Schmollers Jahrbuch 127 (2007), 655 – 668
Duncker & Humblot, Berlin

CAMPUS Files

Free Public Use Files for Teaching Purposes

By Markus Zwick

1. Introduction

Microdata now form the basis of a wide range of scientific analyses. The use of the original statistical information, that is the information on the individual respondents, is almost a matter of course today in empirical social and economic research, as well as in the policy advice which is derived from this. At the recommendation of the *Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik* (Commission to improve the informational infrastructure by co-operation of the scientific community and official statistics – KVI) (cf. KVI, 2001), the most important data producers have now established research data centres which make it possible to access official microdata (cf. e.g. Zwick, 2006). (Hardly a political reform project gains legal force today without its impact having been quantified empirically in advance on the basis of official microdata. Whether it be a fiscal reform project, the healthcare reform or the so-called 'Acts for Modern Services on the Labour Market' (*Hartz-Gesetze*), there is much empirical work available on all these topics. Over and above this, the number of degree and other dissertations, as well as post-doctoral theses, that are written on the basis of official microdata is rising, as the user applications of the research data centres clearly show.

These new possibilities, as well as the demands now being made on empirical social and economic research, also give rise to demands in the area of academic teaching. The KVI's report also called for microdata for scientific training. The Statistical Offices of the Federation and of the *Länder* have reacted to this demand for data by offering special micro datasets for teaching in the shape of the CAMPUS Files series. CAMPUS Files have been available since 2004 free of charge, and now files for the sets of statistics on the microcensus, income tax, cost structure survey and social assistance statistics are available at www.forschungsdatenzentrum.de. The data are offered in each case in SAS, SPSS and STATA, as well as ASCII CSV formats.

These data stocks, which are designed as Public Use Files, are presented in the chapters below, detailing both the data, and especially the methods used to

render them anonymous. The document starts in the next chapter by introducing the various types of individual data which have been anonymised and the methods of access to microdata, and then goes on to explain the special form taken by the CAMPUS Files in official statistics' product portfolios.

2. Anonymised Microdata and Forms of Access to Microdata

The characteristics of individual respondents, which are contained in datasets, are understood as micro or individual data. It is usual for a dataset to portray a respondent, that is an enterprise, household or person, with his or her characteristics that are included in a set of statistics, in the form of the numerical structure of an observed or measured characteristic. The collection, processing and plausibility checking of all the information that is contained in a set of statistics makes available the maximum volume of information from a survey. Each further step, be it processing the results in the form of drawing up tables or rendering the data anonymous for a further transmission, entails information reduction. The information reduction resulting from the manner in which the results are portrayed by the Statistical Offices led to calls for microdata on the part of the scientific community in particular.

The results are processed for statistical purposes by the Statistical Offices in a form which has already been initially anonymised, and hence information-reduced. In accordance with section 12 of the Act on Statistics for Federal Purposes – Federal Statistics Act (*Gesetz über die Statistik für Bundeszwecke – Bundesstatistikgesetz – BStatG*), the auxiliary characteristics of a set of statistics must be separated as early as possible from the so-called collection characteristics. Auxiliary characteristics are information needed for the technical implementation of a set of statistics (cf. section 10 of the Federal Statistics Act). These include, in particular, the personal identifiers of survey respondents, such as the name and address of a person or enterprise. Individual data stocks which do not have direct identifiers are referred to as formally anonymised, or as confidential individual data.

Section 16 subsection 6 of the Federal Statistics Act provided researchers with privileged access to microdata from official statistics from 1987 onwards. Accordingly, individual data may be transmitted to researchers if respondents can only be re-identified by making unreasonable efforts (de facto anonymity). "Unreasonable" means here that the effort involved in re-identification exceeds its benefit. This implies that de-anonymisation of individual information in a de facto anonymous dataset is not ruled out with absolute certainty, but that it would be unattractive for a potential data attacker to attempt de-anonymisation. Data stocks which are transmitted to researchers in a form which has been rendered de facto anonymous are hence information-reduced to such a degree that information can only be attributed to indi-

vidual respondents by making unreasonable efforts in terms of time, cost, and labour.

