Economic Statistics or Statistical Methods in Economics?[1]

Richard Hindls[2] | *University of Economics, Prague, Czech Republic*
Stanislava Hronová[3] | *University of Economics, Prague, Czech Republic*

Abstract

In teaching statistics to economists, it often happens that insufficient distinction is drawn between statistical methods applied on economic data on the one hand, and economic statistics as a special discipline with its own theoretic basis, fundamental notions and a specific concept of indices on the other hand. The authors endeavour to point out pitfalls of this insufficient distinction and introduce didactic ways to resolve this problem.

Keywords

Statistical methods in economics, economic and social statistics, teaching statistics

JEL code

E01, C02, N01

INTRODUCTION

Teaching statistics at universities of economic orientation is subject to a number of popular misconceptions, whose consequences are harmful to both statistics and economics. The worst of it is that students are discouraged by them. It is no secret that, for quite a large proportion of students of economics, mathematics and statistics pose an arduous challenge; they try to evade these disciplines as long as they can in the hope that later, when they work in firms and corporations, they will not need them too badly.[4] What a mistake this is!

For the purposes of the present paper, let us leave aside teaching mathematics and focus on teaching statistics in economics. Our experience is based on forty years of practical teaching statistical methods used in both economics and economic and social statistics at Czech and foreign universities.[5] Although the last two mentioned general areas of understanding statistics in economic fields (economic statistics

[1] This article was written thanks to support from the means of the Institutional Support to Long-Term Conceptual Development of the Research Organisation of the Faculty of Informatics and Statistics of the University of Economics in Prague.
[2] Faculty of Informatics and Statistics, Department of Statistics and Probability, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: hindls@vse.cz.
[3] Faculty of Informatics and Statistics, Department of Economic Statistics, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: hronova@vse.cz.
[4] This is for example discussed in Gatusso (2011).
[5] Statistical methods courses in the faculties of economics are generally two-semester long (lectures + exercises). Economic statistics course is usually one semester long.

and statistical methods in economics) are quite dissimilar to each other, they are often not sufficiently distinguished by either teachers or students of economics at universities.

At the very root of this lack of distinction between statistical methods used in economics on the one hand and economic and social statistics on the other hand is the fact that among statisticians we can find many more of those who specialise in applications of various statistical methods than those who are professionally and primarily interested in economic statistics, national accounts, etc. As a logical consequence, teaching statistics for economists is often reduced to an inorganic combination of learning certain statistical methods with getting some economic data, and applying (more or less mechanically) the former on the latter.

A tempting support of this approach stems from the wide range of statistical software, so that you simply "input" the collected data into suitable software procedures and then just wait for the outcome. And sometimes you do not need even take such a complicated course: instead, simply use Microsoft Excel. At many universities or faculties, that kind of exercise is taken for mastering statistics in economics. At these cases, it is the teachers' fault, because it is unfortunately true that some professors who teach of statistics for economists do not have enough experience from business nor sufficient contacts with the business community or the providers of macroeconomic data.[6]

We can see many instances in which lecturers perseveringly think up pseudo-examples from the national economy, trying to make an impression that real quantitative-economic analysis of certain phenomena is thus achieved. In fact, such pseudo-applications are quite a long way from a realistic economic analysis and, in the long run, discourage both statisticians from economics and economists from statistics. To the detriment of both.

In other words, the relationship between statistical methods (or general quantitative methods in the broad sense) and economics is far more complex in the real world (as opposed to the isolated and virtual realm of pseudo-examples). On the whole, the above-described trivial approach to teaching is shallow and insufficient and leads to irreversible didactic errors. However, problems lie not only in the approach to the application of statistical methods in economics, but also in the fact that such applications are mistakenly considered as economic statistics. There are several reasons – let us go through some of them.

