

## Teaching and Statistical Training

Rendt, Ulrich

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

### Empfohlene Zitierung / Suggested Citation:

Rendt, U. (2009). *Teaching and Statistical Training*. (RatSWD Working Paper Series, 81). Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD). <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-80578-2>

### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Rendt, Ulrich

Working Paper

## Teaching and Statistical Training

RatSWD Working Paper, No. 81

**Provided in Cooperation with:**  
German Data Forum (RatSWD)

*Suggested Citation:* Rendt, Ulrich (2009) : Teaching and Statistical Training, RatSWD Working Paper, No. 81, Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin

This Version is available at:

<http://hdl.handle.net/10419/186189>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

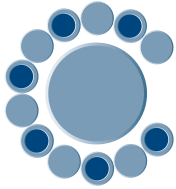
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



German Council for Social  
and Economic Data (RatSWD)

[www.ratswd.de](http://www.ratswd.de)

# RatSWD

## *Working Paper Series*

Working Paper

No. 81

### Teaching and Statistical Training

---

Ulrich Rendtel

---

March 2009

---

## Working Paper Series of the Council for Social and Economic Data (RatSWD)

---

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)

# Teaching and Statistical Training

**Ulrich Rendtel**

*Freie Universität Berlin (Ulrich.Rendtel[at]fu-berlin.de)*

## **Abstract**

The availability of well-educated researchers is necessary for the fruitful analysis of social and economic data. The increased data offer made possible by the creation of the Research Data Centers (RDCs) has resulted in an increased demand for PhD students at the master's or *Diplom* levels. Especially in economics, where we find intense competition among the various individual subjects within the course of study, survey statistics has not been very successful in laying claim to a substantial proportion of the coursework and training. The situation is more favorable in sociology faculties.

This article argues that the creation of new *CAMPUS Files* would help foster statistical education by providing public use files covering a wider range of subjects. It also presents some suggestions for new *CAMPUS Files* along these lines. Additionally, it argues for the establishment of master's programs in survey statistics to increase the availability of well-trained statisticians. An outline of such a master's program is presented and current PhD programs are evaluated with respect to training in survey statistics.

Training courses are also offered outside the university that promote the use of new data sets as well as expanding the knowledge of new statistical methods or methods that lie outside standard education. These training courses are organized by the RDCs, (i.e. the data producers), the Data Service Centers, or by GESIS (Leibniz Institute for the Social Sciences). The current tendency to strengthen ties and collaborate with universities should be supported by making it possible to earn academic credit for such courses.

**Keywords:** master's programs, survey statistics, campus files, statistical training.

## 1. Introduction

A major issue identified by the KVI, the Commission to Improve the Statistical Infrastructure in Cooperation with the Scientific Community and Official Statistics (*Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik*), is the relationship between data access and the ability to analyze these data competently. For this reason, the original KVI proposal voted for the creation of *CAMPUS Files*, free public use files to support academic teaching, as well as new training courses on Scientific Use Files (SUFs) (KVI 2001, 32). In this paper I review the current state of statistical teaching and training in Germany with respect to the use of new information sources that became available during the first phase of the RatSWD, the German Council of Social and Economic Data (*Rat für Sozial- und Wirtschaftsdaten*).

Several aspects of university training in statistics will be addressed. First, both economics and the social sciences are affected by the transition from the educational model of the German *Diplom* to the bachelor's and master's program model. This transition, which is ongoing, has strong implications for university curricula. The impact of this change was not foreseen by the KVI and is analyzed in the opening sections. One new feature of German university education today is the obligatory emphasis on structured doctoral programs and graduate schools. The relationship between these new branches to the present availability of statistical training is discussed in the Section Four. In this context, I will also introduce the role of training courses held outside the university, namely courses run by GESIS (Leibniz Institute for the Social Science) and the Research Data Centers (RDCs). Finally, some concluding remarks will be made with respect to international comparisons.

## 2. Consequences of improved data access

During the first phase of the RatSWD there was a strong emphasis placed on data access; namely, the development and production of Scientific Use Files (SUFs) and their deployment by the Research Data Centers (RDCs). Corresponding to the federal structure of Germany there is a total of sixteen state agencies and one federal state agency that offer SUFs as well as on-site access to data sets where no SUFs exist, such as firm-level data or household data with detailed regional information. This expanding data offer has resulted in a sharp increase of users. From the beginning of 2004 to 2007, the number of new data contracts rose by a factor of seven. Given that this process of improved data access has not yet come to an end, one may reasonably predict an additional dramatic increase in the number of data user contracts.