Overview 1

Information Content and Degree of Anonymisation

Individual data	Information content	Form of access
not anonymous	full information content	no access
formally anonymous	no direct identifiers such as name and address complete analysis content	controlled remote data processing
de facto anonymous	information reduced until effort of deanonymisation greater than benefit medium analysis content	guest researcher workplace Scientific Use Files
absolutely anonymous	information reduced until can no longer be ascribed to respondents as far as it is possible to judge low analysis content	Public Use Files CAMPUS Files

Microdata are only available to users outside the independent scientific community in an absolutely anonymous form. Here, the microdata only contain enough information so that attribution of information is ruled out as far as it is possible to judge. However, as a rule, the reduction in information is so great here that the remaining potential for sound analyses is, in general, no longer sufficient.

Standardised Public and Scientific Use Files can be used outside the Statistical Offices (off-site use). To this end, the microdata stocks are processed using a fixed anonymisation concept, and are then available for transmission to users.

Additionally, the research data centres of the Statistical Offices also offer specially-tailored data access in the form of workplaces for guest researchers and controlled remote data processing. This permits microdata to be used which are less securely anonymised, but these are only available for use in the secure areas of the Statistical Offices (on-site use).

In accordance with the instructions contained in the Federal Statistics Act, further distinctions are made as to the groups of individuals who may be provided with access to the data. Whilst Public Use Files and remote data processing may be used by all interested persons and facilities, Scientific Use Files and workplaces for guest researchers are reserved for use by independent scientific facilities.[1]

[1] Cf. Zwick (2006); Zühlke et al. (2003) as well as www. forschungsdatenzentrum.de.

3. CAMPUS Files as a Special Form of Public Use File

The Public Use Files offered so far by official statistics have met with little interest. Firstly, the fact that the information volume of these data is clearly reduced by the anonymisation procedure leads to lower demand. Secondly, these data were frequently more expensive in the past than Scientific Use Files since no subsidies were available for compiling them, and hence the total cost of compilation was as a rule passed on to the users. The combination of an analysis potential that is further restricted in contrast to the Scientific Use File, and that was more expensive, led to considerable reservation on the part of users.

The series of CAMPUS Files constitutes an entirely new approach. Firstly, this data service does not primarily aim to retain as much of the analysis potential of the data as possible and, secondly, these data are offered free of charge at www.forschungsdatenzentren.de.

The direction taken by the Statistical Offices of the Federation and the *Länder* in providing the CAMPUS Files is for these data to be used in teaching at universities. The data are intended, on the one hand, to assist in training statistical methods, and on the other to make it easier to teach application-related statistics in the field of social and economic statistics. Empirical work using 'real' data is time-consuming and error-prone. The period of time available to students for empirically-orientated degree or other theses is frequently not sufficient to get to know the minutiae and vagaries of a set of data material, and to become familiar with a statistical program at syntax level, in order to then use both for the actual topic. The CAMPUS Files offer the possibility to include major stages of the training process in the lessons themselves. Complex data material such as the microcensus or the income tax statistics can be studied prior to the actual work. It is also necessary here to tackle a statistical program at syntax level. The data material, along with the possible questions are, as a rule, designed so that clicking under SPSS and working in Excel is not sufficient.

In light of this goal, the analysis potential made available by the CAMPUS Files is of subordinate interest. The primary objective is to train using larger, complex sets of statistical data. Then, after this basic training, the much larger Scientific Use Files are available for the academic graduation work, when it will be possible to focus on the actual academic topic since it should already have been possible to become familiar with the tools, data and statistical program.

Compared to their big brothers the Scientific Use Files, the CAMPUS Files show far fewer datasets and characteristics. Thus, the 'microcensus' CAMPUS File holds only roughly 12,000 households in contrast to the Scientific Use File, which holds roughly 230,000 households. But the CAMPUS File is large

in comparison with other household samples such as the German General Social Survey (ALLBUS) or the Socio-Economic Panel (SOEP). For this reason, the CAMPUS Files are currently being optimised. The primary goal is not scientific use, but the characteristics and the sample concept are nonetheless selected in such a manner that specific topical areas can also be processed in teaching with sufficient precision.