1 HAZARDS HIDDEN IN ECONOMIC DATA

Little attention is given to the quality of economic data. Such data are often easily accessible at websites of various institutions. When applying statistical methods, few people carefully study the potential methodological pitfalls hidden in the data, to what extent they are affected by the methods of data collection used in corporations and state administration, the organisational structures and their changes, or the pricing, taxation, exchange-course and other aspects. The procedure is then completed by an accurate and quickly available calculation of the desired characteristics. A more detailed study, however, shows the lack of any informative value of such characteristics; and their deeper analysis by both theoretically and practically oriented economists leads to disappointment and scepticism. The ultimate result is distrust in the potential of statistics in economics felt by both statisticians and economists.

A bridge between statistics and economics is thus destroyed, or even not built at all, from the very beginning of possible cooperation.

The reader might object that nothing like that should happen and that everybody is aware of such risks. However, the opposite is true. This fact is proved by many years of our experience from various universities in the Czech Republic and abroad and from contacts with the economic practice. It is implied by a strong preference teachers put on the formal side of statistical methods, suppressing the parallel

[6] Compare for example Gelman, Loken (2012) or Groth (2013).

need for critical assessment of the economic data to be analysed with the aid of this or that statistical method.

Let us mention a small example from time series analysis: students learn quite a few modelling methods, such as time series trends (the deterministic approach, adaptive methods, the Box-Jenkins methodology, etc.), equip themselves with the corresponding software, get some data and carry out their calculations.

How much information have they obtained about problems that may occur in the respective data environment of the time series in question, and what should their attention be focused on? We have in mind, for example, the issues of spatial, factual and temporal comparability in the data (constant and current prices, methodology of collection, calendar variations, etc.). What about the length of the time series? We keep telling our students that statistics deals with mass phenomena – the more observations the better. It is even the categorical imperative when using, say, the Box-Jenkins methodology. On the other hand, the longer the time series the better for the methods but the higher contamination, often fatal, of the data (due to the changing methodologies, factual discrepancies, pricing recalculation coefficients, etc.). If a trend analysis is mechanically applied to such a time series, total doom follows. Do students know such facts and are they persuaded about them within their study? Well, often they do not and are not.

A solution is: knowledge of economics by teachers of statistics, as well as knowledge of statistics by teachers of economics, should be elevated. More time should be given to quality of data, with less preference given to describing mere methods and tools. This is the only way of reducing the risk of unsuccessful applications of statistics in economics, and also reducing the risk of economists' relying on feeble verbal declarations they are unable to support with relevant numerical illustrations, i.e., to use particular facts in strengthening their argumentations.

2 SIMPLIFIED APPLICATION OF STATISTICAL METHODS – A SOURCE OF DUBIOUS INTERPRETATIONS

Statistical methods and tools applied in economics are often taught in their simplified versions, presenting just basic principles, properties, and utilisation. Why not – after all, most students at economic universities are not specialists in statistics and in the future they are going to become practical users of statistics; very often of just a very restricted range of statistical techniques. Hence this approach is correct. But only until real economic and social situations are encountered. Nothing was explained incorrectly in the teaching, all aspects were given adequate attention, but in confrontations with real situations something seems not to work.

Let us present another trivial story, this time of a correlation coefficient. Students in non-statistical fields (i.e., non-statisticians) are honestly explained what the correlation coefficient is, what its uses are, what the regression concept of its origin is, what the coefficient of determination is, etc. What students actually remember from such explanations is a simple interpretation of the resulting value of the correlation coefficient (hardly anyone deals with the calculations nowadays, a simple MS Excel procedure is sufficient for getting the value): if it holds $|r_{yx}| \geq 0.7$ for its absolute value, we will say that the dependency is strong.