A further consequence of the increased number of research contracts at the RDCs has been an increased number of job openings in the area of applied data analysis. As a register of this increased demand, I have looked to the SOEP mailing list, a forum for advertising job openings in the field of applied data analysis.<sup>1</sup> In this venue, the number of job offers (including academic research) has risen from eleven (in the second half of 2004) to thirty (in the first half of 2008). The positions offered are mainly part-time jobs (half or two-third positions) that include the opportunity of writing a doctoral thesis.<sup>2</sup>

The principal qualifications required for these positions include: competence in handling data generated by complex surveys, background in statistical methodology, the ability to run statistical analysis packages (i.e. STATA, SPSS, R, or SAS), and some familiarity with a special substantial topic, such as labor economics or gender diversity. In these areas, however, representatives of the Federal Statistical Office in Germany have complained of a lack of skills and education, especially among economists who have recently left the university (see Rendtel 2008). Specifically, it is said that young economists aren't familiar with the important surveys in official statistics, that they don't know the framework of survey methodology, and have limited experience handling empirical data - for example in dealing with item nonresponse or coding errors. Sociologists, on the other hand, are regarded as better trained. They seem to profit from mandatory courses on empirical methods in surveying in their field, which are not included in the standard program of economic study.

To summarize, there is a gap created by the increased demand for young researchers with a sound knowledge of important surveys and data handling and an insufficient amount of statistical training. This observation, however, is more characteristic of university programs in economics than it is of sociology departments.

### 3. *CAMPUS Files*

One measure taken to narrow this gap is the use of *CAMPUS Files* (CF) in academic teaching. These files are created for use in statistical training. Because of the lack of controls in their use by students, the level of anonymization should be higher than in the case of SUFs. In general, they are regarded as absolutely anonymous Public Use Files (PUF), which restricts their power for analysis (see Zwick 2008).

---

1 The mailing list has existed in its present form since March 2004. The results reported here should be interpreted with some caution. Other effects, such as a potential increase of list subscribers, may also have induced a larger number of job offers. Help from the SOEP group, especially Uta Rahmann, in providing this information is gratefully acknowledged.

2 Compared to job offers from the private sector, the income earned in these positions is quite unattractive.

At the moment (August, 2008), there are eight *CAMPUS Files* offered by the Research Data Center of the Federal Statistical Office (RDC-FSO): two from the German Microcensus (1998, 2002), two taxation data files (*Lohn- und Einkommenssteuerstatistik* 1998, 2001), a file of employee and firm data on wages (*Gehalts- und Lohnstrukturerhebung* 2001), a file of the German subsample of the European firm-level panel data on the impact of job training, a file of social aid recipients (*Sozialhilfestatistik* 1998), and finally, a file of small and medium-sized firms on cost structure (*Kostenstrukturerhebung kleine und mittlere Unternehmen* 1999).<sup>3</sup> Since they allow no identification of units in the files, there is no control over what is done with the file data.

A more restrictive use of data is offered by the RDC of the German Pension Fund (*Deutschen Rentenversicherung Bund*). In this case, the instructor must apply for a CF and notify each student who receives a copy of the file.<sup>4</sup> There are four of these files offered for teachers: two on the stock of the retired persons (2003, 2005) and two on recently retired persons (2003, 2004). For the social sciences, the ALLBUS (German General Social Survey) is also offered as a *CAMPUS File*, although the title is not explicitly used.<sup>5</sup> Other surveys are offered for a modest fee by the GESIS Data Archive and Data Analysis for use in teaching.

It is relatively surprising that one of the most frequently analyzed data files, the Socio-Economic Panel Study (SOEP), is not represented as a CF. Teachers are allowed to use a 50% subset of the SOEP that they must construct themselves. However, this file cannot be given away to the students, which makes it less than optimal as a teaching option. Moreover, the SOEP is a collection of more than one hundred flat files across six subsamples. The complexity of this data structure is overwhelming for untrained students. Thus it seems desirable to have a SOEP file that can be distributed to students and that has a simpler structure than the full SOEP.<sup>6</sup> Nevertheless researchers who run analyses with the SOEP need to be trained on a SOEP version that has the full complexity of a long-running household panel. A *CAMPUS File* version of the SOEP would arguably present an educational tool at a level somewhere between the full complexity of the original file and that of a collection of mere analysis files to demonstrate the syntax and outcome of statistical program packages.<sup>7</sup>

---

3 These files can be downloaded from the RDC website ([www.forschungsdaten-zentrum.de/campus-file.asp](http://www.forschungsdaten-zentrum.de/campus-file.asp)).

