DESTATIS
wissen.nutzen.

METHODS – APPROACHES – DEVELOPMENTS

Information of the German Federal Statistical Office

Edition 2/2012

The catchword

Methods of federal statistics – Further development

Events

Published by: Statistisches Bundesamt, Wiesbaden

Subject-related information
on this publication:

Division B 2,
tel: +49 (0) 611 / 75 20 77
fax: +49 (0) 611 / 75 39 50
institut@destatis.de

General information
on the data supply:

Information service,
tel:+49 (0) 611 / 75 24 05
fax: +49 (0) 611 / 75 33 30
https://www.destatis.de/.../contact

Release calendar
Of the Press Office:
http://www.destatis.de/.../preview

Periodicity: (generally) twice a year

Published in December 2012

The catchword

Microcensus and labour force survey: results on the employment status as from 2011

The results of the 2011 labour force survey and microcensus regarding the number of persons in employment can be compared with previous years' results only to a limited degree, which is due to modifications in the questionnaire. The modifications aim at a (more respondent-friendly) improvement in the implementation of the labour force concept of the International Labour Organization (ILO) in the questionnaire of the microcensus and of the labour force survey integrated in it. Those modifications lead not only to a real increase in the number of persons in employment compared with the previous year but also to a increase due to methodology.

What has been modified?

In the 2011 questionnaire, the questions on determining employment according to the ILO labour force concept were introduced for the first time by a new filter question on the main status. There the respondents are asked to indicate which social status is best suited to describe their current situation, so that it is then better possible to cover the fact of whether the person performs a paid job. The main status comprises 13 categories describing the respondent's life situation:

- employee, public official, self-employed, freelancer
- person in employment on parental leave
- person in employment in partial retirement
- apprentice
- worker in a family member's establishment/holding
- persons performing compulsory military or alternative service, persons doing a voluntary social year or other voluntary service
- pupil, student
- pensioner
- unemployed
- housewife/househusband
- person permanently incapable of work
- other

Only after this question, other questions will be asked which are based on the criteria of whether a person is in employment according to the definition of the labour force concept. Consequently, people who do not indicate employment for their main status are asked about secondary jobs.

The modification is based on the assumption that the separate question about secondary jobs allows better coverage especially of smaller jobs done by people who do not consider their job as their main activity. This has two advantages: First, it allows a more exact coverage of the ILO employment status and, second, it improves the coherence with reference statistics.

What are the consequences of those modifications?

Apart from the goal of a more respondent-friendly wording of questions, the intention was to cover more completely smaller jobs[1]. Detailed empirical and methodological studies carried out in recent years[2] have shown that small jobs are better covered if a person's main status is determined before the questions about employment are asked. This is confirmed by the first results of the 2011 labour force survey and microcensus. In 2011, a marked reduction was achieved regarding the differences in results between the microcensus/labour force survey and employment accounts as well as the employment statistics of the Federal Employment Agency (*Bundesagentur für Arbeit – BA*), which are used as reference statistics in the following.[3]

Reduction of differences in results and incoherences:

The differences in results regarding the number of persons in employment between the microcensus and employment accounts were markedly reduced in 2011. The difference between the two sources was down from 1.6 million persons in employment in 2010 to 1.2 million in 2011.

Persons in employment	**2010**	**2011**
Microcensus	38 938 000	39 869 000
Employment accounts (residents; as at May 2012)	40 506 000	41 037 000
Difference	1 568 000	1 168 000

Regarding the number of employees subject to social insurance contributions, the differences in results compared with BA employment statistics were down from 1.5 million persons in 2010 to 1.2 million in 2011. Examining that group of people separately by full-time and part-time employees clearly shows that the difference decreased especially for part-time employees.

The methodological studies of the last few years (see also footnotes 1, 2, 3) suggested that, due to the sequence and wording of questions, some people with marginal jobs were covered as employees subject to full social insurance contributions doing a part-time job ("misclassification hypothesis"). The results suggest that this effect has been reduced through the re-arranged questions because the decrease in the number of part-time

1 For the background of the known differences in results see: Körner, Thomas and Katharina Puch (2009): Der Mikrozensus im Kontext anderer Arbeitsmarktstatistiken. Ergebnisunterschiede und ihre Hintergründe. In: Wirtschaft und Statistik 6/2009, pp. 528 - 552.

2 Statistisches Bundesamt (2008): Ergebnisse der Nachbefragungen im Mikrozensus/Labour Force Survey. Ein Gemeinschaftsprojekt der Statistischen Ämter des Bundes und der Länder zur Qualitätssicherung im Mikrozensus/Labour Force Survey. Statistisches Bundesamt, Wiesbaden.

Köhne-Finster, Sabine and Andreas Lingnau, 2008: Untersuchung der Datenqualität erwerbsstatistischer Angaben im Mikrozensus. Ergebnisse des Projekts „Nachbefragung im Mikrozensus/LFS". In: Wirtschaft und Statistik 12/2008, pp. 1067 - 1088.

Körner Thomas, Katharina Puch, Thomas Frank and Holger Meinken, 2011: Geringfügige Beschäftigung in Mikrozensus und Beschäftigungsstatistik. Neue Erkenntnisse zu den Hintergründen der Ergebnisunterschiede. In: Wirtschaft und Statistik 11/2011, pp. 1065 - 1085.

3 The differences in results are due to many causes (see conversion table from employment accounts to the microcensus (2010)). Apart from methodological differences, there are a number of small-scale conceptual differences.

See also Körner Thomas, Katharina Puch, 2011: Coherence of German Labour Market Statistics. The Labour Force Survey in comparison with the Employment Accounts and the Unemployment Register. Statistics and Science, Vol. 19.

employees subject to full social insurance contributions in the microcensus/labour force survey is accompanied by an increase in the number of marginally employed persons.

