

# Towards a Survey Research and Education Center (SUREC) at the Department of Statistics, Stockholm University

---

## Background

Statistics has a strong tradition in Sweden going back to the population registration starting in the 17<sup>th</sup> century. This registration has given Sweden the longest time series in population statistics in the world. Furthermore, the Swedish administrative system, which was also founded in the 17<sup>th</sup> century, has generated a lot of valuable data about society, which have been archived and preserved for research about social and economic conditions in society. During the 20<sup>th</sup> century a wide range of statistical surveys, based on modern survey methodology – in combination with the rapidly developing information technology – has added value to the Swedish statistical system.

Thus Sweden has gold mines of survey data and administrative data, which can be used for statistical purposes and research. These data archives attract researchers from all over the world. However, the data could undoubtedly be used much more, by individual researchers and in research projects concerning all aspects of society. Advanced statistical methods and modern information technology should play an important role in exploiting this potential.

This is of course a great opportunity and a challenge for statistical and social science departments at Swedish universities. The Faculty of Social Sciences at Stockholm University has decided to meet this challenge and try to become a world leading center in data-based survey research and education in social sciences, regarding theories and methods as well as applications.

The name of the Center is Stockholm University Survey Research and Education Center (SUREC). It will be organized as a network with the Department of Statistics in two roles: as coordinator of the Center, and responsible for methodological research and training. All other departments in the Faculty of Social Sciences will be invited to become nodes in the network, with responsibility for applied research on data available within their respective areas, e.g., criminology, sociology, psychology, and economy.

SUREC will conduct both methodologically oriented and applied research projects. Closely related to the research projects there will be seminars and courses given on all levels: basic, advanced (master), doctoral, and post-doctoral.

Research groups and projects from other Swedish and international universities and institutes can become associate members of the network. Interested researchers from all over the world are welcome to cooperate with SUREC in various ways, e.g., by carrying out research entirely or partly based upon Swedish data, by giving seminars and courses on the basis of their research, joint methodological projects, etc.

Swedish official statistics are currently produced by approximately 25 government agencies including Statistics Sweden, our national statistical office. These agencies need a continuous inflow of well-trained statisticians and survey methodologists to be able to produce accurate and timely statistics. The Department of Statistics has Masters and PhD programs for official statistics that are supposed to address that need, but the number of students has so far been lower than expected.

There are several reasons for this state of affairs. First, the programs have not been sufficiently advertised and promoted. Many of the agencies are not aware of the courses that are offered. Second, the number of supervisors capable of helping students writing essays and theses is limited. Third, the courses need a continuing improvement due to the rapid development in the field, and they also need to be more applied and data-based. Still, it should be emphasized that the Masters courses provided by the department are often well-attended, even though only a fraction of the students participate in the Masters program in “Official Statistics Production including Survey Sampling.”

Some of these problems are relatively easy to solve. The department’s website is currently upgraded and contacts will be established with all agencies that produce official statistics so that training needs can be defined. The courses themselves should be developed to reflect recent advances. Course development and student supervision can be improved by inviting researchers and methodologists from abroad. For instance, the U.S. Federal Statistical System’s provider of competence development, namely the Joint Program in Survey Methodology (JPSM), <http://www.jpsm.umd.edu/jpsm/>, has expressed an interest in collaboration.

It should be noted that statistical agencies and other survey data collection organizations all over the world have a lot in common. They share the same methodological challenges, and there is already a very active formal and informal cooperation and sharing of experiences between many organizations and our own faculty members. Our network in the official statistics world is very strong and vast improvements of the department’s efforts could be made very rapidly once SUREC is launched.

## ***Survey-based research at other departments***

### **Some ongoing studies**

The use of survey data, register data and official statistics is an important part of the research carried out at the Faculty of Social Sciences at Stockholm University. The Faculty conducts more longitudinal survey studies than any other university in Sweden.

The *Project Metropolitan*, now the Stockholm birth cohort, originally at the Department of Sociology but now also at other departments and institutes at the faculty, follows a cohort born in 1953.

In the *Örebro project* at the Department of Psychology another cohort born in the early 50:s has been studied since the 60:s. Since 1988 the Department of Psychology also hosts the Project Betula, a longitudinal study analyzing aging, memory and dementia.

In 1995 the Department of Criminology started a *study of self-reported criminality among school children* that has been continued by the National Council for Crime Prevention and with methodological development at the Department of Criminology.

The project *Stockholm Boys* at the department is another longitudinal study where data from the 1950's are now supplemented by interview and register data.

