's sampling probability, we assign sampling probabilities to all members of the existing households, even though these individuals were not initially sampled as anchor persons. This requires accounting for various characteristics used in stratifying the sample. Once the sampling probabilities for each household member are identified, we can calculate the household sampling probabilities within each household. The household weights are then derived as the inverse of these household sampling probabilities. Kroh, Kühne, Jacobsen, Siegert, and Siegers (2017) provide a formal documentation of this procedure.

To address potential selectivity resulting from the field work and nonresponse, we employ models that first estimate the success of contacting a household. Given a successful contact, we model the decision-making process of each household and estimate the probability to participate. The models are based on:

1. contact: all households that have been deployed to the field and have been approached,
2. participation: all households that have been successfully contacted.

Given the limited availability of data on households sampled, we use area-level information regarding the residential environment provided by infas360 (see https://datenkatalog.infas360.de/). In addition, we use design information such as stratification variables and data on processes of the field work including interviewer information and attributes of the

first contact attempt enter the equation. Finally, the information on the anchor persons provided by the population register, such as for example sex or date of birth were used, too.

4.1 Sample Weighting Adjustments

In the second step of adjusting the design weights, it is crucial to identify strong predictors of nonresponse. To achieve this, we conduct an iterative process, examining all variables included in the previously described data. We select variables that significantly influence the participation decision using bivariate regression analyses. Next, we remove variables from the set of significant predictors if their absolute correlation with one another is 0.95 or higher. This step prevents the inclusion of highly correlated variables in the analysis. The remaining variables form the basis of a preliminary nonresponse model. To obtain a final model, we apply a variable selection procedure in both forward and backward directions, using the Bayesian Information Criterion (BIC) as selection criterion. This approach ensures a more parsimonious model by retaining only the most relevant variables. Table 3 presents the resulting model estimates of the probability to be successfully contacted as well as the response propensities used for deriving weighting adjustments.

In the two-stage procedure of contact and participation, we look at contact first. We see that the participation in the pre-recruitment survey, naturally increases the probability for a successful contact. In terms of the interviewers first contact attempt, we find the timing of the first contact in the morning as well as in the evening to be more successful compared to a first contact at noon or afternoon. Moreover, a first contact at the weekend was more likely to be successful than first contact attempts during the week. Looking at the interviewer characteristics, we find Interviewers born between 1957 and 1990 as well as interviewers born in 1991 or later to be more likely to contact a household successfully. In terms of occupation, interviewers having another occupation than working or pensioner and interviewers with an intermediate level of education (indicated by CASMIN) were more effectively in terms of contacting households.

Table 3: Models estimating the probability to be successfully contacted as well as the response propensities used for deriving weighting adjustments.

Variable Value	Contact Estimate (Std. Error)	Participation Estimate (Std. Error)
(Intercept)	−0.025 (0.029)	−2.285*** (0.049)
1. Attributes of first contact attempt		
Pre-recruitment survey participated	0.605*** (0.026)	
Contact time evening	0.344*** (0.018)	
Contact time morning	0.146*** (0.032)	

Table 3 continued.

Variable Value	Contact Estimate (Std. Error)	Participation Estimate (Std. Error)
Contact day weekend	0.177*** (0.022)	
2. Attributes of the interviewer		
Year of birth 1991 or later	0.372*** (0.057)	
Year of birth 1957-1990	0.366*** (0.020)	
Occupation other[1]	0.457*** (0.022)	0.282*** (0.047)
CASMIN 2a, 2b, 2c_gen, 2c_voc	0.277*** (0.020)	
Foreign language English	−0.412*** (0.052)	
Native Language German	−0.506*** (0.022)	
Sex female		0.189*** (0.049)
3. Attributes of the anchor person		
Sex female	0.084*** (0.017)	
Age (categorized) 18-25		−0.445*** (0.098)
4. Attributes of the residence		
Distance to train station highest quartile	−0.117*** (0.022)	
Distance to public transport stop highest quartile		0.226*** (0.049)
Number of inhabitants per census grid lowest quartile		0.181*** (0.051)
5. Attributes of the residential area (block of buildings)		
Share of graduates (Ph.D.) above 2.5%	0.230*** (0.029)	
Area of block in m^2 lowest quartile	−0.080*** (0.020)	
Share of Catholics high	−0.085*** (0.023)	
Purchase index (e/m^2) lowest quartile		−0.302*** (0.060)

Table 3 continued.