4 See [www.RDC-rv.de](http://www.RDC-rv.de)

5 The ALLBUScompact Cumulation 1980–2006 covers 13 biannual cross-sectional surveys, see <http://www.gesis.org/en/services/data/survey-data/allbus/>

6 The complexity of the data structure is to some extent buffered by the retrieval system, the SOEPinfo meta analysis program (see [www.diw.de/SOEP/](http://www.diw.de/SOEP/)). One easy way to reduce the complexity of the full SOEP might be to put aside all subsamples with the exception of the Subsample F, which was started in 2000.

7 See for example the collection of SOEP files in STATA format used to support Kohler and Kreuter's textbook, *Datenanalyse mit Stata* ([www.stata.com/datenanalyse/](http://www.stata.com/datenanalyse/)).



In response to the broader range of data sources that can be analyzed now, the topics covered by *CAMPUS Files* should be correspondingly enlarged. For example, the EVS, or German Income and Consumption Survey (*Einkommens- und Verbraucherstichprobe*), is a basic source of poverty research. Also the German Microcensus, which has followed a continuous sampling scheme since 2005, is not represented by a *CAMPUS File*, nor it is used as a rotating panel over three years.<sup>8</sup>

Generally speaking, a good infrastructure should offer a CF for each subject area. For example, the data of the Federal Employment Agency have become a must for a labor economist, yet there is no CF for the RDC of the Institute for Employment Research (IAB – *Institut für Arbeitsmark- und Berufsforschung*) which distributes the data of the federal employment agency. Alternatively, European data sets such as the EU Statistics on Income and Living Conditions (EU-SILC), a rotating panel started in 2005, would offer another good candidate for a European CF.

For sociologists, the European Social Survey (ESS) is an important international data source. Compared with EU-SILC, the situation is again much more advantageous. The ESS EduNet is an internet-based analysis training program developed by Norwegian Social Science Data Services. It not only provides data access but also a teaching environment.<sup>9</sup> Further areas that deserve more attention include health surveys and educational data.<sup>10</sup>

The downloadable format is very convenient for teachers and students. However, other formats may be equally attractive for the dissemination of data for seminars and projects. For example, the British Economic and Social Data Service (ESDS) offers a data sharing option that allows the teacher to distribute data to his or her students under the condition that the students are registered and have signed an agreement concerning the terms of data usage.<sup>11</sup> More information on this can be found on the ESDS website.

#### **4. After Bologna: The situation of statistical education in Germany**

The most important outcome of the Bologna Process is the transition from a single phase *Diplom* curriculum to a two-phase scheme with a three-year bachelor's and a two-year master's phase. Compared with the German *Diplom* and its four-year schedule, the bachelor's

---

8 For more information, use the search tool on the website of the Federal Statistical Office, (<http://www.destatis.de>) for the Microcensus Panel Project (MZ-Panel).

9 See <http://essedunet.usd.uib.no/cms/edunet/about.html>

10 The large-scale Educational Panel Study (Bildungspanel), for example, might be a good candidate for demonstrating the difficulty of analyzing school data.

11 <http://www.esds.ac.uk/ordering>Data/sharing>Data.asp>

phase is significantly shorter.<sup>12</sup> However, this has given rise to competition between the individual subjects within a faculty over their representation in the shorter bachelor's framework.

There are two models for coping with such a situation: (a) all courses are required to cut their curriculum by approximately 30 percent; and (b) a narrower selection of courses lead to a more specific BA exam. The decision in favour of either model depends on the individual preferences and composition of local faculties. An empirical analysis of the economic curricula in various faculties was presented by Rendtel (2008). This study compared 117 BA degree programs at economic faculties in universities and technical schools (*Fachhochschulen*). In one quarter of the cases the percentage of mandatory credits to be earned from quantitative courses was found to be less than 5.5%! At the other extreme, a small number of some BA programs required 25% or even up to 30% of course credits from mandatory quantitative courses.