The Institute for Social Research (SOFI) is responsible for the *national surveys of level of living* that have been carried out since 1968. Separate level of living surveys are being carried out for foreign born and elderly people at the Department of Sociology, and at the Center for Health Equity Studies (CHESS) and the Aging Research Center (ARC). The projects are based both on interview data and register data from Statistics Sweden.

A large number of researchers at the Faculty use data from official statistics produced by Statistics Sweden and other providers. The value of using official statistics for testing central scientific and policy hypotheses can be illustrated by the following report based almost entirely on official statistics: *Crime and Punishment in Sweden. Historical Criminal Justice Statistics, 1750-2010*, 4<sup>th</sup> edition 2011 (in Swedish with an English summary).

The official judicial statistics has also been central in the work on *European Sourcebook of Crime and Criminal Justice Statistics* – a project where one of the initiators comes from the faculty.

Survey data, register data, and other data from official statistics play an important role in the research carried out at our faculty, the largest social science faculty in Sweden. Similar theoretical and methodological issues are raised in many of the different projects at different departments and institutes, and this calls for collaboration. The Department of Statistics can benefit from the different projects that are being carried out, including contextual analyses. The projects can in turn profit from the expertise at the Department of Statistics on survey design, sampling, nonsampling errors, analysis, etc. as well as the handling of register data. The common problems – and challenges – also suggest joint master and doctoral courses.

The scope and contents of the research is further illustrated by the following list, which, however, should not be regarded as exhaustive. Examples of integrated register data bases are given. Several of the data bases are joint projects of two or more departments or institutes and have been used by researchers from different disciplines.

## **A structured list of ongoing studies**

The studies are classified into four major groups:

- Longitudinal surveys
- Integrated register data bases
- Integrated data bases from Statistics Sweden
- Official statistics

## ***Longitudinal survey studies***

### **ARC, The Aging Research Center**

- *SNACK-K, The Swedish National Study on Aging Care*  
The SNACK population study includes two projects, one on the care of people 60 +years, another on clinical examination of people 60 + years.
- *SWEOLD, The Swedish Panel Study of Living Conditions of the Oldest Old*  
This panel study on living conditions has been conducted three times through face to face interviews or telephone interviews.

### **CHESS, Center for Health Equity Studies**

- *UBCoS, The Uppsala Birth Cohort Multi Generation Study*  
This multigenerational study includes males and females born in Uppsala 1915-1929.

### **Department of Criminology**

- *The Stockholm Boys*  
This is a longitudinal study where data from the 50:s now are supplemented by interviews and register data.
- *Self-Reported Crime among Students in 9<sup>th</sup> Grade*  
This national study, that was started at the Department of Criminology, now continues at the National Council for Crime Prevention, where it is carried out every 2-3 years.

### **Department of Education and Didactics, Institute of International Education**

- *The Malmö Project*  
This project, that started in 1938, studies the impact of education on life chances.

### **Department of Psychology**

- *BETULA*  
This is a longitudinal study that started in 1988 analyzing aging, memory and dementia.
- *The Örebro Project*  
This cohort from the early 1950:s has been studied since the 60:s by means of different types of data.

### **Department of Social Work**

- *From Student to Social Worker. A Longitudinal Study of Transfer from Studies to Work Life in an Australian and Swedish Perspective*
- *Children of Alcohol and Drug Addicts – A Longitudinal Study of Conditions during Upbringing*

## **Department of Sociology including Demography**

- *LIFEINCON, Individual Life Chances in Social Context – A Longitudinal Multi-Methods Perspective on Social Constraints and Opportunities*  
The main focus of this project is on contextual factors explaining differences in young adults' life chances in a longitudinal perspective.
- *YAPS, the Young Adult Panel Study*  
Data from this panel study are used in a number of different projects concerning the development and effects of normative factors.

## **SOFI, the Institute for Social Research**

- *The Level of Living survey, LNU*  
The surveys on living conditions in the adult population, based on a panel, have been carried out five times since the first one in 1968.
- *The Stockholm Birth Cohort*  
This data base contains different types of data for those born in 1953 and living in Stockholm 1963. The project started as The Project Metropolitan at the Department of Sociology and the database is now used by researchers at several departments and institutes within the faculty.

