RatSWD

Working Paper Series

Working Paper No. 101

Migration and Statistics

Sonja Haug

July 2009
The *RatSWD Working Papers* series was launched at the end of 2007. Since 2009, the series has been publishing exclusively conceptual and historical works dealing with the organization of the German statistical infrastructure and research infrastructure in the social, behavioral, and economic sciences. Papers that have appeared in the series deal primarily with the organization of Germany’s official statistical system, government agency research, and academic research infrastructure, as well as directly with the work of the RatSWD. Papers addressing the aforementioned topics in other countries as well as supranational aspects are particularly welcome.

*RatSWD Working Papers* are non-exclusive, which means that there is nothing to prevent you from publishing your work in another venue as well: all papers can and should also appear in professionally, institutionally, and locally specialized journals. The *RatSWD Working Papers* are not available in bookstores but can be ordered online through the RatSWD.

In order to make the series more accessible to readers not fluent in German, the English section of the *RatSWD Working Papers* website presents only those papers published in English, while the German section lists the complete contents of all issues in the series in chronological order.

Starting in 2009, some of the empirical research papers that originally appeared in the *RatSWD Working Papers* series will be published in the series *RatSWD Research Notes*.

The views expressed in the *RatSWD Working Papers* are exclusively the opinions of their authors and not those of the RatSWD.

The RatSWD Working Paper Series is edited by:

Chair of the RatSWD (2007/ 2008 Heike Solga; 2009 Gert G. Wagner)

Managing Director of the RatSWD (Denis Huschka)
Migration and Statistics

Sonja Haug

Bundesamt für Migration und Flüchtlinge (sonja.haug[at]bamf.bund.de)

Abstract

The field of empirical migration and integration research is characterised by a wide range of research questions, theoretical approaches and data sets. Research based on official statistics has to deal with different data sets on migration and foreign population resulting in different numbers. Developments in official statistics concentrate on the improvement of data quality. The census 2010/2011 or the projected central population register are important issues, e.g. in respect of sampling and weighting of migrants in surveys. The concept of migration background in the microcensus is a mayor enhancement in population statistics and has become widely accepted. It is recommended to implement questions on migration background in the census 2010/2011 too. The most important accessible data sets in the field of empirical integration research are the microcensus and the German Socio-Economic Panel; there is still untapped potential for analysis in this area. The supplementation of large surveys with a migrant sample is an appreciated trend. Most important challenges for empirical migration and integration research are the development of sampling methods for migrant population (including onomastics and topomastics), studies on new and small migrant groups, research projects in the country of origin, longitudinal migrant surveys and the development of measurement instruments.

Keywords: Population, Migration, Integration, Migration Background
Research questions

The field of empirical migration and integration research is characterised by a wide range of research questions, theoretical approaches and data sets.

Old Research Questions

Migration research focuses on classic questions: How many migrants are on the move? What are their countries of origin and destination? What are the determinants of migration? Integration research in Germany payed particular attention to the "guest worker" population. Empirical research focused on the descriptive analysis of integration (Mehrländer et al. 1996), or on explanation of the migration and integration process (Esser 1980).

Theoretical Developments and New Research Questions

Migration research is differentiating into several focus points. New research fields in international migration are migration from Central and Eastern European countries, questions relating to demography and the idea of "replacement migration", return migration and irregular migration. New interdisciplinary approaches in the field of sociology or economics employ theories of social capital and social networks to explain migration decisions (cf. Boswell and Mueser 2008). Transnational migration, chain migration, circular migration and migrant communities are further important research issues (Pries 1997). Analysing international migration dynamics from a development perspective using a micro-level decision model and event history method is an example for theoretical and empirical progress (Massey et al. 2008).

New developments in migration and integration theory arise with discussion on a general model on intergenerational integration (Esser 2008; Kalter 2008b). Actual research is also concentrating on different dimensions of integration and the mechanisms underlying the observable types of integration (cf. Kalter 2008b).