In addition to changes in course requirements, the format of the written diploma thesis, a final year project, has changed from a one that would have been finished within four to six months, to one that must be completed in nine weeks.<sup>13</sup> Such a short time frame excludes examination topics requiring substantial empirical data analysis.

The large disparity in required quantitative credits illustrates the extreme diversity of different subjects within economic faculties. In these faculties, business administration recruits the majority of students and often has interest in subjects that do not use statistical inference or survey data. For example, in the bachelor's program in Business Administration at the Free University of Berlin, statistical inference is no longer a mandatory course. As a result, one may expect a large variation in the statistical skills of new BAs graduating from different economic faculties. In the case of sociology, university departments seem more homogeneous. In this discipline the role of survey data and empirical statistical analysis in the educational program seems to be well recognized.

Nonetheless, it does seem that at the moment students with a bachelor's degree are not qualified for research projects in empirical data analysis. The usual qualifications that are listed in the job descriptions correspond rather to the *Diplom* or the master's level of study. Thus, unless there is a substantial progression of students from the bachelor's into the master's level, one may predict a decrease in candidates qualified for high-level data analysis.

---

12 This effect is reinforced by mandatory general occupational skills training, comprised of languages, internships, or word-processing. At the Free University of Berlin this block of required study amounts to 30 credits, equal to the workload of a semester.

13 12 credit points equal a total of  $12 \times 30 = 360$  working hours. With a weekly workload of 40 hours one obtains  $360/40 = 9$  weeks.

The heterogeneity of qualifications increases at the master's level. There is a trend toward highly specialized master's degrees. Again the diversity of master's degrees seems to be much greater in economics than in other faculties. This trend towards tailored master's degrees has given rise to highly specialized courses in the curricula, such as "Quantitative Methods in Finance." These replace statistical courses of general relevance, such as "Multivariate Analysis." Unless the master's program is geared specifically toward to statistics and data analysis there will be no possibility of obtaining a sound education, for example, in survey statistics.

This article therefore proposes the establishment of master's programs tailored to the needs of empirical data analysis with a special emphasis on survey data. Such a program should cover the framework of design-based statistics, i.e. sampling from a finite population with known inclusion probabilities, since most of the RDC files come from surveys with informative sampling. The calibration of survey data - often simplified as "weighting," which is the standard routine in official statistics - should also be given more attention. Furthermore, the issue of nonresponse and some strategies to cope with it is an important topic for everyone who utilizes survey data. In fact, missing data not only occur as nonresponse but they also occur in evaluation studies as one missing observation in treatment-control pairs.<sup>14</sup> Measurement error is another important issue for everyone who analyses survey data.<sup>15</sup> Measurement errors overlap with survey techniques and questionnaire design. This is an area in which social scientists are well-trained but it is much less familiar to economists. Last but not least, there should be extensive training in basic skills (i.e., data management, model selection, data presentation, and interpretation). This can be supplemented, for example, with internships at the RDCs or other research institutes, such as the Institute for Employment Research (IAB) or the German Institute for Economic Research (DIW).

One of the big statistical events in the near future, i.e. the German Census 2011, will be mixture of data from different sources. Such a design, which is complicated by overcoverage (*Karteileichen*) and undercoverage (*Fehlbestände*), is a methodological challenge for the Federal Statistical Office and it will be a challenge for those who analyze a SUF based on this census. Moreover, as regional counts are one of the most important issues of the census, the use of small area estimates is on the agenda. Whether it is accessed on-site or via the installation of platform for remote access, small area estimation will become a topic for data users.

---

14 See Rässler (2006) for an overview of this in the context of data from the Federal Employment Agency (Bundesagentur für Arbeit).

15 For illustrative examples see Raghunathan (2006) and Durrant (2006)

However, none of these topics are the focus of a master's program in Germany. Neither has survey statistics been prioritized at the two German statistical faculties in Dortmund and Munich. Almost no graduates from the Dortmund program, of the some 1000 *Diplom* statisticians the department has produced, are working in official statistics (see Thöne and Weihs 2008). Here biometrics, computational statistics and, not least, the facilities own demand for doctoral candidates were the largest fields where the graduates were employed.