Employees subject to full social insurance contributions		**2010**	**2011**
Microcensus	total	29 212 000	29 579 000
	full-time [2]	23 254 000	23 725 000
	part-time [2]	5 957 000	5 853 000
BA employment statistics [1]	total	27 710 000	28 381 000
	full-time [3]	22 306 000	22 683 000
	part-time [3]	5 389 000	5 670 000
Difference	total	1 502 000	1 198 000
	full-time	948 000	1 042 000
	part-time	568 000	183 000

[1] Reference date always 30 June of any year

[2] A part-time job here comprises not more than 31 weekly working hours usually worked. 32 or more hours worked are considered as a full-time job.

[3] The distinction between full-time and part-time employment is made by the employer when reporting data to the social insurance institutions.

Particular mention should be made of the changed differences in results regarding the number of people with only low-pay and short-term jobs compared with the BA employment statistics. As mentioned above, the coverage of marginal jobs has been improved by the rearranged questions. The difference has been reduced by just under 500,000 persons. While a decrease in the number of marginally employed people is recorded in BA employment statistics in 2011 compared with 2010, the microcensus shows an increase due to methodological reasons.

People with only low-pay and short-term jobs	2010	2011
Microcensus	3 109 000	3 541 000
BA employment statistics [1]	5 261 000	5 239 000
Difference	2 152 000	1 698 000

[1] Reference date: 30 June of any year

However, these improvements regarding the coherence with reference statistics involve restrictions in terms of comparability over time: Comparing the results of the 2011 microcensus with previous years or examining them in a time series shows an increase especially for the current year that is mainly due to methodology. This should be taken into account when interpreting the results.

Assessing the year-on-year change:

Compared with other sources, the results of the 2011 microcensus and labour force survey show a larger year-on-year employment growth. Employment accounts record an increase of 1.3% in 2011 on 2010, which is a smaller relative increase than that shown by the microcensus (+2.4%).

Referring only to employees subject to social insurance contributions, the microcensus shows a smaller increase (+1.3%) than the comparable BA employment statistics (+2,5%). As explained above, this might be due to a modified, and more exact, allocation of respondents as a result of the rearranged questions. Marginally employed persons who, in the microcensus, were classified as part-time employees subject to full social insurance contributions until 2010 are now identified as marginally employed persons and shown as such. As a consequence, the 2011 microcensus shows a decreasing number of part-time employees subject to full social insurance contributions compared with the previous year, whereas BA employment statistics show a marked increase. However, the year-on-year change in the number of full-time employees is similar in the two sets of statistics (BA employment statistics: +1.7%, microcensus: +2.0%).

For those who are only marginally employed, the microcensus records a particularly sharp increase by 14% in 2011 on the previous year, while BA employment statistics show a 0.4% decrease. This suggests that, due to the arrangement of questions applied until 2010, it was not possible in the microcensus to completely cover respondents as persons in employment or to correctly allocate them. The disproportionate increase in the number of marginally employed people confirms the assumption that coverage of smaller jobs is better with the rearranged questions.

Comparisons with other data sources suggest that the larger employment increase in the 2011 microcensus results is mainly due to the changes in the questionnaire.

Other results:

In a demographic breakdown, particularly large year-on-year changes are observed in the 2011 microcensus for younger and older people. There is a quite distinct increase in employment among older people who have reached, or have exceeded, statutory retirement age (+14% for the over 65 year olds). For the subgroup of marginally employed people over 65 years, an increase of as much as 29% is recorded for 2011.

When examining the development by sex, hardly any differences are observed. According to the results of the microcensus/labour force survey, the year-on-year increase in the number of men in employment in 2011 was slightly smaller than that of women. This is probably due to the fact that women, in addition to a different social status (e.g. housewife), often perform part-time or marginal jobs, which are better covered through the new questions.

For self-employed without employees (solo self-employed) and family workers, too, disproportionately large year-on-year changes are observed. These results confirm the assumption made before the modification that the rearranged questions would lead to a better coverage of (smaller) jobs for persons with different social status.

Additional information on other differences compared with the reference statistics:

Other differences, apart from smaller definitional differences, are mainly due to the different survey methods and procedures of the statistics compared with each other. In employment accounts, the total number of persons in employment is estimated using all available sources (currently about 60 sources).1 The main sources are the statistics of employees subject to social insurance contributions and of marginally employed persons of the Federal Employment Agency as well as the statistics of public service personnel.

The BA employment statistics are produced through a count of administrative data at the end of any calendar month, which are based on data reported by employers to social insurance institutions.

1 See Lüken, Stephan: Die deutsche Erwerbstätigenrechnung für Volkswirtschaftliche Gesamtrechnungen und Arbeitsmarktstatistik. In: Wirtschaft und Statistik 5/2012, pp. 385 - 405.

The microcensus is a household survey conducted annually among one percent of the population in Germany. It involves the obligation to provide information and is harmonised across the EU. Experience shows that covering smaller (secondary) jobs or jobs in the grey area between legal and illegal work is difficult in household surveys, which contributes to the differences. For such areas which are difficult to cover statistically (e.g. domestic services), supplementary estimations are also made in employment accounts.

Impact on the results for unemployed and inactives

The above effects also have an impact on the numbers of unemployed and of inactives, which show an unusually sharp year-on-year drop. The number of inactives was down by 366,000 people (–1.3%) in 2011, while the average decrease of the previous years was only just under 100,000 people.

