IAB-SOEP-MIG Survey Data: Outline

Content characteristics	
Topics	**Migration history:** Year of immigration; migration history; search behaviour and information channels; social networks; **Education history:** highest obtained schooling and vocational degrees; years of schooling; education acquired at home and abroad; acknowledgement procedure of foreign credentials; language proficiency; **Employment history:** Employment; self-employment; unemployment in Germany and in foreign countries; **Labour market background:** Earnings; full- and part-time employment; working hours; benefit assistance, reservation wage; participation in active labour market policies; **Return migration:** Return migration intensions; return migration; limited survey of returned immigrants; **Miscellaneous:** Remittances to home countries; life satisfaction; risk preferences; social integration and acceptance;
Data unit	Anchor persons: Persons registered at the federal employment agency the first time after 1995 who immigrated to Germany themselves or who are children of immigrants. Family members: All Family members of anchor persons over 16 years
Number of cases	**Wave 1:** 4964 persons in 2723 households **Wave 2:** 3752 persons in 2021 households **Wave 3:** 4770 persons in 2768 households **Wave 4:** 3832 persons in 2155 households **Wave 5:** 3408 persons in 1913 households **Wave 6:** 2962 persons in 1692 households **Wave 7:** 2515 persons in 1424 households
Period covered	**Wave 1 (bd):** Year 2013 **Wave 2 (be):** Year 2014 **Wave 3 (bf):** Year 2015 **Wave 4 (bg):** Year 2016 **Wave 5 (bh):** Year 2017 **Wave 6 (bi):** Year 2018 **Wave 7 (bj):** Year 2019
Time reference	Date of survey, migration biography retrospectively since first leave of country of birth, retrospective questions on life course
Regional structure	German federal state (Bundesland)
Territorial allocation	As of date of survey

Methodological characteristics

Method of data collection	The initial sample, which was surveyed for the first time in 2013, has been drawn from all individuals in the IEB data that have appeared there for the first time after 1994. Clusters of 2,500 people each were formed using geocoded data from the RDC of the IAB, resulting in a total of 6,725 geographical sample units (PSUs). In the next step, all people with a migration background were identified on the basis of their nationality. Among the people of German nationality, the onomastics procedure was used to search for further people with a migration background. The onomastics procedure was only carried out in a partial sample of 600 out of a total of 6,725 PSUs. The PSUs were selected on the basis of stratified random sampling. All persons without a migration background were removed from the data. A total of 250 out of the 600 PSUs were taken into account for the field survey. In the selected 250 PSUs, the gross sample consists of 80 addresses (households). Simulated random walk procedures were implemented for the sampling. Through a so-called screening interview with the anchor people, the households that meet the criteria of the sample design were identified.
	Since 2015 there has also been a refreshment sample to take account of panel attrition and changing immigration pattern in younger years. The new respondents are only those who immigrated between 2009 and 2015. The refreshment sample has also been drawn from the IEB data and is disproportionate to the country of origin, as certain countries are overrepresented (e.g. Bulgaria) and other countries are underrepresented (e.g. Turkey).
Institutions involved	**Commissioned by:** Cooperation between the Institute for Employment Research (IAB) and the Socio-Economic Panel Study at DIW Berlin **Carried out by:** TNS Infratest Sozialforschung
Frequency of data collection	Yearly (Panel
File format and size	STATA (several files)
File architecture	Person-related, household related and biographical datasets. The structure of the datasets is organised as the GSOEP from DIW Berlin

Data access

Data access	Scientific Use File
Degree of anonymisation	Factually anonymous
Sensitive variables	None
Standards for the citation of data and data documentation	**Data:** This study uses the factually anonymous data of the IAB-SOEP Migration Sample. Data access was provided via a Scientific Use File supplied by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) or the SOEP at DIW Berlin.

Data documentation:
Brücker, Herbert, Martin Kroh, Simone Bartsch, Jan Goebel, Simon Kühne, Elisabeth Liebau, Parvati Trübswetter, Ingrid Tucci and Jürgen Schupp (2014), The new IAB-SOEP Migration Sample: an introduction into the methodology and the contents, SOEP Survey Paper 216.