THE LECTURE SERIES "ECONOMIC STATISTICS. DATA PRODUCTION AND DATA ANALYSIS IN THE OFFICIAL STATISTICS" FROM THE FEDERAL STATISTICAL OFFICE

Tim Hochguertel and Dr. Markus Zwick
Federal Statistical Office, Research Data Centre, Germany
markus.zwick@destatis.de

Since winter semester 2007/2008, the Federal Statistical Office offers the lecture series "Economic Statistics. Data Production and data analysis in the official statistics" to lead advanced students to the official statistics. The lecture presents methodological aspects of the official statistics and the extensive possibilities of empirical analyzing. The lecture deals with the essential tasks, concepts and problems of official statistic. Measuring theoretical concepts to describe reality is one part of the lecture; another one is sampling theory, which is necessary to understand the design of different sample methods. Furthermore, the lecture presents the main surveys as well as their expressiveness and limits. Working with absolute anonymous micro data, the students learn about the potential of analyzing official data for their own scientific work. Using such files students can make their own analysis and simulation. The lecture finishes with a written examination, which result is included in the students' final score.

INTRODUCTION

The major task of the Federal Statistical Office is the production of information about the social and economic life in Germany. Several data concerning population, income, tax, business, households and more are collected to be prepared for further analysis and reports. With the lecture series "Economic Statistics. Data production and data analysis in the official statistics", the Federal Statistical Office is engaged to instruct students from different universities and courses of studies about the methodologies and the possibilities of the analyzing of the data from the official statistic.

HISTORY

In winter semester 2007/2008 the Federal Statistical Office started the lecture for the first time at the Johann Wolfgang Goethe University in Frankfurt. While the first lecture in that term was visited by about 15 students, the students' interest in learning about statistical methods and empirical work grew more and more over time. Regarding the course two years later, there are 84 students who registered for the final examination in February 2010 in Frankfurt. Simultaneously, the lecture series was established at the Technische Universität Dortmund, the Albert-Ludwigs-Universität Freiburg and the Ludwig-Maximilians-Universität München. With different universities, the lecture series reach students of different courses of studies. At the universities of Frankfurt and Freiburg, students of business studies and economics participate in the lecture. However, at the universities of Dortmund and München students of statistics attend in the lecture.

MOTIVATION

Lectures concerning statistics in Germany usually teach prior a stochastic-orientated methodology. Beyond computing the arithmetic mean, standard deviation, a linear regression, an analysis of variance and their assumptions, students also get to know several distributions and other basics. Unfortunately students only learn the methodology and how to use them in hypothetic cases, but often without any link to an application in real life. In general, this statistic lectures do not discuss different data sources. Without a doubt teaching a fundamental knowledge in statistics is irreplaceable. But for an empirical analysis the knowledge of theoretical statistic is not enough.

An empirical analysis needs a theoretical frame. This theory is connected with concepts, which can be used for the empirical analysis. In the process of the empirical analysis, the researcher has to decide, how he or she likes to measure this concepts. The possibilities of an operationalization depend on the data which are used for the analysis. It is necessary to find an operationalization for the concepts which can be realized with the given data. Therefore, the researcher needs a way from the general idea what he or she likes to study to the measurement in the real world. The lecture deals with the problem, how to develop a strategy for the measurement

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php [© 2010 ISI/IASE]

of theoretical concepts. On this way, the lecture wants to close the gap between theory on the one hand and statistical methods on the other hand.

Accordingly, important statistical parameters are introduced to the students. These parameters are for example the birth rate, the jobless rate or the price indices. To do a statistical computation the underlying concepts of these parameters have to be conveyed to the students. Additionally, the students learn how to deal appropriate with the measure.

The demand for such empirical lectures increases over time. The necessity for an enlargement of such a lecture is evident. The Federal Statistical Office developed the lecture „Economic Statistics. Data Production and data analysis in the official statistics" to fulfill that outstanding need. To support the students and react to that lack of empirical experience, the lecture demonstrates the possibilities and limits of analyzing official micro data. Working with empirical micro data, students are taught to interpret and understand the results of different analysis under methodological aspects.

LECTURE CONTENT

The universities of Frankfurt, Dortmund, München and Freiburg offer the lecture once a year. In general, scientific members of staffs from the Federal Statistical Office work for one week with the students. The lecture includes different important themes concerning statistics which are applied on official statistics. Starting with the process of developing an official statistic, several central official statistics are presented. Here, important concepts of different areas are covered. Problems with interpreting the data in a correct manner are discussed. Finally the students are asked to work with selected micro data to apply their acquired knowledge.

The lecture starts with the discussion of data production. Here, the lecture shows in which way data are collected in Germany. The collection process is demonstrated from the start with the legal basis to the end with the publication in different forms.