The number of unemployed fell by 445,000 from 2010 to 2011 (–15%). This decrease is much larger than that of the registered unemployed as covered by the Federal Employment Agency (–263,000 or –8.1%). However, that difference may also be due to the different concepts of measurement.[1]

Summary

A consequence of the modified questionnaire, apart from better coherence with the reference statistics, is an excessive year-on-year increase in the number of persons in employment, which is due to methodological reasons. This also has an impact on the numbers of unemployed and of inactives. Although the impact of the methodological changes on the results for labour market subgroups cannot exactly be quantified, the modified methodology should definitely be taken into account when examining and interpreting the results over time.

Katharina Puch, tel: +49 (0) 611 / 75 41 06, e-mail: katharina.puch@destatis.de

1 Hartmann, Michael and Thomas Riede (2005): Erwerbslosigkeit nach dem Labour-Force-Konzept – Arbeitslosigkeit nach dem Sozialgesetzbuch. Gemeinsamkeitenund Unterschiede. In: Wirtschaft und Statistik 4/205, pp. 303 - 310.

And Hartmann, Michael (2009): Umfassende Arbeitsmarktstatistik: Arbeitslosigkeit und Unterbeschäftigung. Methodenbericht der Statistik der Bundesagentur für Arbeit.

Methods of federal statistics – Further development

New design for census of crafts
Administrative data evaluations replacing surveys

In July 2011, results of a census of crafts were published for the first time since 1995. For those statistics, it was the first time that only administrative data were evaluated.

Crafts, which traditionally have been very important in Germany, are a heterogeneous economic sector. To obtain reliable information on the businesses that are active in that economic sector, censuses of crafts were conducted as complete counts in 1949, 1956, 1963, 1968, 1977 and 1995. In those censuses, detailed information on the volume and structure of crafts in Germany was collected from the craft enterprises and then evaluated.

In the context of efforts made to reduce the burden both on respondents and on statistical offices caused by official statistical surveys, the above-mentioned census of crafts was conducted as an evaluation of the statistical business register for the first time for reference year 2008. This means that the results of the census of crafts were entirely obtained from administrative data. Consequently, about 577,000 craft enterprises are no longer required to comply with statistical reporting obligations.

Administrative data used from the business register

The business register is a database updated regularly and containing enterprises and establishments from nearly all economic branches with a taxable turnover from deliveries and other performances and/or with employees subject to social insurance contributions. The administrative data available in the business register are, first, turnover data from the procedure applied in the fiscal authorities regarding the enterprises' advance turnover tax returns. Second, there are data of the Federal Employment Agency on employees subject to social insurance contributions, based on reports to social insurance funds. The information on whether an enterprise is a craft enterprise is entered into the business register from data of the Chambers of Crafts.

When evaluating the business register for the census of crafts, all craft enterprises are included which had a taxable turnover from deliveries and other performances in the reference year and/or had employees subject to social insurance contributions on 31 December of the reference year. The evaluation concept applied aims at representing the total number of craft enterprises in a specific reference year and, consequently, is mainly in line with the concept of the censuses of crafts which in the past were based on surveys. The evaluations are based on the administrative data for the reference year which were last processed in the statistical business register. It should be noted here that, due to register maintenance work required for other purposes of the business register, part of the evaluation variables (for example, the Official Municipality Code for the regional breakdown) may be more recent than the reference year of the administrative data evaluated.

For the census of crafts, information is available from the business register on the following variables: number of enterprises, employees subject to social insurance contributions and taxable turnover in the reference year. In addition, data from the Federal Employment Agency on marginally employed people are included, which are available on the basis of the Act on using administrative data[1]. The number of working proprietors is estimated. In addition, classifying variables available are the main branch of trade according to the Crafts Code, the economic

[1] Act on using administrative data for the purposes of economic statistics (*Verwaltungsdatenverwendungsgesetz – VwDVG*) of 4 November 2010 (Federal Law Gazette I, p. 1480).

branch according to the Classification of Economic Activities (2008 edition) and the Official Municipality Code.

As the data of the business register are available on an annual basis, evaluations for the census of crafts can be done every year, whereas in the past surveys it could be done only at intervals of several years.

Evaluating the business register does not allow showing all the variables covered in the previous censuses of crafts. That problem can in part be solved by showing crafts separately in other sets of statistics. This provides additional information on crafts. For example, crafts are shown separately in the statistics of construction establishments and in the quarterly earnings survey. It is planned for the medium term to represent crafts separately in other statistics, too.

Census of crafts and other economic statistics

The census of crafts results published previously differ from the statistics of other economic sectors in terms of how the enterprises' economic activity is represented. In the past, the results of the census of crafts were not broken down according to the Classification of Economic Activities, 2008 edition (WZ 2008) – as is usual for other business statistics – but by branch of trade in crafts with compulsory or non-compulsory approval according to Annexes A and B1 of the Crafts Code. To compare the census of crafts results with other economic statistics, results are currently being processed by selected economic branch of the WZ 2008.

When new businesses are included in the business register, information on the relevant economic branch is taken from the above administrative data. However, the quality of coding by economic branch in external sources does not always meet the standards of official statistics. Generally, the economic branch data are checked only when an enterprise reports data for an official statistical survey. As especially among craft enterprises there are many small and medium-sized enterprises which do not often report data in surveys, the share of unchecked economic branch information is particularly large for the group of responding businesses in crafts statistics. To take account of that, it had first to be checked which economic branches were eligible for publication from quality aspects. Those checks are finished now.

The results by selected economic branches of WZ 2008 for reference year 2009 are expected to be published in September 2012. The results of the census of crafts by branches of trade will be released earlier, that is, by July 2012.

Conclusion

With the 2008 census of crafts results, data on the volume and structure of crafts in Germany have become available from official statistics for the first time since 1995. This demonstrates that, basically, it is possible to obtain even detailed results from administrative data and to reduce the burden on respondents, in particular on small and medium-sized enterprises. Also, evaluations can now be done every year, which is a considerable improvement compared with previous censuses of crafts.