So far so good. But in social sciences a characteristic feature is that they are based on an objectified outcome of people's subjective efforts and motivations; human behaviour is by far not linear and the resulting data may, due to considerable differences in people's characteristics and abilities, have a high level of variability and a low level of consistency. In other words, real data will, to a great extent, be contaminated with subjective features of human behaviour. As an example we can mention an outcome of an opinion poll or consumer behaviour in marketing. Do we really encounter in these – rather usual – applications values of dependency leading to correlation coefficient values in the area of $|r_{yx}| \geq 0.7$? We dare say that it is practically never the case. Its value is much more likely to achieve something like

r_{yx} = 0.3. A student not specialised in statistics, equipped with formalism and lacking real knowledge of social data, will conclude that the dependency between the analysed phenomena is weak. But regarding the data quality, even 0.3 may be quite enough for strong dependency. Students, however, rarely learn about such a conclusion, unless they go into a deeper analysis of the underlying problem and, possibly, employ some methods of qualitative survey, as usual in marketing, to name one example. If they do, it may turn out that even $|r_{yx}| \leq 0.3$ is not such a small value in the given situation. We usually do not encumber non-statistician students with such explanations, leaving them at the mercy of simplified techniques for interpretations of statistical results.

Teaching statistical methods for economists is limited by the relatively small number of hours (usually 4 hours per week for 2 semesters), which does not allow to introduce to students all application possibilities of statistics in depth. This leads on the side of the teachers to superficial interpretation of the nature and conditions of applicability of each method, and on the side of the students to their misunderstanding and consequently to their improper use of a simplified interpretations. The solution to this situation can only to reduce the number of topics and methods presented in the basic course of statistics for economists. This will allow to deepen the explanation of selected methods with the emphasis on the conditions of their applicability and interpretation of the findings.[7]

3 ECONOMIC AND SOCIAL STATISTICS

What we said above was concerned with the use of statistical methods in economics. But there is also economic and social statistics, as a special and quite large part of statistics, which requires a different approach. It is the one that is perhaps most neglected by economists. A similar observation is valid for national accounts, which should be taught to every student of economics to provide them with a plastic view of what is globally going on in both national and worldwide economics.

Both these disciplines "sit on the fence" between statistics and economics. In order to be able to cope with them, students must have good knowledge of economics (both theoretical and practical); and it is impossible without a good command of quantitative techniques and the ability to interpret economic data (not only form the viewpoint of statistics but even that of accounting). Let us mention a small example here – the Keynesian economic theory, which is the economic basis of the national accounts. And vice versa: the national accounts as a statistical model of the national economy lead students to understanding mutual relationships between major aggregates and, more generally, how these aggregates work. Moreover: the national accounts cannot be understood without explanations of the fundamentals of statistics because, without those, students cannot get a proper insight into the data provided by the national accounts as a system of economic information and cannot process such data. The circle is thus closed and we are back at the beginning.

The story of the economic and social statistics is quite similar. Economic theory often employs notions such as inflation, unemployment (or employment) production and productivity, etc. Of course. But are we able to quantify and estimate such notions, or are we to rely on mere theoretical meditation?[8]

Therefore, the economic and social statistics is a special discipline of statistics (similar to testing hypotheses or regression and correlation analysis). It has its theoretical basis, a system of fundamental notions, methods and tools and, above all, a specific concept of indices viewed as variables, including definitions of their contents.

---

[7] Similar consideration can be found for example in Hernandez (2006) or Brown, David (2010).

[8] Please note that we have deliberately avoided the term of "measuring" any of these economic variables. Not much can actually be measured in economics; hence we leave measurements to physics, anatomy and similar, more "measurable" fields of human knowledge and endeavour.

In no case can it be reduced to mechanical applications of selected statistical methods (as mentioned above – e.g., time series analysis, statistical inference, regression and correlation analysis, descriptive statistics, and multidimensional statistical methods) to real economic data, and such a reduced version must not be passed off as economic statistics, even though we sometimes see exactly that within teaching at economically oriented universities.

It is the specific concept of the indices and definitions of their contents that make up the crucial framework for the economic and social statistics. Going back to the above-mentioned relationships between theoretical notions of the science called economics on the one hand and possibilities of their relevant quantification on the other hand, we necessarily come to a notion called adequation gap. This adequation gap lies in the core of the matter: many notions utilised within theoretical economics cannot simply be fully quantified and a certain quantitative approximation to such notions must be accepted.