4. The CAMPUS Files of the Statistical Offices of the Federation and of the *Länder*

CAMPUS Files, as a specific form of Public Use File, are not designed for scientific use. Because of their size and of their ongoing evolution, tailored to user requirements, it is possible to execute research applications with these data, but the primary goal of the CAMPUS Files is to facilitate training on 'real data' from the Statistical Offices of the Federation and of the *Länder*. This consideration is also followed by the anonymisation strategy applied to these data. With customary Scientific as well as Public Use Files, the anonymisation procedure must accommodate two competing goals. As a result, material is to be available that contains as great an analysis potential as possible for the data, whilst meeting the data protection requirements. The CAMPUS Files do not make this claim. The remaining analysis potential was secondary in preparing the anonymisation concepts. Because of the goal of offering files which are free of charge and available without restriction on the Internet, and therefore are available worldwide, the information reduction applied to the CAMPUS Files goes far beyond that of other Public Use Files. As a result of the primary goal of the CAMPUS Files, namely to provide additional empirical resources in teaching, the considerable information reduction does not constitute a restriction. The data make it possible to become familiar with statistics and procedures in academic training, whereas the Scientific Use Files are then available for research theses.

The CAMPUS Files of the microcensus, of the income tax statistics and of the cost structure survey, presented below, are hence similar as to their anonymisation conception. Firstly, they are all much smaller samples than can be found in the respective Scientific Use Files. Furthermore, the number of characteristics is clearly reduced, and as a rule they are only available in broad categories. What is more, earlier years are selected for the respective statistics, and hence time is used as an anonymisation measure (cf. Müller et al., 1991; Ronning et al., 2005).

No mention is made below of the CAMPUS File of 1998 social assistance statistics. As a result of the 'Fourth Act for Modern Services on the Labour Market' (*Hartz IV-Gesetz*), and the transition of ongoing social assistance into basic security benefits for job-seekers which it entailed for recipients who are

in principle able to work, there are currently no statistics available which could depict social benefits in a form comparable to the data material of 1998. One part of the statistics is currently handled by the Statistical Offices, whilst another part is handled by the Federal Employment Agency. Therefore there is currently no uniform data stock, so that the series of CAMPUS Files cannot be continued at the moment for the social assistance statistics.

4.1 The Microcensus 1998 CAMPUS File

The microcensus data are official, representative statistics of the population and the labour market in which 1% per year of all households in Germany are obliged to provide information (ongoing household sample). All in all, roughly 370,000 households with 820,000 persons take part in the microcensus (2005); of whom roughly 160,000 persons in approx. 70,000 households are in the new Federal *Länder* and Berlin East.[2] All households have the same selection probability in the microcensus. A one-tier, stratified area sample is carried out, meaning that areas (selection districts) are selected from the entire country in which all households and persons are asked. The selection districts are formed from the material of the 1987 census; a comparable selection basis was drawn up for the new Federal *Länder* on the basis of the "statistical population register".

The Scientific Use File is a de facto anonymised 70% sub-sample of households from the microcensus (in the case of microcensuses with the additional program on the housing situation: a 70% sub-sample of housing).[3] The de facto anonymous sub-sample is drawn from the original material as a systematic, random selection. The microcensus SUF is distinct over and above the sub-sample from the original file of the microcensus in that, as a result of anonymisation, certain variables are only available in the Scientific Use File in a categorised form which is less detailed.

The CAMPUS File is a 3.5% sub-sample of housing of the 1998 microcensus, taken from the original material, and contains information on 25,410 persons from 11,771 households and 11,668 dwellings. All in all, 199 variables of the original material and of the Scientific Use File were integrated into the absolutely anonymous basic data file. Three new variables, adjusted expansion factors for persons, households and dwellings, were created. The measures taken to achieve the absolute anonymity of the CAMPUS Files build on anonymisation measures to achieve the de facto anonymity of the

[2] On the microcensus cf. for instance Wirth / Müller (2006) as well as http: // www.destatis.de / themen / d / thm_mikrozen.php, cf. also http: // www.gesig.org / Dauerbeobachtung / GML / data / MZ / index.html.