## ***Integrated register databases***

- The Swedish Generation and Gender Survey (Sociology/Demography)
- The Swedish Work and Mortality Database, HSIA/WMB (CHESS)
- Social Policy Indicator Database, SPIN (SOFI)

## ***Integrated Databases, Statistics Sweden***

- LISA, Longitudinal database for analysis of health insurance and labor market
- LOUISE, register for taxation, education and employment
- The Multi Generation Register

## ***Official Statistics***

Considerable amounts of data available

## **SUREC activities**

### ***Methodological research***

The methodological research conducted by SUREC will initially focus on the following three areas that together cover a large part of survey research:

- statistical use of administrative and other nonstatistical data
- total survey error
- international and other comparative studies

## **Statistical use of administrative and other nonstatistical data**

### **Background and definitions**

*Administrative data* are data generated by administrative processes, typically government processes on different levels. Administrative processes are often supported by *administrative registers*. Sweden has a long tradition in using administrative registers and administrative data for statistical purposes. Huge datasets, including long time series, exist, but are still remarkably underutilized. There is a great interest from social science researchers all over the world, to get access to these data, and to cooperate with Swedish researchers.

*Other nonstatistical data* might include data generated by societal processes other than administrative processes. These processes could take place in real life or on the *Internet*. The Internet offers enormously rich new sources of data, waiting to be “tamed” by statisticians for statistical purposes.

### **Goal**

Swedish universities have good opportunities to become world leaders in international research and education networks within this area. We think that the Department of Statistics should have the ambition to become a world leader in research and education on how to use administrative data and other nonstatistical data in combination with survey data for statistical purposes.

In order to reach this goal, many theories and methods from traditional survey theory and practice can be used, sometimes after some adjustment and reformulation, but completely new theories and methods must also be developed and implemented. Well-known quality issues have to be revisited, and new quality issues have to be studied by researchers and practitioners in cooperation.

### **Possible partners**

- Statistics Sweden (SCB)
- Other departments at Stockholm University, e.g., Criminology, Psychology, and Sociology
- Swedish Institute for Social Research (SOFI) at Stockholm University
- Karolinska Institutet (KI)
- The National Board of Health and Welfare (Socialstyrelsen)
- Swedish National Data Service (NSD) at Gothenburg University

### **Possible initiatives**

- Invite internationally well-known guest lecturers to give seminars and short courses
- Invite guest researchers and doctoral students to do research and give seminars and short courses
- Organize an international conference
- Start an open access journal

## **Total survey error**

The *Total Survey Error (TSE) paradigm* embodies the best principles, strategies, and approaches for minimizing the survey error from all sources within time, costs, and other constraints that may be imposed on the survey. The TSE paradigm can be viewed as resting on four pillars of survey methodology corresponding to survey (a) design, (b) implementation, (c) evaluation, and (d) data analysis. Regarding (a), the TSE paradigm specifies that surveys should be designed to maximise data accuracy subject to budgetary and other constraints by minimizing the cumulative effects of error from all known sources. For (b), the paradigm specifies that strategies should be in place to monitor the major error sources, adapting the survey design as necessary to minimize the TSE through real-time interventions and design modifications. Regarding (c), the TSE paradigm emphasizes the importance of regularly assessing the joint effects of survey error on estimation and analysis so that continuous improvement and future design optimizations are possible. Finally, for (d), the TSE paradigm specifies that data analysis should appropriately consider the complex sampling design and the effects of nonsampling errors on the analytical results.

Examples of research areas include:

- the evolution of the TSE paradigm, Total Survey Quality, and how they relate,
- the design of surveys to reduce the effects of nonresponse, frame coverage error, measurement error, and data processing error,
- survey error measures, metrics, and monitoring strategies based upon real-time paradata,
- incorporating adaptive total design (ATD) during data collection to control survey costs and quality,
- the design of studies for assessing specific error sources such as nonresponse, frame noncoverage, interviewer errors, mode effects, and data processing error, and
- methods for analyzing complex survey data that compensate for nonresponse and measurement errors, including latent variable modeling.

The Department of Statistics has very strong ties with the yearly Workshop on Total Survey Error and most research activities would be conducted in collaboration with these colleagues. It is very important that our department is able to serve the studies conducted within the Faculty regarding the best ways to conduct the studies so that the total survey error is minimized given constraints in terms of budget limitations and comparability requirements. It is also important that the department can help other data providers enhancing the quality of their surveys through more consultation and collaboration.

## **International and other comparative studies**

In the past 20 years, cross-national research has burgeoned in every field that uses survey data. As global measurements have become key in policy planning, resource allocations and interventions at national and supranational levels, multiregional and multinational surveys have become standard in diverse fields. At national levels, too, increasing cultural diversity and a heightened concern to measure minority needs and preferences have resulted in an increase in survey research in different languages.