Empirical research into migrant integration is an emerging field and frequently overlaps with issues like social inequality and exclusion (Kalter 2003); theories relating to the sociology and economy of migration are increasingly being integrated in this area of research. Additionally, integration of the second generation is a core issue (Haug and Diehl 2005). The PISA study and other research highlights the problems of migrant children in the education system (Stanat and Christensen 2006; Kristen 2005); labour market integration is another
topic (Granato 2003). New research therefore focuses on the transition from primary to secondary school and/or to vocational education and to occupation. Analysis of the integration process is a key trend in this area, and panel surveys and longitudinal studies play an ever greater role, although adequate data sets are rare. One important issue is the evidence that naturalised migrants show greater integration success but still differ from people without a migration background (Haug 2002; Salentin and Wilkening 2003). Therefore, the formerly widespread concept of "foreigners" is losing ground and the concept of "migration background" is becoming generally accepted (see for an overview on measurement aspects Diefenbach and Weiß 2006).

Research on internal migration is another topic, for example migration decision making in terms of the labour market and the life cycle (Huinink and Kley 2008; Kalter 1997; Wagner 1989; Windzio 2004).

Policy-Related Research Questions

Analysis of legal migration and estimation of illegal migration is one of the central policy-related issues (BAMF 2008). The national contact point of the European Migration Network (EMN) at the BAMF aims to improve the availability of and access to information on migration and asylum at EU level, facilitating work on annual policy reports, small scale studies and annual reports on migration, asylum and return migration statistics to support the policy and decision-making process. The target population for policy-related surveys is made up of former labour migrants and their families (Venema and Grimm 2002; Babka von Gostomski 2008; Weidacher 2000) or ethnic German repatriates (Haug and Sauer 2007), while special research projects focus on female migrants (Boos-Nünning and Karakasoglu 2006), the "second generation" (Haug and Diehl 2005) or Muslims (Brettfeld and Wetzels 2007). Integration policy-makers often tend to think of migration in terms of deficits, like deficits in German language knowledge, educational success or the labour market integration of migrants. What is needed, however, is a totally new approach towards prioritising resources for migrants in areas such as language and professional skills. One of the emerging fields of policy-related data analysis is the development of integration indicator sets in municipalities, the federal states and on national level (Siegert 2006; Worbs and Friedrich 2008; KGSt 2006; Filsinger 2008). The BAMF integration report working paper series provides an overview of official statistics and empirical social research covering a wide range of fields of integration. The Commissioner for Migration, Refugees, and Integration has published a set of integration indicators and will publish a more detailed report at a later date.
Comprehensive data analysis on integration aspects is also conducted in different fields of policy-related research; examples include the official reports on education (Bildung in Deutschland), poverty (Armuts- und Reichtumsbericht) or families (Familienbericht).

**Status Quo: Data Bases and Access**

Different data bases are available for research purposes (Diehl and Haug 2003; Haug 2005; Schönwälder et al. 2008).

*Data Bases for Research into Immigration and Emigration (Population Flow)*

Since registration with the local authorities is obligatory in Germany, legal migration can be measured directly. Flow data result from decentralized local resident registration offices, which supply aggregate data on immigration and emigration, categorised by citizenship, country of origin or destination, age and gender to the statistical offices of the states and to the Federal Statistical Office. A lack of reliable figures on the number of irregular migrants is inherent to this process, and this leads to underestimation of immigration and the size of the foreign population. A further problem is that register information generally produces case-based statistics, so that official aggregate migration statistics tend to overestimate migration. Restructuring of migration statistics for foreigners is ongoing, with the result that persons crossing the border several times a year are no longer counted multiple times. The migration data does not include specifics on the duration of stay or the residence permit status, so it is not possible to differentiate between long-term and short-term migrants.

A second source of migration flow data is the Central Register of Foreigners (AZR, BAMF 2008). The major drawback of this register is that it only records the immigration of foreigners, and it cannot therefore be considered a comprehensive migration register. The main advantage of the register is that it provides person-based statistics and the option of distinguishing between short-term and long-term migrants (persons who have stayed for at least one year). In 2004 the register was adjusted, and this resulted in a reduction in the number of foreigners in Germany from 7.3 million to 6.7 million (Opfermann et al. 2006).

