Resource Provision of the Methodological Grounding of Students in Informatics, Statistics, and Econometrics on the Basis of ESS Research Methodology

Venelin Boshnakov[1], Valentin Goev[2]

University of National and World Economy – Sofia, Bulgaria[1,2]

Abstract

The paper informs about the range of alternatives for utilization of various methodological tools implemented at a major European Research Infrastructure – namely the European Social Survey programme – to provide modern teaching resources for students in applied informatics, statistics, and econometrics. A contemporary study of database management, statistical analysis of data, and reporting of research results at all university levels (undergraduate, graduate, and postgraduate) can be substantially boosted by incorporation of particular components of the methodological grounds of the European Social Survey programme. A specific focus of the paper is put on the variety of options to provide valuable resources for teaching empirical analysis courses that utilize cross-sectional data. Particularly important for the students are the options for deriving cross-country comparative analyses of socio-economic topics using large sample representative data for numerous European countries. However, before doing any analytical work the student has to gain substantial knowledge about database management (with appropriate examples form the area of social surveys), data processing, and preparation of cross sectional data for analysis using specialized statistical or econometric software. The paper underlines the major advantageous alternatives for utilizing the methodological knowledge achieved in the framework of the European Social Survey Programme for the goals of providing real-world innovative study resources at the required high methodological standards level.

Keywords: *European Social Survey, survey design, survey data analysis, teaching empirical analysis.*

1. Introduction

Design and implementation of practice-oriented teaching in the areas of Informatics, Statistics, and Econometrics is frequently based on student resources that are specifically developed for this purpose. The content of these resources is generated in order to provide the required variety of examples related to the analytical situations defined in the respective courses curricula. Generally, this approach provides satisfactory educational results due to the targeted illustrations that facilitate the achievement of particular study goals. The challenge here is the incorporation of real-world data and methodologies into the teaching process – an approach that can achieve not only the teaching goals. It can provide direct evidences to the students about the practical orientation of the courses curricula – they will be assured that their studies will provide them with expertize required by the employers and directly implementable in the profession of interest.

Here we assert that a substantial variety of modern teaching resources and methodological tools for students in Informatics, Statistics, and Econometrics can be provided by the European Social Survey (ESS) Program – one of the main European research infrastructures supporting social research in Europe and worldwide. The headquarters of ESS are located in City University (London) and up to year 2018 nine waves of social surveys (biannual) have been accomplished since the first wave conducted in 2002 [1]. A significant advantage of ESS methodological and data resources originates from the fact that it is designed and operated as a multi-country standardized social survey of personal opinions and attitudes of European citizens toward a variety of social, political, and welfare issues [2]. On a regular basis, ESS is nationally implemented in order to generate cross-country datasets that are comparative not only between countries but also in the time span. It is our contention here that contemporary studies in survey design, survey sampling, data collection and processing, database management, statistical analysis of data, and reporting of research results at all university levels (undergraduate, graduate, and postgraduate) can be substantially boosted by incorporation of particular components of ESS methodological knowledge and data resources. Moreover, ESS resources, including cross-country datasets can be downloaded from ESS website for free and analysed using popular statistical software systems.

2. ESS resources in support of teaching empirical research

ESS provides a diversity of valuable resources that can facilitate the educational process within various courses in empirical analysis, particularly based on cross-sectional data – as far as ESS is performed by a repeated sampling procedure. This can naturally be put in the context of the search of best practices and innovative approaches that stimulate the interest of the students in such study field [4]. However, the standardized methodological setting provides valuable opportunities to derive empirical results not only about a particular country but also in a comparative context, e.g. about a selection of European countries (in case that each one within this set has participated in a particular wave of ESS). Table 1 contains a selection of typical applications of statistical methods for empirical analyses that can be implemented using ESS data and the relevant knowledge and skills to be developed by students.

Table 1. Statistical analyses based on ESS data resources

Methods	**ESS data**	**Knowledge and skills for:**
Descriptive statistics	*Detailed set of personal and household characteristics: Demographic, Social, Economic*	*Statistical processing and presenting summary data by one, two or more variables (tabular and graphic forms)*
Hypothesis testing: non-parametric methods	Numerous variables with nominal or ordinal measurement scales related to: Human values; Politics; Media and social trust; Subjective wellbeing, Health, etc.	Testing for distribution form (normal) Testing with crosstabs (Chi-square) Testing with independent samples (Mann-Whitney, Kruskal-Wallis, Wilcoxon, etc.) or paired samples
Hypothesis testing: parametric methods	Numerous variables with quantitative measurement scales related to: Demography (age, children); Economic activity (income, labour market participation);	Testing about population means Testing with independent samples (t-test; ANOVA) Testing with paired samples Testing about proportions / shares
Classical regression	Simple and Multiple / Linear and Nonlinear regression models. Quantitative dependent variable	Building and estimating regression models (mix of nominal, ordinal, and scale IVs). Interpretation of results
Correlation	Measurement of correlations between any combination of nominal, ordinal, and interval/ratio scale variables	Estimation, hypothesis testing, and interpretation of: classical Pearson product-moment correlation; variance ratio (eta); non-parametric correlations (Spearman, Kendal, etc.)
Regression with nominal or ordinal dependent variable	Models of binary DV (logit / probit). Multinomial and ordinal logistic regression. Ordinal logit / probit models	Building and estimating regression models. Variety of utilization (mix) of nominal, ordinal, and scale IVs. Interpretation of results
Multivariate analysis: Factor	Principal component analysis. Factor analysis models	Performing a factor analysis procedure. Derivation of latent variables. Interpretation of results
Multivariate analysis: Discriminant	Binary (two-group) and K-group discriminant functions	Performing a discriminant analysis procedure. Standardized discriminant function, Z-scores, testing of hypotheses. Classification table
Multivariate analysis: Clustering	Non-hierarchical and hierarchical cluster analysis (e.g. deriving clusters of respondents with specific profiles)	Performing a cluster analysis procedure. Derivation of clusters. Interpretation of results (cluster centres, degree of homogeneity)
Multivariate analysis: Classification trees	Classification and Regression Trees (CHAID, CART). Decision Tree algorithms	Performing a CHAIND or CART procedure. Solving classification predictive modeling problems. Segmentation studies