Together with the quarterly reports on crafts, which provide rates of change and index numbers on turnover and persons employed in crafts and whose results are entirely obtained from administrative data, too, a statistical system is available for crafts statistics that is used to determine structural and short-term developments of an important part of the economy without putting additional burdens on craft enterprises.

Literature

Feuerhake, Jörg: *Handwerkszählung 2008* in *Wirtschaft und Statistik* 1/2012, p. 51 ff.

More detailed results of the census of crafts are contained in the GENESIS-Online database and additionally in Fachserie 4 "Produzierendes Gewerbe", Reihe 7.2 "Unternehmen, tätige Personen und Umsatz im Handwerk"*, which is available on the website of the Federal Statistical Office*

(www.destatis.de). Results in a more detailed regional breakdown by Chamber of Crafts districts and possibly by administrative districts are published by the statistical offices of the Länder.

Jörg Feuerhake, tel: +49 (0) 611 / 75 41 16, e-mail: joerg.feuerhake@destatis.de

Introduction of a rolling sample for the quarterly earnings survey

Following the entry into force of the Earnings Statistics Act, the quarterly earnings survey (QES), which is the successor to the continuous earnings survey, was conducted for the first time in 2007, using a stratified, representative random sample. The survey enables economic analysis because the data reflect short-term trends in average earnings in nominal and real terms. It also provides the data basis for a number of short-term economic and structural statistics as well as for national accounts. In the autumn of 2006, 40,500 local units were selected to provide information for these statistics for the first time. After the statistics on food and accommodation service activities were reorganised in early 2010 to use a rolling sample (cf. Methods Approaches Developments, 2/2010), this concept has now been adapted to and introduced for the quarterly earnings survey.

Reasons for using a rolling sample instead of replacing the complete sample

Since 2007, the QES has come to be used in many important areas such as national accounts at federal and Land level, the labour cost index, EU short-term economic statistics, and by a variety of research institutes. Besides, the QES results have aroused increasing interest among the media and the political community.

Data quality

Long-term data comparability is one of the indicators of good data quality. Comparability is ensured if there are as few breaks in the time series as possible. The main reason for breaks is that the sample must be refilled to the statutory maximum size when local units have ceased to exist. In 2011, only about 34,000 of the initial 40,500 local units were still included in the survey. In the case of the predecessor survey, the continuous earnings survey, it was possible to solve the problem by processing both the old and the new sample. As a result, link factors could be computed to smooth the breaks. However, this is no longer permissible under the new Earnings Statistics Act, so that an alternative method of reducing breaks had to be found. The rolling sample is such a method because it permits refilling the sample every year.

The representativity and accuracy of results are other important quality indicators. To give the most accurate picture possible of the constantly changing reality – local units are closed down while others are set up – newly established local units must have a reliable chance of being selected. This could previously not be ensured because samples of business start-ups did not exist. Now that the rolling sample has been introduced, new sampling plans are set up each year which take business start-ups into account. As a result, undercoverage and structural distortions can be avoided.

The representativity and accuracy of the statistical results are significantly improved.

Burden on the respondents

As part of the rolling sample, only a certain proportion of local units are released from their reporting obligations each year while others are newly included. This approach is based on a pre-defined rotation schedule which also specifies conditions on the expected maximum

length of time that a local unit remains in the sample. Accordingly, the respondent local units have some certainty about how long they will be obliged to provide information.

Workload for the specialised units

The rolling sample allows the workload to be spread more evenly across the specialised units in the statistical offices of the Länder, so that the increasing shortage of resources can be taken into consideration. Work that was previously performed once a year becomes routine work and the staff doing it benefit from learning effects.

Principles of rotation

With due regard to the conflicting interests of the respondents, the statistical offices of the Länder and the data users, a third of the local units included in the sample since 2007 will be replaced in each stratum in the transition period from 2012 to 2014. From 2015 onwards, a rotation schedule will be used according to which a sixth of the local units will be replaced every year.

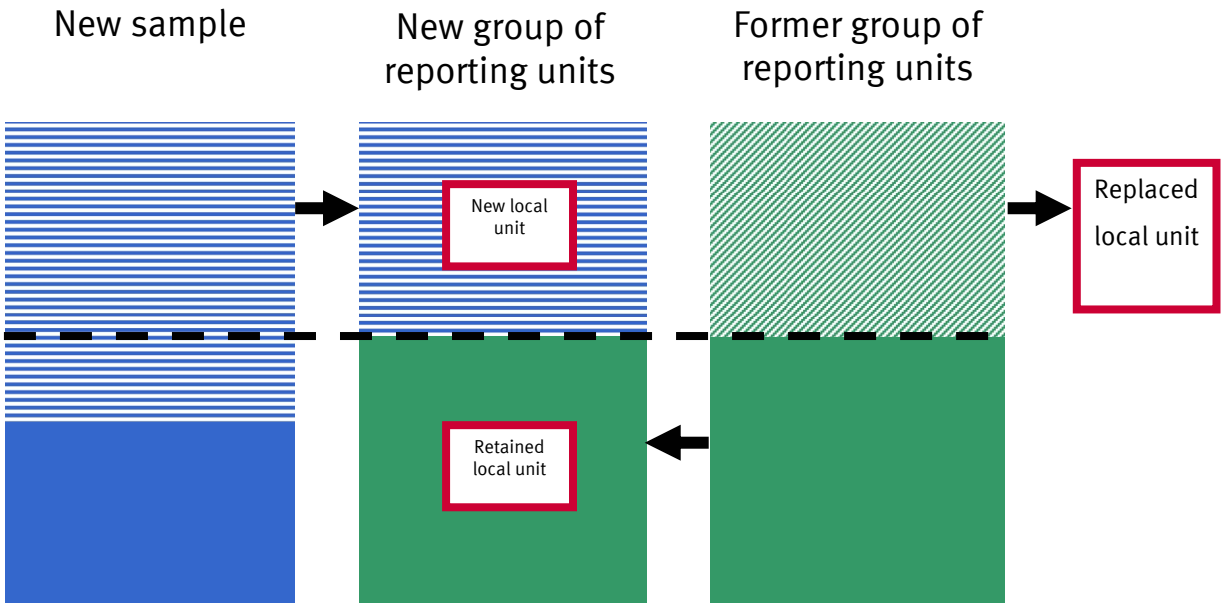
Local units that have left the sample should not be required to provide QES data for a minimum of three years or, if possible, even for six years (minimum and target pause length). Local units that have been newly included are intended to report QES data for a minimum of three years and a maximum of six years (minimum and target participation time).