This "approximation" is thus a necessary trade-off between the theoretical economics' needs for quantification of its notions and our practical ability to quantify them as desired. This trade-off between "possible and required" is the structural content of the above-mentioned adequation gap. To provide a tangible example of this gap, we can mention inflation as a theoretical economic category and the index of consumer prices as a quantification of this theoretical notion.

These considerations are not, however explained to students of economy with sufficient emphasis; and if it comes to the worst, they are not mentioned at all. In consequence, students are at a loss when looking for the "proper" statistical data, they do not understand the data they get, use them in inadequate ways and, finally, interpret the results incorrectly.

CONCLUSIONS

Teaching statistics to economists, or more generally at universities and faculties with economic orientation, more attention should be given to interrelation between statistics and economics. This attention should, above all, be demonstrated by using real data from the economy (whether national or corporate) with a strong emphasis on understanding the substance of such data. Similarly, teaching economics should be more attentive to quantifying theoretical notions – we can hardly prove what we cannot quantify, having to rely on mere hypothetical claims which may later – in the light of real data – turn out to be disputable and unprovable, or even doubtful. By no means should applications of statistical methods in economics be confused with economic and social statistics; unfortunately, this is often the case within the teaching process.

A way to solve the problems outlined in the article is certainly not only to reduce the number of topics contained in the basic course of statistics for economists and thus to allow more profound explanation of a smaller number of methods, but also the inclusion of the basic course of economic and social statistics in the mandatory curriculum of students of economics. A separate course of economic and social statistics will enable students to understand the nature and characteristics of statistical data and allow them to avoid some errors in the application of the methods and the interpretation of the conclusions of the analysis. To increase the quality in perception of statistics for students of economics would also undoubtedly contribute deeper economic and economic-statistical knowledge of statistics teachers.

References

BROWN, J., DAVID, I. Teaching critical thinking to first year university students. In: *Data and context: towards an evidence-based society. Proceedings of the 8th International Conference on Teaching Statistics*, ICOTS-8, 2010.

GATTUSO, L. L'enseignement de la statistique: où, quand, comment, pourquoi pas? *Statistique et Enseignement*, Vol. 2, No 1, 2011. ISSN 2108-6745.

GELMAN, A., LOKEN, E. Statisticians: When we teach, we don't practise what we preach. *Ethics and Statistics*, 2012, Vol. 25, pp. 47–48.

GIOVANNINI, E. *Understanding economic statistics: an OECD perspective*. Paris: OECD, 2008. ISBN 978-92-64-03312-2.
GROTH, R. E. A day in the life of a statistical knowledge for teaching course. *Teaching Statistics*, 2013, Vol. 35, No 1, pp. 37–42.
HERNANDEZ, O. Teaching statistics to undergraduate students of the social sciences. In: *Working Cooperatively in Statistics Education. Proceedings of the 7th International Conference on Teaching Statistics*, ICOTS-7, 2006.
HINDLS, R., HRONOVÁ, S. Quarte conférences critiques qui peuvent détourner les non-statisticiens de la statistique. *Proceedings of the 55th Session of the International Statistical Institute*, Sydney, 5.4.2005–12.4.2005. ISBN 1-877040-28-2.
HINDLS, R., HRONOVÁ, S. How teaching discourages non-statisticians from statistics (In Czech: Jak výuka odrazuje nestatistiky od statistiky). *Statistika*, 2005, Vol. 42, No. 2, pp. 168–172. ISSN 0322-788X.
HINDLS, R., HRONOVÁ, S., ŠTĚPÁN, J. University Statistical Education from the Unifying Europe's Point of View. *Statistics in Transition*, Warszawa 2001, Vol. 5, No. 2, pp. 249–257, ISSN 1234-7655.