Alongside these comprehensive migration statistics, there are also statistical systems for the registration of different groups of immigrants using specific entrance options such as those for ethnic German repatriates, Jewish migrants, family re-unification or asylum seekers, short-term labour migrants, highly qualified labour migrants or students (BAMF 2008).
Three kinds of official databases exist. Data on population stocks are collected by the local authorities, who adjust the census data of 1987 by adding or subtracting national and international migration flows and natural population events (births and deaths) to or from the estimated central population update (Bevölkerungsfortschreibung). The population update contains basic data on the demographic development of the German and foreign population, including statistics on population differentiated according to German and foreign citizenship. It is known that the figure of 7.3 million foreigners listed in the population update is higher than the actual figure. Another key shortcoming is the important concept of "foreigners", whereby German citizens like the large groups of naturalised migrants and ethnic German repatriates are not identified as migrants. Moreover, the new legislation on citizenship including elements of "ius soli" for children of foreigners will result in a significant decline in the number of "foreigners" over the next few years, which means that the population update is less and less suitable as a basis for depicting the migrant population.

The BAMF is responsible for the central register of foreigners (AZR). The data of the foreign population is collected by the local authorities responsible for alien registration and includes the personal details used for administrative purposes like name, gender, date of birth, date of immigration, country and place of birth, and citizenship. The register has several advantages, the most important of which is the "legal status" parameter, which is useful for the differentiation of temporary or permanent residents. Additionally, unlike the population register, this register permits identification of first and second generations as well as duration of stay. The quality of the data was enhanced in 2004 by the clearing up, and this resulted in a decrease in the number of foreigners from 7.3 million to 6.7 million (Opfermann et al. 2006).

A third database for migrant information is the microcensus, a mandatory survey of a one per cent sample of the German population conducted each year. The last census in Germany was conducted in 1987, and the next one will be in 2011. Until then, the sole official source of information on migrant population households is the microcensus (Statistisches Bundesamt 2008). The microcensus is also part of the European labour force survey. Before 2005, migrants were identified by foreign citizenship. To ensure that the figures also covered German migrants, the new concept of "migratory background" was incorporated in the 2005 microcensus. This concept covers all foreigners, naturalised foreigners, ethnic German repatriates and immigrants as well as their descendants (Statistisches Bundesamt 2008). Under the new definitions, around 15 million persons have a migration background.
There is an even broader range of options in the field of migrant integration. All the data sets including those providing information on nationality or migration background can be used for social and economic research on integration issues.

At national, regional state and municipal level, various models are in place for the classification of aggregate data collected for administrative purposes for the implementation of an integration monitoring system. Various data sets exist on structural integration, such as the education statistics, the employment statistics of the German Federal Employment Agency (BA), the data sets of the Institute for Employment Research (IAB), the statistics on social security recipients and the crime statistics, which also cover German repatriates in some states (Länder).

The microcensus is the largest official microdata file and one of the most important resources for integration research. It primarily permits research on the structural integration of migrants as well as – since 2005 – analysis of persons with a migration background. Researchers interested in aspects other than structural integration are likely to conduct their own surveys or to turn to unofficial data sources.

A second important data set for analysis of different aspects of integration is the Socioeconomic Panel (SOEP), a representative longitudinal study of private households conducted annually since 1984 by the German Institute for Economic Research (DIW) (Wagner et al. 2007). Migrants are generally included in the former labour migrant household sample, supplemented by samples of ethnic German repatriates and migrants from a variety of countries.