[3] On the Scientific Use File of the microcensus cf. Müller et al. (1991) as well as Schimpl-Neimanns (2004); cf. also the information in footnote 2.

Scientific Use File. Over and above the measures used in the Scientific Use File, additional measures were implemented, such as taking a much smaller sample than in the SUF, as well as a further reduction in the detail of characteristics, and the additional deletion of variables. In particular, critical characteristics which occur less frequently in the original material were given even less detail. By way of example, one should refer here to Top and Bottom Coding of the variables age, income, nationality, etc., which has been adjusted to the Scientific Use File. The *Länder* breakdown by Federal *Land* was made less detailed to become a regional structure by East (new Federal *Länder* and East Berlin) and West (old Federal *Länder* and West Berlin). A third measure applied in the CAMPUS File for absolute anonymisation was the deletion of 370 variables from the original material.

Overview 2

CAMPUS File 1998 Microcensus

Sample size	25,410 persons from 11,771 households
Data collection	The starting point is the original material of the 1998 microcensus. The microcensus is a representative 1 % population sample. The CAMPUS File is a 3.5 % sample of the starting data material.
Content	In addition to information on the reference person, this Campus File contains amongst other things information on employment, job-seeking, maintenance, income and family concepts. All in all, 199 characteristics of the microcensus are available.
Data access	http: // www.forschungsdatenzentrum.de / bestand / mikrozensus / cf / 1998 / index.asp.

The expansion factors for persons, households and dwellings were adjusted in the CAMPUS File according to the bounded expansion method. The bounded expansion factors were created according to adjustment classes, thus making it possible to expand the values from the CAMPUS File to the total population with virtually no distortion.

The CAMPUS File of the microcensus was used from the outset by the Centre for Survey Research and Methodology (ZUMA). It was possible to develop the microcensus CAMPUS File on the basis of this experience (cf. Wirth / Schimpl-Neimanns, 2004). This made it possible to solve problems occurring in the sample and to correct difficulties in the expansion factors. The use of the CAMPUS File within the workshops on the microcensus, which ZUMA regularly offers, also led to a requirement to revise the list of selected characteristics. The microcensus 2002 CAMPUS File, which is published in spring 2007, has a programme of characteristics which is optimised in this respect.

After the experience of ZUMA in training (workshops), it appears to be necessary for lecturers to have a CAMPUS File of an earlier microcensus survey at their disposal. This is the only way during teaching to prepare users for the difficulties occurring in comparative analyses. Ideally, the available older Scientific Use Files of the 1976 microcensus should be used because for instance a larger number of characteristics is available for this microcensus in comparison to the 1973 microcensus (in particular educational variables), and most differences emerge in comparison to the more recent data (sample plan, survey programme, interruption in the variables, etc.). This data material is to be created in cooperation between the research data centre of the Statistical Offices of the *Länder* and ZUMA.

Further information on the 1998 microcensus CAMPUS File, as well as the data, can be found at http: // www.forschungsdatenzentrum.de / bestand / mikrozensus / cf / 1998 / index.asp.

4.2 The CAMPUS File of 1998 Wage and Income Tax

The 1998 wage and income tax statistics are a set of decentralised, secondary statistics implemented every three years. This means that the information is not collected for this statistical purpose, but is created in another context, in this case as part of the taxation process, and is put to statistical use in a second step. For these statistics, the finance administrations provide the respective information from taxpayers to the *Land* Statistical Offices at pre-set dates. The latter generate the respective *Land* results and transmit the resultant tables to the Federal Statistical Office. In a further step, the Federal Statistical Office then compiles the *Länder* results to form the federal result. Since the 1996 Annual Tax Act (*Jahressteuergesetz*),[4] in addition to the table data the individual information of the *Land* Statistical Offices is also transmitted to the Federal Statistical Office. Due to the possibility that is open to taxpayers to declare their income up to three years after its accrual, depending on the case, there is always a delay before the wage and income tax statistics are available.