Variable Value	Contact Estimate (Std. Error)	Participation Estimate (Std. Error)
6. Attributes of the residential area (neighborhood)		
Index on availability of public transport low	-0.324^{***} (0.052)	
Share of Persons with a Polish migration background highest quartile	-0.087^{***} (0.020)	-0.392^{***} (0.059)
Share of Persons with an Italian migration background highest quartile	-0.168^{***} (0.021)	-0.253^{***} (0.057)
Share of Persons with a Turkish migration background lowest quartile	0.123^{***} (0.023)	
Share of Persons with a Soviet migration background lowest quartile	0.152^{***} (0.023)	
Share of Persons with a Soviet migration background highest quartile	-0.100^{***} (0.021)	
Purchasing power (€per year) lowest quartile	0.090^{***} (0.022)	
Number of apprentices lowest quartile	0.104^{***} (0.024)	
Number of students highest quartile	0.147^{***} (0.022)	
Area of district in m^2 highest quartile	-0.075^{***} (0.021)	
Rent index (€/m^2) highest quartile	-0.176^{***} (0.023)	
Purchase index (€/m^2) highest quartile		0.382^{***} (0.055)
Number of persons aged 15 without professional qualification lowest quartile		0.192^{***} (0.051)
Index of socio-economic deprivation 2	0.205^{***} (0.037)	
Index of socio-economic deprivation 3	0.219^{***} (0.028)	
N	25,504	17,644

Notes: Dependent variable: Success in contacting the household (1 = yes, 0 = no), participation of the household (1 = yes, 0 = no). [1] other than working or pensioner (e.g. unoccupied, student, or housewife). Significance indicated by $*** \equiv p < 0.001$, $** \equiv p < 0.01$, and $* \equiv p < 0.05$. The model is estimated using the function `glm()` with a cloglog link function in R (R Core Team, 2025).

Interviewers with German as a native language as well as interviewers who can speak

English as a foreign language were less likely to contact a household. When the anchor person sampled, was a woman, a contact was more likely to yield positive outcome. We also find regional characteristics at different levels to influence the likelihood of a successful contact. When the household resides in a long distance to the nearest train station there is a lower chance for successfully contacting it. When being located within a block of building a household is more likely to be contacted when the share of graduates is above 2.5% and less likely when the area of the block is low or the share of catholic people is high. The most effects are at the level of the neighborhood of a household. Here we see that a low index of availability of public transport relates to lower success rates in contacting households. We find the share of persons from different groups of migration background to affect successful contacts in different directions. For neighborhoods with high shares of Polish and Italian migrants the contacting a household failed more frequently. In neighborhoods with a low share of Turkish migrants the likelihood for contact is higher. The coefficients for the share of Soviet migrants (positive for lowest and negative for highest quartile) indicate a negative linear relationship, i.e., the higher the share of Soviet migrants the less the likelihood for a successful contact of the household. When the household is located in a neighborhood with low purchasing power interviewers are more likely to contact the household. The same holds for households in neighborhoods with a low number of apprentices and a high number of students. Households in larger areas are less likely to be contacted successful and also households in neighborhoods with a high purchasing index are contacted less. Finally, households in neighborhoods with a low values for the index of socio-economic deprivation (GISD) are less likely to be contacted successfully.

When it comes to participation we find fewer predictors. In terms of interviewer characteristics we find female interviewers as well as interviewers with an occupation other than working or pensioner to be more likely to pursue a household to participate in the panel. Looking at attributes of the anchor persons, we find that persons aged 18 to 25 are less likely to participate in the survey. Households residing in areas with low numbers of inhabitants per census grid (100m x 100m) and a large distance to the nearest public transport stop have a higher propensity to participate. For the purchasing index we find different on block and neighborhood level. On the block level a household is less likely to participate, when the purchase index is low. In contrast, a household in a neighborhood with a high purchase price index is more likely to participate. Households located in neighborhoods with high shares of Polish and Turkish immigrants have a lower participation propensity. Finally, households in a neighborhood where the number of persons aged 15 without professional qualification is low are more likely to participate in the panel.

4.2 Population Weighting Adjustments

In the final step of the weighting process, we use post-stratification and raking techniques to adjust the weights obtained previously. This adjustment aligns the weights with known population totals, as well as joint and marginal distributions. The specific method used depends on the available population data, with a comprehensive overview provided by Kalton and Flores-Cervantes (2003). The resulting weights from this step form the basis for deriving both cross-sectional and longitudinal weights for subsequent survey waves, beginning with wave 2. The population parameters and distributions used in these adjustments were provided by the Federal Statistical Office, based on data from the German

Microcensus. Margins used in the post-stratification process include the number of households with at least one person of the population by

- household typology (single vs. other),
- municipality size, and
- regions (federal states categorized by north, east, south, and west); and

the number of persons of the population by sex and age group. For information on the categories of the different variables, see Siegers, Steinhauer, and Schütt (2022).