The QES sampling plan, which specifies the size of the strata and the estimation procedure, is adjusted to the current business register every year. In the process, the characteristics relevant to stratification (Land, branch of economic activity, size class) are taken into account for the local units reporting to the survey, using the QES results for the second reference quarter of the year concerned.

As in earlier years, local units with more than 1,000 employees form fully enumerated strata because only the statistical information regularly provided by these local units ensures that reliable data can be provided on the levels of, and changes in, earnings.

Composition of the new group of reporting units

In accordance with clearly defined rules of decision, the new QES reporting group is set up every year partly from the existing group of reporting units and partly from the newly drawn sample. The business register is used as the basis for drawing the sample, which is done centrally at the Federal Statistical Office. The sample contains as many local units as possible of those that have never, or not for a long time, been surveyed by the QES in order to create sufficient potential for replacing the currently reporting local units. As part of the subsequent rotation, it is decided how many local units in each stratum can be released from their reporting obligations (replaced local units), how many local units will have to continue reporting to the QES (retained local units) and how many local units are newly selected to report (new local units). The local units to be replaced are randomly selected from the group of currently reporting local units, as are the new local units from the newly drawn sample.



The new group of reporting units is composed of new and retained local units.

Conclusion

Now that the rolling sample has been introduced, the selection of the reporting group is based for the first time on a concept that offers planning dependability and reliability to all parties involved. Data quality is increased, and the higher demands on the QES are taken into consideration. The transition to routine work enables the statistical offices of the Länder to design resource-friendly and more efficient work processes. Last but not least, the yearly feedback of information from the survey contributes to improving the quality of the business register. And this again is a benefit to the QES as a register-based survey.

It remains to be seen whether all these benefits can be achieved to the extent expected. The new methodology will be improved and optimised on the basis of the experience gained.

Sandra Klemt, tel: +49 (0) 611 / 75 27 06, e-mail: sandra.klemt@destatis.de

"Surveys for special purposes" on doctoral students in Germany in winter semester 2010/2011

Project background and purpose

Data on doctoral students in Germany are required for educational reports at the national and the international level. However, due to the provisions in the enrolment and doctoral degree regulations at higher education institutions, systematic undercoverage of the data on doctoral students is observed in the statistics of students. Doctoral students who are not enrolled at a higher education institution are not covered there. This is why the Federal Statistical Office had been asked by the Federal Ministry of Education and Research to conduct a survey on doctoral students in Germany. The project purpose was to develop a concept allowing to provide reliable information on doctoral students in Germany. National data were required, for example, for the Federal Report on Promoting Young Scientists. Another goal was to obtain information on

doctoral students that would include the variables required for international reporting[1]. The legal basis for the surveys was Article 7 Paragraph 1 of the Federal Statistics Law[2]. As is usual for "surveys for special purposes", the survey was conducted in co-operation with some Land statistical offices.

Survey concept

In the survey concept, a two-stage procedure was defined. In the first stage, 20,000 professors at German higher education institutions with a right to award doctorates were asked in a voluntary survey to indicate the number of doctoral students supervised at the beginning of winter semester 2010/2011 as well as their type of doctoral studies and sex. Doctoral students covered were those guided by the professor as a (primary) supervisor. Another item covered was the professor's willingness to pass on the survey documents to their doctoral students at the second stage. This was crucial for the further process. The professors were either contacted directly or got the survey documents through an internal distribution procedure applied at higher education institutions. In a second stage, another "survey for special purposes" was conducted among 20,000 doctoral students. In the second survey stage, the professors were asked to pass on the survey documents to their doctoral students. The doctoral students were questioned on socio-demographic variables, on when they started their doctoral studies, on their employment status, and on any support programmes.

Sampling concept

A two-stage procedure was applied in the sampling concept, as was the case for the survey concept. The information obtained in the survey of professors at the first stage was needed for sampling in the survey of doctoral students (stage 2). At a first stage, 20,000 professors at German higher education institutions with the right to award doctorates were sampled. For the purpose, a cluster sample was formed, the clusters being the higher education institutions. For the participating higher education institutions, a Bernoulli sampling procedure with an inclusion probability proportionate to "doctorates per chair and year". To avoid large design weights, care was taken to have an inclusion probability of at least 1/10. Large higher education institutions were sampled with a probability of 1 to avoid excessive variations in the number of professors questioned. A total of about 19,500 professors were actually questioned at the first survey stage. In the sampling concept for the survey of professors, it was defined that not all higher education institutions in Germany with a right to award doctorates had to be included in the survey.