There are obvious problems in terms of teaching capacity with trying to establish this type of intensive focus on surveys statistics at one site. To run a master's program like this would require the teaching power of at least three chairs in statistics who had a close affiliation to survey statistics. At the moment no German university offers such a concentration of energies in survey specialization. However, one might assemble the teaching resources residing at different universities in a joint master's program as a second-best solution. There are still problems with teleteaching from different sites, but given the technical possibilities that exist, teleteaching survey statistics at the master's level seems a feasible solution.

## **5. After the Master's: Vocational training and PhD programs**

It is clear that the new data sets that have been generated by the RDCs require some introduction for interested users to acquaint them with the potentials and risks of the data set. In general, this type of training units last about three days and includes practical exercises with the data. The standard clients are young researchers who are at the beginning of some empirical project and/or their thesis project. Most participants have just finished their *Diplom*. The level of statistical proficiency is quite mixed. Very often researchers lack even an elementary knowledge of the design-based approach, and models beyond the linear regression model (e.g. Logit model or Loglinear models) are unknown. To my knowledge there is no systematic test of the statistical knowledge given to participants of such training courses.

The need for data training courses was recognized early on by the SOEP project, which has offered an annual training course at the German Institute for Economic Research (DIW) in Berlin since 1989. This opportunity has now been expanded by its integration into the university framework and for the two years they have organized a workshop series, SOEP@campus, in collaboration with other universities.<sup>16</sup> The participation is partly sponsored by the Federal Ministry for Education and Research (BMBF) as a part of the KVI process. The RDC of the Institute for Employment Research (RDC-IAB) has offered a

---

<sup>16</sup> See, for example, <http://www.uni-due.de/soziologie/-soepatcampus/index.php>

workshop on spell data on the basis of some of its test data. Again with sponsorship from the BMBF, the RDC of the Federal Statistical Office and the Data Service Center of the German Microdata Lab at GESIS have been offering workshops on newly released data files. Here the Microcensus and the Microcensus Panel have played an important part.

Often data producers have established user groups that convene for regular meetings (annual or biannual), where results can be presented and discussed. The user group can be regarded as an academic network for the exchange of knowledge and experience. Therefore it can support statistical training in multiple ways.

Within the framework of statistical training, GESIS—the Leibniz Institute for the Social Sciences plays an important part. GESIS is a member of the Leibniz–Gemeinschaft and provides statistical education on subjects that are not routinely offered at university. Thus it supplements university education, for example, by presenting courses on latent class analysis, multi-level models or mixture models. Their “Spring Seminar” is devoted to an intensive training on special methods, usually presented in a sequence of three blocks of one week each.

The ZUMA branch of GESIS offers workshops in different fields (new data sets, interview techniques and questionnaire design, sampling, editing, and statistical methods). The workshops are presented in a three-day format at Mannheim. The participants have to pay a moderate fee. The number of participants is limited (14–18), so waiting lists have been created. The demand is such that the average waiting list is as long as the number of participants, numbering about 400 per year. There is nothing similar to GESIS in economics. One reason is probably the greater heterogeneity of the research areas.

At the international level, there are similar bodies that offer training and statistical instruction in survey statistics and the analysis of survey data; however, their organization differs. The National Survey Research Center in the US is affiliated with the University of Michigan and involves the development, refinement, and propagation of the scientific method of survey research through teaching and training.<sup>17</sup> The National Center for Research Methods (NCRM) in the UK is a network of research groups, each conducting research and training in an area of social science research methods.<sup>18</sup> It acts under the auspices of the Economic and Social Research Council (ESRC), the British funding organization that integrates research activities in this field.<sup>19</sup> The network is organized according to a “hub-model,” where the Southampton School of Social Science serves as the hub that connects six nodes. These nodes are located at separate universities and each specializes in certain topics or methods. The

---

17 See [isr.umich.edu/src/](http://isr.umich.edu/src/)

18 See [www.ncrm.ac.uk](http://www.ncrm.ac.uk)

19 In Germany this council would cover the activities of DFG, the BMBF and the RatSWD.

whole project runs under a four-year funding scheme.