Table 1. Exemplary Schedule of lecture content

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
09.00 – 11.00		IV. household statistics	VII. Price theory	X. population and labour market statistics	XIII. National Accounts
11.00 – 13.00	I. Intro-duction	V. Confidentiality of micro data	VIII. Price statistics	XI. business statistics	Work with micro data/ final session
Break					
14.00 – 16.00	II. Samples in the official statistics	VI. Campus-Files	IX. Microcensus	XII. disclosure control of tables	
16.00 – 18.00	III. Access to data	Work with micro data	Work with micro data	Work with micro data	

As one important aspect of data production, the lecture discusses sampling theory. The most surveys in Official Statistics are sample surveys. Therefore, the lecture shows different methods of drawing a sample. First of all the benefits of elevating a sample are presented, before different sample methods are demonstrated. Here the lecture focuses on the cluster sample and the stratified random sample. These sample theories are then applied on the official statistics.

The third part demonstrates, in which way micro data in Germany can be used for a scientific analysis. The Research Data Center offers different ways for the work with micro data. With Public-Use-Files, Scientific-Use-Files, Safe Centres and Remote execution are four ways

established. Here, the lecture discusses the potentials and limits of the different way of micro data access.

The next sessions deal with selected household statistics. Each of these statistics is presented including its characteristics. Different household statistics such as the Income and Consumption statistics, the Social Economic Panel (SOEP), statistics on income taxes and further more are described and compared.

Part five demonstrates the need and methodologies of confidentiality of micro data. Aimed at the protection of personal data and retaining the probands' trust in the Official Statistic, the confidentiality is one of the most important aspects in handing out micro data and its results. The different levels of confidentiality and their differing analysis potentials are highlighted.

These Campus-Files, as a special form of Public-Use-Files, are topic of the next lecture. Being only prepared for teachings, Campus-Files are absolute anonymous and available for free. The students use these files to learn handling empirical data. The lecture demonstrates which anonymisation techniques are used to generate absolute anonymous Campus-Files and discuss the influence of the anonymisation techniques to the analysis potential.

The fundamental aspects of price theory in part seven are needed to understand the illustrated price statistics in part eight. The lecture starts with a discussion about the economical theory of prices. Based on this theory, the lecture shows how prices and changes of prices are measured by the official statistic.

The biggest household statistic in Europe, the Microcensus is presented with a separate session. The Microcensus is a representative stratified cluster sample which covers different aspects, concerning the German population. This survey includes variables to labor market participation, family life, household income and many more. One percent of all households are selected and obligated to take part on the elevation. An additional module, which changes yearly, is elevated to cover special topics, which are not part of the Microcensus, such as retirement, integration or work accidents.

The part ten and eleven demonstrate several further selected statistics. Here, the lecture present data sources, classification and statistical parameter for population, labour market and business.

If a scientific user work with micro data via Safe Centres or Remote execution, the result of the analysis have to checked for disclosure risks from the Official Statistic. Not every result of an analysis can be published. The lecture demonstrates the methods which were adopt for the disclosure control of tables.

As the last theoretical session the lecture introduce the National Accounts. In this part, the lecture demonstrates the calculation of the domestic product as an indicator of the activity of a national economy.

EMPIRICAL ANALYSIS WITH CAMPUS-FILES

The empirical analysis with Campus-Files is an essential idea and part of this lecture series. Campus-Files are absolutely anonymized micro data from the official statistics. The Files are established for empirical statistic lessons on universities. The Research Data Center offers Campus-Files from different social and economic statistics. These Campus-Files are available as a free download from the web (http://www.forschungsdatenzentren.de/campus-file.asp). Meta data as the description of the variables and the anonymization concept are part of every Campus-File.

The Campus-Files serve as a first possibility for students to get in contact with a huge data set, without any fees. Based on these data they get the chance to gain their own experience in working with empirical data and get to know what it means to analyze real life data. This purpose includes especially the possibility to demonstrate methodological issues like sampling design effects and the effect of anonymization routines. Real life data include some problems like missing data and lack of information, which the researcher is like to have. Headed by the members of staffs from the Federal Statistical Office, the students start with an own analysis. After concluding and discussing the results of their analysis students learn how to prepare for their own scientific work.

CONCLUSION

Since several terms, the lecture series „Economic Statistics. Data production and data analysis in the official statistics“ from the federal statistical office is established in four German universities. Members of staff from the federal Statistical office present the lecture periodical. In 2010 all of these four universities offer the lecture series at the third time for their students.

After the week, in which the lecture occurs, the students evaluate the lecture in an open discussion with the members of the Federal Statistitical Office and in an anoymous response of a questionary. The evaluation generates always good results. A few weeks after lecture participation the students write an examination. The result of the examination includes in the students' final score.

REFERENCES

Babbie, E. (2001). *The Practice of Social Research* (9th Edition). Belmaont: Wadsworth.

Wagner, G. (2003). Anmerkungen zu "Mehr Wirtschaftsstatistik in der Statistikausbildung für Volks- und Betriebswirte". *Allgemeines Statistisches Archiv, 87*, 473-478.

Zwick, M. (2008). CAMPUS Files: Free Public Use Files for Teaching Purposes. *Schmollers Jahrbuch 127*,655-668.