At a second stage, 20,000 doctoral students of those professors were sampled who at the first stage had agreed to pass on the survey documents to their doctoral students. So the results of the first stage of the survey among professors were the sampling basis for the survey among doctoral students. The professors who had agreed to participate in the second stage were stratified by Land, higher education institution, as well as teaching and research area or subject of the chair. The sample size was distributed proportionately to the strata; when selecting the professors, care was taken that small groups, such as doctoral students with an external type of doctorate, were sufficiently represented in the sample. A total of 19,998 doctoral students were sampled at random from the sampling frame with the condition that all doctoral students of a specific professor are selected.

Extrapolation method

1 Due to Regulation (EC) No 452/2008 concerning the production and development of statistics on education and lifelong learning, there is an obligation in the European Union (EU) to provide statistics on the systems of general and vocational education. In Commission Regulation (EU) No 88/2011 of February 2011 implementing Regulation (EC) No 452/2008, it was specified that the Member States have to provide data on doctoral students in September 2012. In a Commission Decision of December 2010, however, derogations were granted regarding the implementation of the Regulation (OJ L 335 of 18 December 2010, p. 66). In a derogation, Germany was allowed to supply data on doctoral students a year later, that is, in 2013.

2 Article 7 Paragraph 1 of the Federal Statistics Law: "In order to meet short-term data requirements for the purpose of preparing and motivating impending decisions of highest federal authorities, federal statistics may be performed without an obligation to provide information if a highest federal authority demands such federal statistics." In a "survey for special purposes" according to that Article, not more than 20,000 respondents may be covered. Participation in the survey is voluntary.

In line with the other concepts, the results of the two “surveys for special purposes” were extrapolated in two stages, too. The general goal of extrapolation was to take the sample parameters (total value, mean value, proportion, variance) as a basis for drawing conclusions for the parameters of the population, using suitable estimators. An unbiased estimate of the unknown total value of a relevant variable *Y* is obtained through a generalised regression estimator. The regression estimator is a linear estimator and one of its characteristics is that the benchmarks are met during extrapolation from the sample. This methodological approach was applied for extrapolation in the “surveys for special purposes”. Furthermore, the regression approach has the advantage that the approximate formula of the variance of the estimator has a closed representation. In turn, the variance can be estimated from the sample. Consequently, the estimation error can be quantified for any target group after every extrapolation. The relative standard error can be indicated as a measure of the random error. The standard error is calculated by means of the CLAN SAS programme package of Statistics Sweden.

In the first stage of the survey, about 9,400 (48%) professors had responded and were available for extrapolation. About 6,800 of them were ready to pass on the questionnaires to their doctoral students. The basis for the first extrapolation stage was benchmarks from the statistics on public service personnel. Referring to 2010, those figures covered all 33,880 professors at German higher education institutions with a right to award doctorates. From them, benchmarks for the professors were derived regarding the study fields or subject groups, main or secondary occupation, sex, distribution by former territory of the Federal Republic as well as new Länder and Berlin-East, and the common distributions between subject groups and the professor’s activity.

In the second stage of the survey, about 8,700 (43%) doctoral students supervised by 1,800 professors had responded to the survey. The basis for the second extrapolation stage was benchmarks of the enrolled doctoral students, taken from higher education statistics for winter semester 2010/2011. To ensure that both extrapolations refer to the same number of doctoral students, the number of doctoral students calculated in the first stage was taken as an additional benchmark in the second extrapolation. A strong correlation was found between the number of a professor’s doctoral students and the probability that at least one of them participated in the survey. During extrapolation, the correlation was taken into account by weighting.

The surveys on doctoral students in Germany thus provide results both at the level of professors and at the level of doctoral students.

Results

The surveys on doctoral students in Germany provide comprehensive results on doctoral students in Germany in winter semester 2010/2011. They show that about 200,400 doctoral students were supervised by professors at German higher education institutions in winter semester 2010/2011. In addition to the 104,000 doctoral students enrolled at German higher education institutions, nearly the same number of people did doctoral studies without being enrolled at a higher education institution. A total of 25,600 people finished their doctoral studies in Germany in 2010.

Most doctoral students worked on a doctorate in the area of mathematics, natural sciences. About 58,400 people (29%) obtained their doctorate in that area in winter semester 2010/2011. Engineering was the second largest area of study (44,500 doctoral students or 22%), followed by languages and arts (38,700 doctoral students or 19%). About 35,900 people (18%) completed their doctorate in legal, economic and social sciences. In the human medicine area of study, there were just 8,300 doctoral students (4%). The main reason is that doctoral studies in human medicine are relatively short.

About 64,900 doctoral students (32%) started their doctoral studies in 2010. About 47,100 doctoral students (24%) started in 2009 and 37,500 (19%) in 2008. In the years from 2005 to 2007, the number of doctoral students taking up their doctoral studies was 44,100 (22%), while 6,800 doctoral students (3%) had started before 2005.

More detailed results and the comprehensive description of the concepts are contained in the special report on *Promovierende in Deutschland*, which is available at www.destatis.de.

Miriam Wolters, tel: +49 (0) 611 / 75 41 13, e-mail: miriam.wolters@destatis.de

Sven Schmiedel, tel: +49 (0) 611 / 75 22 75, e-mail: sven.schmiedel@destatis.de

"Survey for special purposes" on income and expenditure of non-government schools and schools for nurses, midwives etc.