5 Characteristics of Weights

Due to stratification and disproportional allocation of households, there is some variance in the design weights. Adjusting the initial design weights for unit nonresponse by multiplying them with the inverse of estimated participation probabilities increases variation in the second weighting step resulting in the adjusted design weights. The population weighting adjustments to derive calibrated weights further add to the variation and magnitude of weights, see Table 4. The resulting calibrated weights are provided in the variable `hhrf0` included in the data set `hpathl` as well as in the variable `bohhrfm10` in the data set `hhrf`. After the integration step, a further post-stratification step was carried out. Here, weights of all households in the SOEP are adjusted with respect to the standard marginal distributions. Using the resulting standard SOEP weighting factors (`hhrf` included in `hpathl` and `phrf` included in `ppathl`), the sample M10 cases can then be analyzed jointly and comparatively in combination with all other SOEP observations.

Table 4: Characteristics of weights after the steps of the weighting process (rounded to integer values).

		Quantiles							
Step	Min.	10%	25%	50%	75%	90%	Max.	Mean	SD
DW	8	36	53	81	114	181	2608	107	126
SWA	63	297	474	768	1279	2046	30,184	1109	1413
PWA	19	225	382	688	1274	2233	10,672	1048	1154

Abbreviations: SD = standard deviation, DW = design weighting, SWA = sample weighting adjustment, PWA = population weighting adjustment.

6 Summary

Sample M10 refreshes the previous IAB-SOEP-Migration samples adding 2,128 households with at least one dual citizen having a nationality of one of the member states of the European Union (EU), except Germany. Unlike to previous migration samples of the IAB-SOEP-Migration survey the sample is drawn from the population register using a stratified two-stage sampling design. Concerning field work, we find a few potential selectivities. Using the information available from field work, interviewers, anchor persons,

and regional attributes of the household, we account for possible selectivities due to contact and participation in our weighting strategy. We find attributes of interviewers as well as regional characteristics to have an impact on the success of contacting households. In contrast, participation is mainly impacted by characteristics of the direct neighborhood of the households. We adjust design weights by the inverse of the estimated probabilities for contact and participation to correct for unit nonresponse. In the final step we use post-stratification to make the distributions conform to those provided by the microcensus. Weights for this sample are provided on the household as well as on the person-level. Moreover, households have been seamlessly integrated into SOEP-Core.

References

American Association for Public Opinion Research. (2023). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (10th ed.). Retrieved from https://aapor.org/wp-content/uploads/2024/03/Standards-Definitions-10th-edition.pdf

Brick, J. M., & Kalton, G. (1996). Handling missing data in survey research. *Statistical methods in medical research*, *5*(3), 215–238. doi: 10.1177/096228029600500302

Brücker, H., Kroh, M., Bartsch, S., Goebel, J., Kühne, S., Liebau, E., … Schupp, J. (2014). *The new IAB-SOEP migration sample. An introduction into the methodology and the contents.* (SOEP Survey Papers No. 216). Berlin: DIW-Berlin.

Kalton, G., & Flores-Cervantes, I. (2003). Weighting methods. *Journal of official statistics*, *19*(2), 81–97.

Kalton, G., & Kasprzyk, D. (1986). The treatment of missing survey data. *Survey methodology*, *12*(1), 1–16.

Kroh, M., Kühne, S., Jacobsen, J., Siegert, M., & Siegers, R. (2017). *Sampling, nonresponse, and integrated weighting of the 2016 IAB-BAMF-SOEP Survey of Refugees (M3/M4)—revised version* (SOEP Survey Papers No. 477). Berlin: DIW/SOEP. Retrieved from https://www.diw.de/documents/publikationen/73/diw_01.c.572346.de/diw_ssp0477.pdf

Kroh, M., Siegers, R., & Kühne, S. (2015). Gewichtung und Integration von Auffrischungsstichproben am Beispiel des Sozio-oekonomischen Panels (SOEP). In *Nonresponse bias* (pp. 409–444). Springer.

R Core Team. (2025). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Siegers, R., Steinhauer, H. W., & Schütt, J. (2022). *SOEP-Core v37 Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2018)* (SOEP Survey Papers No. 1106). Berlin: DIW/SOEP.

Steinhauer, H. W., Trübswetter, P., Fendel, T., Ivanov, B., Cardozo Silva, A., Süttmann, F., … Zinn, S. (2025a). *IAB-SOEP-Migration – 2022: Sampling, Nonresponse, and Weighting in Sample M8b* (SOEP Survey Papers No. 1497). Berlin: DIW-Berlin.

Steinhauer, H. W., Trübswetter, P., Fendel, T., Ivanov, B., Cardozo Silva, A., Süttmann, F., … Zinn, S. (2025b). *IAB-SOEP-Migration – 2023: Sampling, Nonresponse, and Weighting in Sample M8c* (SOEP Survey Papers No. 1536). Berlin: DIW-Berlin.

Steinhauer, H. W., Trübswetter, P., & Zinn, S. (2022). *SOEP-Core - 2020: Sampling, Nonresponse, and Weighting in the IAB-SOEP Migration Studies M7 and M8* (SOEP Survey Papers No. 1105). Berlin: DIW-Berlin.