The experience gained from numerous utilizations of ESS data resources in teaching statistical or econometric analysis reveals the potential of these resources to provide real-world empirical data for

effective teaching purposes. A study focused on frequently met negative perceptions among students about statistical methodology outlines three major areas for enhancing the effectiveness of the learning process [5]:

1) fostering classroom environment to become “dreadless”, friendly, and encouraging;
2) explaining the statistical concepts in plain terms;
3) developing practical skills relevant to the area of specialization.

In this line of reasoning, learning from the social research experience appeared to be soundly inspiring for the students. Solving different empirical problems has been identified as one of the major empirically supported strategies for teaching disciplines related to applied statistics: repetition, immediate feedback, and use of original data [6]. Putting the exercises in a particular research context and testing hypotheses by real survey data effectively contributed to the building of technical, analytical, and reporting skills by the students. The knowledge and competences for performing empirical analysis using genuine data proved to be a solid fundament for the students valuable for their eventual professional development as social, economic, business analysts.

3. Methodological resources of ESS

Teaching in applied informatics and statistics is substantially facilitated by the utilization of the methodological achievements of the ESS programme. The general methodology for implementation of a full cycle of an ESS wave is presented on figure 1 (more information can be obtained from [2]). The main areas of ESS methodology that proved to be particularly valuable in the aforementioned study fields are:

- developing the sampling design;
- performing sample selection procedures;
- planning and executing the procedures for data collection;
- preparing software applications and performing primary data processing;
- creating standardised datasets ready for analysis;
- data archiving and linking to online extraction software.

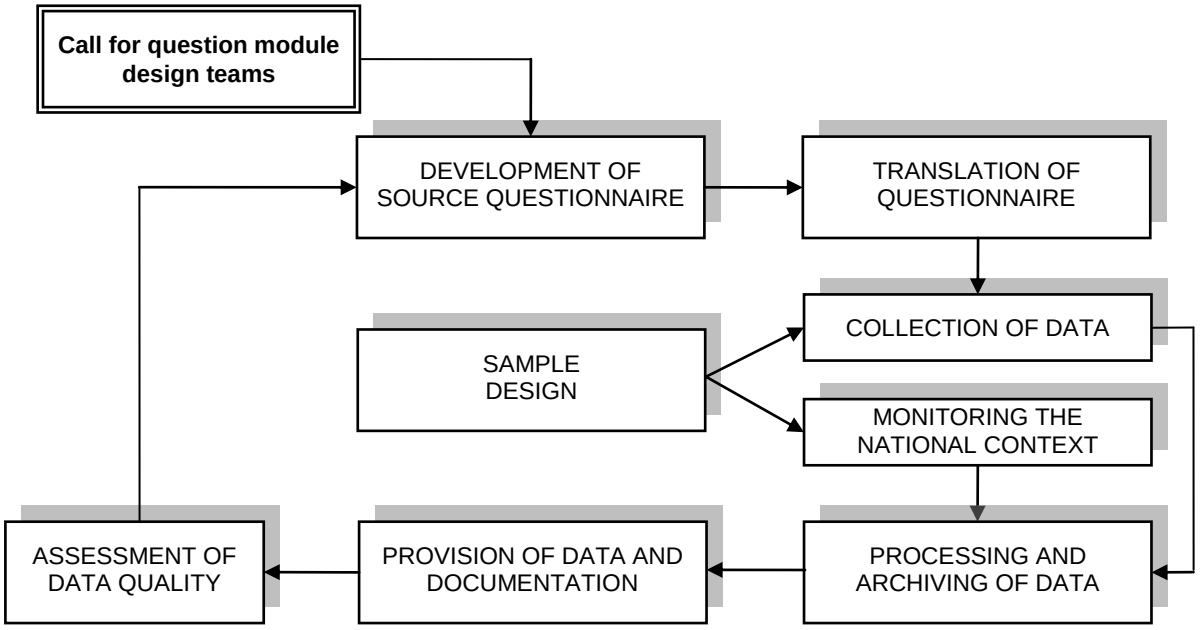

Figure 1. ESS methodological frame
Source: ESS Methodology Overview [2]

3.1. Sampling

Teaching of sampling methodology in specialised courses of applied statistics can be substantially facilitated by utilization of ESS sampling technology. It generally includes the design and implementation of feasible sampling plan in any participating country. Sampling designs and procedures in ESS are developed under several major principles extensively defined in the methodological documents (e.g. ESS Survey Specifications [3]).