Project background and purpose

In the last few years, the quantitative importance of non-government schools has continuously increased in the German education system. In school year 2010/2011, 8.4% of the pupils in Germany attended a private school. Despite their importance, the only financial data available for non-government schools came from a survey of 1997 for reference year 1995. Data availability was similar for schools for nurses, midwives etc. Therefore, income and expenditure data were updated every year, based on the survey results for reference year 1995. However, robust information on expenditure and income of the health sector institutions is urgently required for official statistics. The Federal Ministry of Education and Research commissioned the Federal Statistical Office (Destatis) to conduct a survey among non-government schools and schools for nurses, midwives etc. Financial data on non-government schools are urgently required to assess the degree to which the 10% target has been achieved. The results are also included in the national budget for education, science and research and in the national Education Finance Report. Due to Regulation (EC) No 452/2008 of the European Parliament and of the Council of 23 April 2008 concerning the production and development of statistics on education and lifelong learning, the Federal Republic of Germany is now obliged to annually report data on government and non-government educational institutions. The purpose of the survey was to create a data basis that can be used to develop a solid procedure for annual updating which meets the quality requirements of international organisations. The legal basis for the survey was Article 7 Paragraph 1 of the Federal Statistics Law. As is usual in such surveys for special purposes based on that Article, the survey was conducted in co-operation with some Land statistical offices.

Survey concept

According to the survey concept, all alternative schools as well as all government and non-government schools for nurses, midwives etc. in Germany should be covered by the survey. Items to be covered – in addition to general information on the institutions such as government/non-government responsibility, teachers and number of pupils – were in particular data on income and expenditure in reference year 2009. For the purpose, the current lists of schools and their addresses were requested from official school statistics of the relevant statistical offices of the Länder. A problem that occurred already in processing the address lists was the fact that the term "school" is not used in the same way everywhere. In official statistics, institutions are usually defined on the basis of the number of qualifications offered. This means that an institution offering at the same time intermediate school, secondary general school and grammar school certificates is in part counted as three schools in official school statistics. However, definitions sometimes differ between the Länder; for vocational schools and in particular for schools for nurses, midwives etc. the relevant criterion is not always the number of qualifications offered but often the legal or local/geographical unit. When covering financial data in general and, more specifically, investments, it appears to be more suitable to apply a local definition of the term "school" because it is difficult to classify such data by type of qualification. When preparing the survey, a decision was taken to use the local-unit definition of the term "school" and to process

the address lists according to that definition. A consequence of the different ways of counting institutions is that this figure did not seem very reliable as a benchmark for extrapolation and that the numbers of pupils grouped by Land and school type was envisaged as a valid alternative.

Extrapolation method

To extrapolate the raw data of the survey to the population of private schools and schools for nurses, midwives etc. and thus to compensate for non-response, further steps had to be taken. As data on income and expenditure were covered for reference year or financial year 2009, while data on the numbers of teachers and pupils were available by school year (2008/2009 and 2009/2010), adjustments had to be made. Therefore, the data on teachers and pupils were converted to financial year 2009 in accordance with the methodology of international education statistics. In another step, the data on teachers that had been converted to reference years 2009 were transformed into full-time equivalents by means of data on the teachers' scope of employment. In some Länder, there are teachers at private schools who are financed by the public Land budget. To obtain robust results in such cases and to represent the entire resources available to independent schools, the volume of free teaching staff covered was weighted by means of empirical average cost rates and added to total staff costs. Also, data reports containing numbers of pupils for several school types were separated through calculation because the level of breakdown of the benchmarks applied in actual extrapolation is individual school types and Länder. Another intention was to separately extrapolate schools of general education, vocational schools, and schools for nurses, midwives etc., which again required dividing the reports containing numbers of pupils by several school types. For the purpose, all data within a report were divided proportionately by the number of pupils. The only exception was pupils at special needs schools. They got a larger weight when the reports were divided through calculation because it is assumed that this school type requires more resources.

The results of the survey were extrapolated using the most suitable approach. That approach includes as absolute figures the numbers of pupils and the number of teachers in full-time equivalents by Land and school type for non-government schools of general education and vocational schools. The information on those benchmarks is obtained from official school statistics for the relevant two school years and were converted applying the same procedure. For the calculation of teachers in full-time equivalents, the weighting used for the survey data was applied again. The schools for nurses, midwives etc. were extrapolated only on the basis of numbers of pupils because for those schools no reliable all-German data on teachers are available. The goal of extrapolation was to draw conclusions from the survey parameters for the parameters of the population, using suitable estimators. The regression estimator used is a linear estimator and one of its characteristics is that the benchmarks are met during extrapolation from the sample. That approach was applied for the extrapolation of the number of teachers and the total numbers of pupils. Also, the relative standard error was calculated to indicate the random error.

Schools of general education with a total number of 695,000 pupils were contacted. Data reports covering 317,000 pupils were received, which were then available for extrapolation. The basis for that stage of extrapolation was benchmarks determined from official school statistics. They included the number of pupils at schools of general education and a total of 69,000 teachers (headcount), who were weighted and converted into 51,000 full-time equivalent teachers. The benchmarks were grouped by eleven school types and 16 Länder. As regards non-government vocational schools, schools with a total of 237,000 pupils were contacted. From those schools, data reports covering a total of 61,000 pupils were received. The basis for the second extrapolation stage was again official school statistics data, which were converted to 2009. They contained the number of pupils at non-government vocational schools and a total 30,000 teachers (headcount), who were weighted and converted into 15,000 full-time equivalent teachers. The large difference between the two values is due to a larger relative share of teachers who work part time or on an hourly basis. The benchmarks were grouped by eight school types and 16 Länder. Regarding schools for nurses, midwives etc., schools with a total of some

140,000 pupils were contacted. Data reports covering a total of 41,000 pupils were received from them. The data were extrapolated using numbers of pupils from official school statistics – converted to 2009 – and some data from additional sources. Also, five empirical size classes regarding the number of pupils were stored in the survey data and, analogously, in the benchmarks. This provided additional information for the extrapolation, which makes the data material more homogeneous. The benchmarks were grouped by size class and Land. The three extrapolations of this "survey for special purposes" were carried out through SAS implementation, using the CLAN macro package programmed by Statistics Sweden.