The wage and income tax statistics show the primary income generation in detail for roughly 29 million taxpayers with more than 1,000 characteristics. A large amount of information is collected, starting with the seven types of income, through the various income quantities for income tax, to the assessed income tax (cf. Kordsmeyer, 2004). Limited amounts of socio-economic characteristics are also available. Results based on the individual data of wage and

[4] Reform of the "Law on Fiscal Statistics" (*Gesetz über Steuerstatistik – StStatG*) with Article 35 of the 1996 Annual Tax Act of 11 October 1995 (Federal Law Gazette [*BGBl*] I, 1250), most recently amended by Article 56 of the Act of 23 December 2003 (Federal Law Gazette I, 2848).

income tax statistics are put to intensive use in the policy arena and in the scientific community (cf. Zwick / Merz, 2007).

A Scientific Use File has been available since 2004 for the 1998 wage and income tax statistics (FAST98) (cf. Merz et al., 2006; Vorgrimler, 2006). The Scientific Use File is based on a disproportionally-stratified 10 % sample with roughly 3 million datasets (cf. Zwick, 1998).

The 1998 income tax CAMPUS File is an almost 1 % sample with 234,510 taxpayers. The complete sample, encompassing all characteristics, which was taken analogously to the sample plan of the 10 % sample, forms the basis for the tax simulation model of the Federal Ministry of Finance (cf. Lietmeyer, 2007). The 1998 income tax CAMPUS File was developed building on the anonymisation procedure that had been applied to FAST98.

Overview 3

CAMPUS 1998 Income Tax File

Sample size	234,510 taxpayers
Data collection	The starting point is the 1998 income tax statistics with roughly 29 million taxpayers. The CAMPUS File is an approximately 1% stratified sample of those unrestrictedly subject to tax.
Content	30 characteristics of the tax procedures were carried across from the income tax declarations for the CAMPUS File. All in all, 38 characteristics are available.
Data access	http: // www. forschungsdatenzentrum. de / bestand / lest / cf / 1998 / index. asp.

In a first step, the data material was structured according to three characteristic categories.

Characteristics of the first category:

– Income total (if the splitting procedure is applied: men (M) and women (F))
– Sum of all income
– Income
– Taxable income
– Income tax based on taxable income
– Assessed income tax

Characteristics of the second category:

– The seven types of income, broken down by M and F
– Special expenses which are not expenses of a provident nature

– Special expenses: expenses of a provident nature
– Extraordinary burdens, deductible – with separate assessment – M –
– Extraordinary burdens, deductible – with separate assessment – F –
– Incentives for home ownership: total tax concessions

Characteristics of the third category:

All other constant characteristics

All of the characteristics of the first category are available. The characteristics of Category 2 are only available in the CAMPUS File as dummy variables. This means if the characteristic is available for a taxpayer or tax case, that the characteristic shows the value Òne', otherwise the value 'zero'. The characteristics of Category 3 are not contained in the CAMPUS File. For the ten taxpayers with the highest income stated, measured against the total income, the constant characteristic values of Category 1 are replaced by the respective arithmetic mean of these ten taxpayers. The same procedure was followed with the ten taxpayers with the lowest income stated. This is a necessary procedure because of the instances of loss in the income tax statistics, which in some cases is very considerable, because this information, as with high income, is placed at greater risk.

Further metadata on the CAMPUS File, as well as data and data upload routines for the programs SAS, SPSS and Stata, can be found at http://www.forschungsdatenzentrum.de/bestand/lest/cf/1998/index.asp.

4.3 The CAMPUS File of the 1999 Cost Structure Survey

The anonymised microdata of the cost structure survey were actually the first CAMPUS File, but the term had not yet been coined when this dataset was compiled. A Public Use File of the cost structure survey came about in a first phase within the project entitled Ànonymisation of economic statistical individual data', the goal of which was to compile an initial dataset of economic statistical data. In order to achieve this goal, the information potential of the cost structure survey data was considerably reduced. Over and above this, the characteristics of the various characteristics were treated with data-changing procedures. The resulting dataset is only of restricted use for well-founded scientific analyses. However, for the first time this file provided an anonymised dataset on enterprises from official statistics (cf. Ronning / Gnoss, 2003 on these first project results). A Scientific Use File was then developed in the further course of the project, which provided much more analysis content.[5]

[5] Cf. Ronning et al. (2005) on the results of the project "Anonymisation of economic statistical individual data".