Many institutes and universities are active in the field of integration research; an overview on ongoing projects and publications can be found in the GESIS-IZ research database (2008). Some institutes have conducted surveys on the largest migrant groups (e.g. the integration survey of the Federal Institute for Population Research (BiB, Haug and Diehl 2005), the foreigner survey of the DJI (Weidacher 2000) or the representative survey of selected migrant groups (RAM) (Babka von Gostomski 2008). One current trend is to include a sample of Turkish migrants parallel to an ongoing study, like the study on participation (Halm and Sauer 2007). Another method of research into integration is to survey migrants in studies on pupils conducted in schools, like the PISA-study or the IKG and KfN studies.
Future Developments in Germany

Data Collection and Data Provision

The most important actors in the field of official statistics are the Federal Statistical Office, the statistical offices of the states and the BAMF. As in other demographic fields, the most important future developments for migration research are, firstly, the 2011 census and, secondly, the plans currently under discussion to compile a central population register. Both developments will result in revision of the official number of foreigners.

An important role in data collection play private research institutes, being involved in almost every large empirical research project (Mohler and Rosenbladt 2008). There is a trend towards including a migrant sample in new survey projects, like the National Educational Panel Study (NEPS) at the university of Bamberg, or the panel analysis of intimate relationships and family dynamics (PAIRFAM) at the university of Bremen or the generations and gender survey (GGS) of the BiB. A methodology which is most appropriate for research on migration process rests upon combined surveys in the country of origin and destination, for example the Mexican-US migration and the Polish migration to Germany (Massey et al. 2008) or the SOEP study on emigrants (Schupp et al. 2008).

In view of the growing migrant population with German citizenship, another trend is the use of name-based sampling method (onomastics, see Humpert and Schneiderheinze 2000). Another sampling method is based on the place of birth of migrants (topomastics), see for example studies on ethnic German repatriates in the local population register (Haug and Sauer 2007; Salentin 2007).

Data Usage and Data Access

For data protection reasons, there is no free access to local population registers and to the AZR. The register data is used for administrative purposes, and options for statistical analysis are strictly regulated by law. Aggregate data is available at each office responsible for the register. The local statistical offices give access to population register data for research purposes which is especially the case for registration-based sampling in survey research.

Aggregate migration and population data can be found in the annual publications of the Federal Statistical Office. Other migration data is published by the BAMF or the Federal Employment Agency (BA). Access to the data files can be requested at the local statistical offices or the statistical offices of the states. The Federal Statistical Office publishes data on migration at local level (Statistik local). Population data sets on district level are available at the Federal Statistical Office (Statistik regional) and the Federal Office for Building and
Regional Planning (BBR, INKAR PRO, current regional monitoring including forecast). There are many options for researchers who want to access aggregate data on structural integration. Statistics are published by the relevant authorities and the Federal Statistical Office. For scientific purposes, researchers can access the 2005 microcensus file at the research data centres (FDZ) of the statistical offices. Access is also possible at the German microdata lab at the Centre for Survey Research and Methodology (ZUMA) in Mannheim. Access to the microcensus is exemplary, and this option is widely used by researchers.

The most important actors in the research field are the following: the European Forum on Migration Studies (efms), the Federal Institute for Population Research (BiB), the German Institute for Economic Research (DIW), the German Youth Institute (DJI), the Hamburg Institute of International Economics (HWWI), the Institute for Employment Research (IAB), the Institute for Interdisciplinary Research on Conflict and Violence (IKG), the Institute for Migration Research and Intercultural Studies (IMIS), the Institute for the Study of Labour (IZA), the Mannheim Centre of European Social Research (MZES), the Social Science Research Centre in Berlin (WZB) and the BAMF research group. Moreover, several universities conduct empirical projects on migration or integration.

Research institutes collecting data sets have their own release strategy. Data surveyed for policy-related reasons is generally not available to researchers. The exceptions are the SOEP data, which can be directly ordered from the DIW, or numerous files of the FDZ at the Institute for Employment Research (IAB), which are prepared for scientific purposes. The central data archive (ZA) of the GESIS lists other migrant surveys which can be ordered. The most important of these is the repeated Marplan survey on foreigners in Germany (latest data set: 2002). Since 1991, the German General Social Survey (ALLBUS) conducted every second year by ZUMA has also included foreigners if they speak German. Even if the migrants subgroup is too small to use it in integration research, the ALLBUS is an important data set for the analysis of attitudes of population on migrants.