Under the auspices of the German Research Foundation (*Deutsche Forschungsgemeinschaft*), or DFG, several PhD programs are offered for statistical teaching and training. However, these programs are only open to those few students who were accepted in the program. Moreover, most of these graduate schools, such as the Mannheim Graduate School of Economic and Social Sciences or the Bremen International Graduate School of Social Sciences are integrated into the German Universities Excellence Initiative.<sup>20</sup> Thus they are not oriented toward broader participation like the GESIS training courses.<sup>21</sup>

A different approach was proposed in the DFG-funded Priority Program on Survey Methodology which was started in 2007. Here the intention was to establish survey methodology as an independent subject. For this purpose the program plans to establish a “German School of Survey Methodology.”<sup>22</sup> The proposal incorporates international experts in survey methodology as teachers and includes a nationwide recruitment of students. This proposal is similar to the proposal for the establishment of a master’s program in survey statistics.

A few comments need to be made concerning the relationship of university teaching within the bachelor’s and master’s scheme and those training programs that lie outside this scheme:

- 1) The two levels should be adapted to each other. It is my impression that sometimes the participants of training courses lack both an elementary knowledge of statistics and experience with empirical data analysis.
- 2) Quite often the motivation of students to participate in a training course is low because they cannot earn credits toward their master’s degree. The credit system is very flexible, however, which makes it easy to grant credit for participation in training courses. A necessary prerequisite to this, of course, would be some kind of examination of the attendees by the trainers.

## **6. Conclusions and Recommendations**

The availability of well-educated researchers is necessary for the fruitful analysis of social and economic data. The increased data offer made possible by the establishment of the

---

<sup>20</sup> Gess.uni-mannheim.de; www.bigsss-bremen.de

<sup>21</sup> One may regard low admission numbers as intrinsic for excellence. However, with respect to the need of a higher number of well trained researchers this might be also regarded as a kind of luxury.

<sup>22</sup> See www.survey-methodology.de

Research Data Centers (RDCs) has resulted in an increased demand for PhD students at the master's or *Diplom* level. Even today it is not an easy task to recruit young researchers with a sound education in the methods of data analysis who also have some practical experience in this business. Especially in economics, where we find intense competition among the various subjects within the course of study, survey statistics has not been very successful in laying claim to a substantial proportion of the course work and training. The situation is more favorable in the sociological faculties.

This article proposes the creation of new *CAMPUS Files*, free public use files that would help foster statistical education by covering a wider range of subjects. It also advances some specific suggestions for new *CAMPUS Files* along these lines. Additionally, it argues for the establishment of master's programs in survey statistics that can help increase the availability of well-trained statisticians, and provides an outline of such a master's program.

There is also a widespread network of training courses that address the needs of young researchers. These programs provide training in the introduction of new data sets as well as in non-standard analysis techniques. These training courses are organized by the RDCs (i.e., the data producers), the Data Service Centers (DSZs) or by GESIS. Recently, there has been a greater tendency toward collaboration with universities. In order to attract students before their exam - and thus enlarging the number of applicants for research projects - one should investigate the possibility of granting academic credit for bachelor's and master's students.

The close cooperation between the SOEP group and universities is regarded as a fruitful model of this approach. Likewise, the RDCs offer not only data but also support for the analysis of these data. They should be encouraged to reinforce and expand their training activities. This will not only improve statistical education in the university but will help widen the scope of official statistics from a mere data producer to an information provider.

## References:

- Durrant, G. (2006): Missing data Methods in Official Statistics in the United Kingdom: Some Recent Developments, *ASTA Advances in Statistical Analysis*, Vol. 90, 577-593.
- KVI (2001): Wege zu einer besseren informationellen Infrastruktur, Baden-Baden.
- Rässler, S. (2006): Der Einsatz von Missing Data Techniken in der Arbeitsmarktforschung des IAB, *ASTA Advances in Statistical Analysis*, Vol. 90, 527 - 552.
- Raghunathan, T. (2006): Combining Information from Multiple Surveys for Assessing Health Disparities, *ASTA Advances in Statistical Analysis*, Vol. 90, 515 – 526.
- Rendtel, U. (2008): Statistikausbildung und Amtliche Statistik. Kritik und Perspektiven, *ASTA Wirtschafts- und Sozialstatistisches Archiv*, (2), 5-20.
- Thöne, M. and Weihs, C. (2008): Vielseitig und gefragt. Absolventen und Absolventinnen der Dortmunder Studiengangs Statistik, *ASTA Wirtschafts- und Sozialstatistisches Archiv*, (2), 75-92.
- Zwick, M. (2008): Campus-Files - Kostenfreie Public-Use-Files für die Lehre. *ASTA Wirtschafts- und Sozialstatistisches Archiv*, (2), 175-188.