Updating procedure

Data to be supplied for international education statistics, the education budget and the Education Finance Report have to be compiled every year. As the "survey for special purposes" cannot be conducted every year among non-government schools and schools for nurses, midwives etc., an updating procedure has been developed for the years from 2009 on the basis of the data material available. What is used for the procedure is the numbers of teachers and pupils of the current reference year as a quantitative framework as well as updated expenditure rates that have been determined in the survey of income and expenditure of non-government schools for 2009. It consists of three steps and is briefly outlined below.

Step 1:

First, four ratios are determined for reference year 2009 on the basis of the survey:

- PAL_i = staff expenditure per full-time teacher per school type *i*
- SAS_i = material expenditure per pupil per school type *i*
- IAS_i = investments per pupil per school type *i*
- ISS_i = investments funded from special programmes per pupil per school type *i*

Step 2:

To estimate changes that have occurred in the meantime, auxiliary variables are used, i.e. results of national accounts on the price development of the gross domestic product and of public finance statistics on special programme funds used:

- ΔDEF_j = rate of change of the deflator of the gross domestic product in year *j*
- ΔISS_j = rate of change of special programme funds used in year *j*

The ratios of survey year 2009 are then multiplied by the auxiliary variables for survey year *j* to obtain ratios for reference year *j*:

- $PAL_{ij} = PAL_i \times DEF_j$
- $SAS_{ij} = SAS_i \times DEF_j$
- $IAS_{ij} = IAS_i \times DEF_j$
- $ISS_{ij} = ISS_i \times ISS_j$

Step 3:

In the next step, staff expenditure per full-time teacher is multiplied by the number of teachers determined in school statistics for the reference year (converted to full-time employees). This takes account of the volume of expenditure and of changes in staffing. Material costs per pupil and investments per pupil, including the share of investments from special political programmes, are multiplied by the numbers of pupils determined for school type *i* and reference year *j* to

obtain the volume of expenditure in reference year *j*. In this way, the expenditure of non-government schools are updated, taking account of the development of the numbers of teachers and pupils. By adding up the individual items, total expenditure for non-government schools and schools for nurses, midwives etc. in reference year *j* is obtained.

- $GA_{ij} = (PAL_{ij} \times L_{ij}) + (SAS_{ij} \times S_{ij}) + (IAS_{ij} \times S_{ij}) + (ISS_{ij} \times S_{ij})$

With:

- S_{ij} = number of pupils at school type *i* in school year *j*
- L_{ij} = teachers at school type *i* in school year *j*
- GA_{ij} = total expenditure at school type *i* in school year *j*

Results

The "survey for special purposes" conducted among non-government schools and schools for nurses, midwives etc. provides information at different levels of detail on the availability and use of resources. It shows that in 2009 a total Euro 6.1 billion was spent for non-government schools of general education and vocational schools and Euro 0.9 billion for schools for nurses, midwives etc. The largest part of the expenditure was made on staff remuneration. An average of approximately Euro 7,000 was spent per pupil at non-government schools of general education in 2009. Expenditure differed considerably by school type. Average expenditure at primary schools amounted to Euro 5,900 in 2009, at intermediate schools Euro 4,900, at special needs schools Euro 15,800 and at grammar schools Euro 5,900. Expenditure per pupil at vocational schools was Euro 5,400 and at schools for nurses, midwives etc. Euro 6,500. Differences in financial resources are due to differences regarding the school structure, care offered, pupil-teacher ratios, and remuneration structure as well as to differences in the number of lessons.

More detailed results on non-government schools and a description of the survey concept are contained in the special report *Finanzen der Schulen in freier Trägerschaft und der Schulen des Gesundheitswesens*. It is available for download at www.destatis.de.

Carsten Haider, tel: +49 (0) 611 / 75 42 82, e-mail: carsten.haider@destatis.de

Sven Schmiedel, tel: +49 (0) 611 / 75 22 75, e-mail: sven.schmiedel@destatis.de

Events

Workshop on "The microcensus as a data source for empirical social and economic research: introduction to working with the data" at GESIS in Mannheim from 29 to 30 November 2012

The microcensus, the largest continuous household survey in Germany, is available to the scientific community as a 70% subsample (Scientific Use File) for research projects. The list of questions focuses on socio-demographic, socio-economic and employment variables. In addition to the questions asked regularly, supplementary programmes are included at four-year intervals on various topics and ad-hoc modules are added as part of the EU labour force survey. The supplementary programme of the 2010 microcensus contains questions on the housing situation, while the 2010 ad-hoc module deals with the work-family life balance (0.1% sample).

The purpose of the event is to provide basic knowledge on how to work with the data and to give an overview of the special features and analytical possibilities of the microcensus. On the first day, speeches will be held on the concept, survey design and focal issues of the microcensus. On that basis, selected issues will be dealt with in more detail in practical exercises on the second day. The SPSS statistics software will be used. Those participating in the exercises are expected to have good knowledge in syntax-driven work.

The event will be held in close co-operation with the Federal Statistical Office and the Hessian Land Statistical Office.

Venue: GESIS, Mannheim

Participation fee: Euro 80 (students: Euro 60)

Maximum number of participants: 15

Those interested are asked to register with the conference office not later than four weeks before the workshop starts. E-mail: **workshop-mannheim@gesis.org**; tel: +49(0)621 / 12 46 221

More information and the detailed workshop programme will soon be available at: http://www.gesis.org/

GESIS contact:

Andreas Herwig

GESIS - Leibniz-Institut für Sozialwissenschaften

German Microdata Lab (GML)

P.O. Box 12 21 55

D-68072 Mannheim

e-mail: andreas.herwig@gesis.org

Thomas Haustein, tel: +49 (0) 611 / 75 81 34, e-mail: thomas.haustein@destatis.de