**Future Developments: European and International Challenges**

International migration researchers are confronted with differences regarding concepts, data collection methods and databases. As has long been criticised (Lederer 2004; Sachverständigenrat 2004), the German statistics system does not conform with UN recommendations on the concepts of international migration statistics (long-term and short-term migrants) and international statistics for population stocks (foreign-born persons) (UN
1998; 2007). Standard EU statistics regulations also incorporate the concepts of long-term migrant and foreign-born person (EU 2007). The aim of several follow-up international projects (COMPSTAT, THESIM, PROMINSTAT) was to collect “meta-data” for the comparison of migration and integration of migrants in the EU (Poulain et al. 2006; Fassmann et al. 2008). As a result, it is concluded that attempts to harmonise migration statistics in the EU have not been very successful to date, making comparisons difficult. The variations reflect differences in the definition of what constitutes a migrant and differences in the data sources - surveys or administrative records (Thierry 2008).

In the field of international comparative integration research, the SOEP is one of the most suitable resource, including the German section of the household panel of the EU (ECHP). Other data sets, which include some information on migrants are EU-SILC, EU Labour Force Survey, the OECD/SOPEMI dataset, the Eurobarometer and the European Social Survey (see Kahanec and Zimmermann 2008). International projects focusing on migrants like EFFNATIS (efms) or PIONEUR (ZUMA) incorporate data on migrants in a comparative perspective. The research programme of the IMISCOE network (International Migration, Integration and Social Cohesion), comprised of 23 European research institutes, includes a number comparative projects, for example the TIES project on the Integration of the European Second Generation.

Conclusions and Recommendations

Overall, migration and integration research infrastructure in Germany is comprehensive. Nevertheless, there are several problems that should be mentioned.

Relating to official migration and integration statistics:

1. When using data on migration and migrant population, researchers have to deal with undercoverage of certain groups of international or internal migrants due to non-registered migration. This problem is inherently; studies on irregular migrants can complete knowledge on migration. A more serious problem is the substantial difference in the number of foreigners listed in the population update and the microcensus on the one hand and the AZR on the other. The lawful adjustment of the local resident registration offices and the central register of foreigners will improve the data quality in the future. The problem may also be tackled by the census in 2011 and the population update revision. The census will be a new basis for the extrapolation of the migrant population in the microcensus and also for the weighting of
migrants in surveys. Researchers and private research institutes should take care of results for present and future projects on migrants.

2. The creation of a **central population register** would be another way of improving migration statistics. It should be ensured that researchers will have access to a central register for the analysis and sampling of foreigners and foreign-born persons – as is currently the case with local population registers. A central population register is cost and time effective and would be a great improvement for survey research; due to the decentralised register a nationwide population sampling procedure, for example the ALLBUS, takes five or six months (Babka von Gostomski and Pupeter 2008).

3. The new **concept of migration background** has become widely accepted. Partial replacement of the "foreigner" concept by the concept of "migration background" in official statistics is a necessary improvement which reflects the reality of the population. In line with the recommendations of the working group on coordinated household and population statistics in the Union of Communal Statistical Information Systems (KOSIS), statisticians analyse local population registers for persons with migration background (Härle 2004; Bömermann et al. 2008). Furthermore there are approaches to implement the migration background in the education statistics, and the labour market statistics. Yet there is no agreement or common use of the concept regarding several aspects, like the migration status of third-generation migrants or children with only one parent with a migration background. A clear-cut identification of ethnic German repatriates is still not possible in the microcensus 2007 (Seifert 2008). It can be observed that researchers use slightly different definitions of migration background. The differing operationalisation of migration background complicate the interpretation of results and researchers should carefully discuss the implications of their definition of migration background.

4. In general, data for **international comparative research** into migration is unsatisfactory. Unlike in other countries, the concept of ethnicity is not used in official German statistics (see Schönwälder et al. 2008 for an overview on measurement of ethnicity in empirical research). The extension of the "foreigner" and "foreign-born" concepts and the inclusion of descendants of foreign-born and ever-international migrants is suggested in the context of the 2010 census in the EU (UN 2006). The concept of migration background of the census 2011 will therefore be internationally comparable. Unfortunately, questions on migration background are not
scheduled in the census 2011 in Germany. As a consequence, the extrapolation of persons with migration background in the microcensus will be without solid foundation in the future. It is recommended to reverse this decision.