Many on-going multi-national programmes collect data across continents on values, opinions, student achievements, and behaviors. Examples are the International Social Survey Programme, the European Values Survey and the World Values Survey, PISA, the Comparative Election Study surveys and, most recently, the European Social Survey. The

family of multi-regional BAROMETER attitudinal surveys now organised in various ways in Asia, Africa, South America and Eastern Europe have followed the model of the EU Eurobarometer studies that have been collecting data on political and social issues across Western Europe since the early 1970's.

Whereas at one time cross-national research was commonly thought to be much like any other survey research, it has become clear in recent years to many researchers that cross-national research calls for knowledge, procedures, and tools beyond the so-called monocultural repertoire. At the same time, the recent emphases on quality and total error in general survey research have increased expectations that comparative research be seen to be addressing these issues too. In sum, the on-going increase in cross-national research coupled with innovations in general survey methodology have resulted in new interest across disciplines in how best to undertake and analyse comparative survey research projects. Awareness of the need to secure a better understanding of what cultural difference means for survey instruments and survey data has possibly never been greater.

Methodological improvements and innovations are underway on many fronts. Improved standards and procedures in multicultural survey research programmes have already had carry-over effects into general practice. Study documentation standards are a good example. The documentation available on certain studies is considerably better than in the past (cf ISSP and European Social Survey). In addition, improved process documentation is increasingly seen as the key to process quality monitoring. Joint efforts and collaborations are developing in this respect. A second good example is the considerable investment undertaken within the framework of the European Statistical System towards improving comparability through input and output harmonisation. As the recent series of quality conferences in official statistics reflect, European statistical offices have clearly recognized the need for new methodological developments to enhance comparability within a given multi-national study and also to allow better aggregations of existing regional estimates to establish estimates of multi-regional parameters.

Multicultural survey research is necessarily multidisciplinary and sound guidelines and protocols combining practice with appropriate theory are not simple to construct. In order to be effective, they need to be informed by cultural difference and local requirements. Otherwise they may fail to be accepted, may not be viable, or if followed, may not produce the intended and desired effect. Time lags in knowledge transfer across disciplines are thus vital issues in projects that need to tap into multidisciplinary expertise. Numerous disciplines and institutions are focusing increasingly on producing guidelines, protocols or monitoring systems for comparative research projects in general or for specific studies. In educational testing, for example, the International Testing Commission has been improving guidelines on design, adaptation and testing of instruments for over a decade. Starting in 2000, the European Social Survey has developed a comprehensive package of specifications and guidelines for participating countries. Another example is the guidelines for input and output harmonisation developed within the European Statistical System. Typically these guidelines have been survey-specific rather than general in nature.

Different as the substantive topics may be, there is considerable overlap across disciplines in the basic challenges faced in engaging in multi-cultural survey research. Increased exchange and collaboration between disciplines would stand to benefit all concerned. At the same time, there are numerous areas in which concrete proposals for solutions are not. Concerted collaborative efforts could lead to research providing solutions or insights that would

otherwise be beyond the means of one location, project, or discipline. In sharing problems, we might thus move closer to sharing solutions.

The Department of Statistics has very good ties with the Comparative Survey Design and Implementation Network (CSDI) and most research activities would be conducted in collaboration with these colleagues. SUREC will host CSDI's 2013 workshop March 21-23. The department also have strong ties with some specific international surveys as well as the European Statistical System.

### ***Education methods***

New web-based or web-supported education methods, enhanced by innovative pedagogical models and disruptive information technologies are rapidly gaining ground at Swedish universities and internationally. SUREC with its collaborative network organisation and international orientation will take advantage of this development and offer students and researchers to reap the benefits of the new education methods and technologies. By investing in high quality, web-based courses, developed and executed by the best available teachers, SUREC will improve the quality, efficiency and student satisfaction of its courses and programs.

### ***Activities and staffing***

SUREC is led by the Department Head and staffed by those department members who have an interest in the survey methodology and official statistics area. A critical mass of interested and competent Masters and graduate students that can act as researchers and research assistants will be recruited.

Short courses on (a) weighting, (b) recent nonresponse research, (c) paradata, and (d) structural equation modelling are lined up during February-May 2013. As mentioned, SUREC will also host the 2013 CSDI workshop. Other activities including research will be decided after consultation with the Faculty of Social Sciences and agencies producing Swedish official statistics.