Overview 4

1999 Cost Structure Survey CAMPUS File

Sample size	500 enterprises with a maximum of 250 staff
Data collection	As a stratified random sample, the cost structure survey of manufacturing covers roughly 13,000 enterprises per year with 20 and more workers.
Content	The survey program covers the workers, the total turnover, the costs by cost types, as well as investment.
Data access	http: // www.forschungsdatenzentrum.de / bestand / kse / cf / 1999 / index.asp.

In terms of content, the cost structure surveys of manufacturing provide comprehensive information on enterprises in the field of the production industries. They serve as a starting point for many types of structural survey in the policy arena and in administration, as well as in the field of economics, and here in particular in the associations. The information from the cost structure survey also constitutes a data basis for the National Accounts. Here, the results are mainly used to calculate the added value and its components by economic areas in the context of the output-based method; finally, they also provide important information for the input-output accounts.

The 1999 manufacturing cost structure survey – reduced to enterprises with from 20 up to and including 249 staff (small and medium-sized enterprises) – as a sample contains about 13,000 enterprises. The questionnaire is handled centrally by the Federal Statistical Office and completed by the enterprises themselves. The results obtained in the sample are expanded to cover the whole range of enterprises with between 20 and 249 staff. A new sample is taken every four years as a rule, so that the burden on smaller and medium-sized enterprises can be reduced by rotation.

The data of the cost structure survey basically depict the added value process:

Total turnover

\+ stock changes of unfinished and finished goods from own production

\+ own-account production

= gross production value (total)

– material consumption, use of goods for resale, cost of wage work

= net production value

– other intermediate consumption

= gross added value

– indirect tax (not incl. turnover tax) minus subsidies

= gross added value at factor costs
– depreciation
= net added value at factor costs

The CAMPUS File on the cost structure survey also makes it possible to comprehend the net added value for the participating enterprises. In addition to this possibility, this CAMPUS File offers the opportunity to illustrate the effect of alternative anonymisation procedures. Data-changing procedures were also applied to this dataset, in addition to the classical information-reducing procedures mentioned above.[6] It is shown here that the procedures of coincidental overlapping of characteristics can have serious disadvantages, depending on the data conception. The application of these anonymisation procedures to the CAMPUS File on the cost structure survey has led to the loss of the defining context between total turnover and net added value. This problem particularly makes it possible to explain to students the advantages and problems of different anonymisation procedure with 'real' data.

Both microdata and metadata of the cost structure survey of manufacturing can be found at http: // www.forschungsdatenzentrum.de / bestand / kse / cf / 1999 / index.asp.

5. Usage and Outlook

There is currently only restricted knowledge of the degree to which the CAMPUS Files are being used. If one searches on the Internet with this term, however, one does now find various applications within education. The CAMPUS Files are being used by the author at the Johann Wolfgang Goethe University in Frankfurt in the "Economic statistics" and "Multivariate statistics" courses. Over and above this, little is known as yet about the use of the CAMPUS Files beyond their application at ZUMA. This will change as a result of two measures.

During the course of the year the introduction of technical measures will lead to the Statistical Offices of the Federation and the *Länder* learning something more about the users when they download the CAMPUS Files. There will still be no charge to download the data from the CAMPUS Files, but it is planned to introduce registration before giving access to the data, and users will be able to choose whether they wish to be informed in future of news or events.

A further measure to improve communication with users will be targeted workshops on the CAMPUS Files, and an initial scheme is in operation at ZUMA on the microcensus CAMPUS File. As to the CAMPUS File on in-

[6] Cf. inter alia Rosemann (2006) on classical and data-changing anonymisation procedures.

come tax, the Federal Statistical Office is likely to offer a first workshop in the spring of 2008. In addition to passing on the content, these events will also be used to evolve the content of the CAMPUS Files by entering into a dialogue with users.

The dialogue is also to help identify the need for further CAMPUS Files. We are currently aware that there are calls for the Statistical Offices to also provide a CAMPUS File on the income and expenditure survey. This project is therefore to be tackled this year.