5. German migration statistics does not include information on duration of stay, so the concept of short-term/long-term migration is not practicable. The AZR permits analysis of long-term foreign migrants but excludes German citizens. A new law on population statistics was passed in 2008, stipulating that immigrants must be asked about their date of emigration. In this way, long-term emigrants staying abroad for longer than one year can be identified when they return to Germany, and circular migration can be identified. This is an improvement for German migration statistics, but a comprehensive solution for analysis of long-term migration would be preferable. A central population register would enhance the data situation.

6. The implementation of the "foreign-born" concept has also been inadequate to date. From 2009, this problem will be addressed by recording the country of birth of foreigners in the local registers. This solution is a real improvement in the population statistics, as the current practice is to record only the place of birth.

Relating to empirical migration and integration research:

7. In order to improve the validity of social research and social structure analysis, it is advisable that migrants are generally included in representative population surveys. One reason for the implicit or explicit exclusion of migrants is that especially first generation migrants frequently have no sufficient German language knowledge to participate in a survey. This implies that translations of questions should be available for interviewers, at minimum the most prevalent migrant languages Turkish and Russian.

8. The freely accessible data sets of the microcensus, the SOEP or the IAB employee sample are adequate for research into integration of the large migrant groups, and there is still untapped potential for analysis in this area. But even if population surveys like the microcensus cover the diversity of immigrant groups living in Germany, subgroup analyses soon run into case number problems, as the data protection regulations make research into smaller migrant groups or small regions difficult. To realize a case number of migrants that can be analysed it is recommended that specific large survey studies are supplemented by a migrant sample. Examples are the NIEPS, PAIRFAM or GGS. The advantage is the
possibility of the comparison of non-migrants and migrants. Most projects with this kind of research design focus on Turkish migrants; it should be taken into account that the number of immigrants from Russia and other CIS countries is even larger.

9. Some of the questions being important in migration and integration research can not adequately studied in general surveys or surveys. For example, complex questions on the migration biography only refer to migrants. A further shortcoming is the lack of data on indicators of emotional integration or the meaning of religion in large surveys. Instruments for the analysis of special constraints and resources of migrants are important for the analysis of the causal mechanism of integration processes. It is vital to develop adequate instruments for these topics in empirical migration and integration research and to implement these instruments into large studies.

10. Most research results give information on the integration of the large and important group of Turkish migrants. Over and above integration research has to face the challenges presented by new groups like the large and rapidly growing group of Polish migrants or the groups with unique profile like Jewish or Vietnamese migrants or refugees. A possibility to cope with this problem is the conduction of special surveys for these migrant groups. The sampling of small migrant groups in a nationwide context is even more complicated. Difficulties also arise with the sampling of migrants groups which are German citizens, like ethnic German repatriates and naturalized persons. Identification of the minority population using name-based sampling methods are the preferred option in all these cases. A challenge for future research is therefore the development and methodological assessment of new sampling methods based on surnames or place of birth.

11. For the analysis of causal mechanisms in the integration process, more panel studies like the SOEP are needed. Therefore, the implementation of a longitudinal migrant study in Germany is discussed.

12. Access to the most important data sets, like microcensus, SOEP or the IAB data sets, can be considered satisfactory. Even if access to data sets can be granted by request and decision in each individual case, the general access to official data or the data sets of research institutes can be seen as a serious problem for researchers.
13. International comparable data sets on migration are rare and migration research is in most cases restricted on the country of destination. The fact that **circular migration** is of growing importance, the collection of data in the country of origin of migrants or of biographies of transnational migrants is a wide field of future research.

15. As data collection is in most times delegated to **private sector agencies**, a competent partner for the complex field work is required for data quality and research (Rosenbladt 2008; Mohler and Rosenbladt 2008). The continuous improvement of migrant survey methodology in cooperation of researchers and private research institutes is a field of work.