- Samples are designed to be representative of all persons aged 15 and over (no upper age limit) that are residents in private households – regardless of their nationality, citizenship or language;
- Individuals need to be selected by strict random probability methods at every stage.
- Sampling frames of individuals, households and addresses may be used.
- Any country must aim for an "effective achieved sample size" of minimum 1,500 (or 800 in countries with populations of less than 2 million, after discounting for design effects).
- Quota sampling is not permitted at any stage.
- Substitution of non-responding households or individuals (e.g. refusals, non-contacts, or ineligibles) is not permitted at any stage.

Each National team performing the ESS must generate a sample design that sticks to ESS sampling principles and in the same time is appropriate for implementation in the respective country. Detailed information about, for example, stratification and clustering of ESS data, is provided to the users in the so called "Sample Design Data File". The experience and practices in different countries provide a variety of examples and real-life contexts in support of the effective teaching of stratified cluster sampling designs.

3.2. Data collection

Studying ESS standards and practices in data collection procedures additionally facilitates the development of knowledge in this professional area. Typically, ESS data is collected through face-to-face interviews conducted by the implementation of the CAPI recording method. By exploring the ESS specifications in this respect the student can recognize the approach for ensuring accuracy of data collection procedures (such that provide strict data comparability across countries). The major standards of ESS data collection include:

- response rate target of at least 70%;
- non-contact rate target of 3% maximum;
- fieldwork period of at least 1 month (within the 4 months September-December);
- detailed briefing of interviewers in face-to-face sessions;
- restriction on the workload of the individual interviewer (maximum 48 sample units gross);
- interviewer call schedule (minimum 4 contact attempts among which at least 1 in the evening and 1 at the weekend);
- specific contact forms (to record and document information about the fieldwork processes);
- quality control back-checks on completed interviews and ineligible cases (minimum 10%);
- organization of a close monitoring of the progress of fieldwork.

3.2. Data processing and archiving

The structure and principles of development, update, and maintenance of the ESS Data Archive provide indispensable resource for studying applied database management and statistical data processing. ESS standards in this respect provide overall guiding principles that produce harmonised data files. These files must be not only user-friendly but should reflect the original reliability and quality of the data. A major document supporting these processes is the so called "ESS Data Protocol" which is a key specification document concerning the structure of the data files. It provides definitions that must assure cross-national uniformity of different country data (as far as it provides specifications for the coding of variable values /answer categories/, classifications, and other important issues of the production and delivery of data files). Since ESS practices a strictly probability-based sampling, its methodology provides reliable weighting tools in order to facilitate the empirical analyses. Students can learn practical knowledge and skills in the topics of complex samples and related weighting variables (design weights: dweight; post-stratification weights: pspwght).

5. Conclusions

This paper outlined the major directions for deployment of ESS resources – both data and methodological – in the teaching of specialized courses related to applied statistics and database management. The main proposition here is that using real-world data from social research infrastructures will prompt the students – activating their curiosity and interest – to develop knowledge and skills for designing surveys, planning and organizing data collection and archiving procedures, and implementing a variety of empirical analyses. ESS resources provide the instructor with an rich arsenal of practical tools to formulate educational goals, assign individual or group course works, and other study components in the abovementioned knowledge areas. The data resources and detailed

methodological instruments of ESS cover the major stages of the empirical research processes in an integrated way. Incorporating these resources into the teaching of academic disciplines can contribute effectively to the development of professional competences of the students at all levels of the tertiary education.

Acknowledgment

The authors of this paper acknowledge the support obtained from “ESS for Bulgaria” Research Infrastructure developed by ESS Consortium in the framework of the National Road Map for Research Infrastructures 2017-2023 of the Republic of Bulgaria through the Ministry of Education and Science.

References

[1] ESS-a, “The History of the ESS ERIC” (www.europeansocialsurvey.org/about/history.html), 2019.
[2] ESS-b, “ESS Methodology” (https://www.europeansocialsurvey.org/methodology/), 2019.
[3] ESS-c, “Round 9 Survey Specification for ESS ERIC Member, Observer and Guest Countries”, V.3 /24 July 2018/, European Research Infrastructure Consortium, London, 2018.
[4] Hulsizer, M. R., Woolf, L. M., “A Guide to Teaching Statistics: Innovations and Best Practices”, Wiley-Blackwell, 2008.
[5] Lalayants, M., “Overcoming graduate students' negative perceptions of Statistics”, Journal of Teaching in Social Work, 2012, Vol. 32 (4), pp. 356-375.
[6] Marson, S. M., “Three empirical strategies for teaching statistics”, Journal of Teaching in Social Work, 2008, Vol. 27 (3-4), pp. 199-213.