Furthermore, further consideration is to be given to the idea of extending the concept of the CAMPUS Files towards a SCHOOL File, discussions having taken place for instance on whether and how CAMPUS Files can be combined with learning programs such as the 'statistics laboratory'. One idea in this context is to develop CAMPUS Files into smaller SCHOOL Files which could then be more easily integrated into learning programs, and also could be worked on with the aid of Excel.

References

Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und amtlicher Statistik (Commission to improve the informational infrastructure by co-operation of the scientific community and official statistics – KVI) (eds. 2001): Wege zu einer besseren informationellen Infrastruktur, Baden-Baden.

Kordsmeyer, V. (2004): Die Einkommensteuerstatistik als Mikrodatenfile, in: J. Merz / M. Zwick (eds.), Mikroanalysen und amtliche Statistik – MIKAS; Reihe 'Statistik und Wissenschaft' Statistisches Bundesamt, Vol. 1, Wiesbaden.

Merz, J. / *Vorgrimler,* D. / *Zwick,* M. (2006): De facto anonymised microdata file on income tax statistics 1998, Schmollers Jahrbuch – Zeitschrift für Wirtschafts- und Sozialwissenschaften 126 (2), 313 – 327.

Müller, W. / *Blien,* U. / *Knoche,* P. / *Wirth,* H. et al. (1991): Die faktische Anonymität von Mikrodaten, Schriftenreihe Forum der Bundesstatistik, Vol. 19, Stuttgart.

Ronning, G. / *Gnoss,* R. (2003): Anonymisierung wirtschaftsstatistischer Einzeldaten, Forum der Bundesstatistik, Statistisches Bundesamt, Vol. 42.

Ronning, G. / *Sturm,* R. / *Höhne,* J. / *Lenz,* R. / *Rosemann,* M. / *Scheffler,* M. / *Vorgrimler,* D. (2005): Handbuch zur Anonymisierung wirtschaftsstatistischer Mikrodaten, Statistisches Bundesamt, Statistik und Wissenschaft, Vol. 4.

Schimpl-Neimanns, B. (2004): Anwendungen und Erfahrungen mit dem Scientific Use File des Mikrozensus, in: J. Merz / Zwick, M. (eds.), Mikroanalysen und amtliche Statistik – MIKAS; Reihe 'Statistik und Wissenschaft' des Statistisches Bundesamt, Vol. 1, Wiesbaden.

Vorgrimler, D. (2006): Anonymisierte Daten der amtlichen Steuerstatistik, FDZ-Arbeitspapiere No. 13.

Wende, T. / *Zwick,* M. (2003): Research data centres of the official statistics; Statistical confidentiality and access to microdata, Proceedings of the Seminar Session of the 2003 Conference of European Statisticians United Nations Economic Commission for Europe.

Wirth, H. / *Müller,* W. (2006): Mikrodaten der amtliche Statistik – Ihr Potential in der empirischen Sozialforschung, in: A. Diekmann (ed.), Methoden der Sozialforschung, Kölner Zeitschrift für Soziologie und Sozialpsychologie, special issue 44.

Wirth, H. / *Schimpl-Neimanns,* B. (2004): Anregungen zur konzeptionellen Überarbeitung des Campus File Mikrozensus 1998, German Microdata Lab, Arbeitspapier 03.

Zühlke, S. / *Zwick,* M. / *Scharnhorst,* S. / *Wende,* T. (2004): The research data centres of the Federal Statistical Office and the statistical offices of the Länder, Schmollers Jahrbuch – Zeitschrift für Wirtschafts- und Sozialwissenschaften 124 (4), 567 – 578 ff.

Zwick, M. (2006): Forschungsdatenzentren, Nutzen und Kosten einer informationellen Infrastruktur für Wissenschaft, Politik und Datenproduzenten, Wirtschaft und Statistik 12.

Zwick, M. (2001): Individual tax statistics data and their evaluation possibilities for the scientific community, Schmollers Jahrbuch – Zeitschrift für Wirtschafts- und Sozialwissenschaften 121 (4), 639 ff.

Zwick, M. / *Merz,* J. (eds.) (2007): Mikroanalysen und Steuerpolitik, Statistik und Wissenschaft 7, Wiesbaden.