

Dear Dr. Kostenko:

Jan. 14, 2013

I appreciated your thoughtful review of my book “Interpreting Economic and Social statistics – A Foundation of Descriptive Statistics”. Understandably you limited the review to the Forecasting and time series chapters although the concepts basic for the Foundation in the subtitle of the book are developed in chapters 1 – 4 to which you may have paid less attention. Your concluding remarks that “most statisticians ... would find the book somewhat pedestrian and simplistic” is an astute observation: indeed the explanations in the text proceed step by step patiently walking on the ground of socio-economic reality, keeping in close touch with it. Nor are there any spectacular mathematical high-wire-acts as the book is intended for a wide range of users of socio-economic data, people with different backgrounds not only forecasters or professors teaching statistics. Unaccustomed new concepts particularly in the chapters 1 – 4 had to be presented as simple as possible<sup>1</sup>. This contrasts with the current tendency to upstage statistical research with esoteric terminology<sup>2</sup> for academicians talking to each other in a rarified intellectual environment, but hardly relevant to data-users outside of academe.<sup>3</sup> Though the book appears to be “pedestrian and simplistic” it is, I believe, not irrelevant, and hopefully will help to promote a change of the academic “climate”. You criticized the failure to make explicit the parallel with A.S.C. Ehrenberg’s quoted observation that “*Statistics courses are largely irrelevant – not just boring or technically difficult, but irrelevant*”. Although there appears to be some parallels, our thinking was not so close. Concerning the teaching of statistics in American universities I would not go along with Ehrenberg’s statement, as expressed in Figure 1.2 on page 2 of my book. Such courses for economists and social scientists are aimed at educating professional statisticians and natural scientists, seemingly unawares that most of the students in these courses do not intend to become professional statisticians or bio-scientists but economists, business managers and administrators, etc. For the most part these courses offer watered-down versions of mathematical statistics and are relevant only for a minority among their students. Because of that partial overlap with the needs of socio-economic statistics these courses are not the complete waste implied in Ehrenberg’s statement. I intended my book to help change this situation by re-orienting socio-economic statistics to recognize its particular nature, different from statistics in the sciences.

Toward the end of your review you mention that “*It is this subtitle alone - A Foundation of Descriptive Statistics - which is responsible for the present review having been written*”. It is precisely these ideas about the foundation of socio-economic statistics that were not even part of the vocabulary when discussing socio-economic statistics during these last decades. The concepts describing the nature of and the roles played by “economic phenomena”, the “real-life-objects” and their statistical counterparts, the “statistical counting units”, and then their “aggregates<sup>4</sup>” developed in chapters 1-3, and ratios of these, in chapter 4, were hardly treated in the journal literature. It was as if these concepts did not belong to socio-economic statistics. Yet when developing a theoretical framework of socio-economic statistics from these basic facts a different view of statistics will emerge, glimpses of which are developed in this book.

The word ‘foundation’ in the subtitle of the book, in analogy to the foundation of a building, indicated that it has to be strong, connecting the structure to be built on it with the ground on which the building is to be erected and able to support the weight of the building properly speaking which will be built on it. Besides being robust and strong, foundations also are simple. Analogously the foundation of socio-economic statistics is supposed to be anchored in the socio-economic reality “out there”, forming the solid conceptual underpinning of a more complex theoretical structure. In other words, its logic should be strong and able to uphold later theoretical constructs, regardless of how simple that foundation may be.

You rightly noted that data obsolescence had been considered in the form of ‘exponential smoothing’ and ‘exponential forecasting’ as early as 1950. I had used exponential smoothing and forecasting since then, but found that the coefficients of the formulas were arbitrary and not connected with the underlying historic facts. Lacking was an awareness of the need to tie the weights of the smoothing and forecasting formulas to the variety of socio-economic events causing data to become obsolete. Missing was a discussion of what constitutes ‘obsolescence’ and its effects. I also found it difficult to make sense of the smoothed time series – do the smoothed data represent what would have happened to the data if those actual socio-economic forces had been absent? Or different? Does smoothing reveal something valuable in society that would be overlooked otherwise? A strong argument for the importance to take into account the actual historical situation is evident in

the abrupt breaks in the time series of 'mad cow disease' (on p. 93). Using a smoothing formula just did not make sense because of the immediate, strong impact of the sequence of laws enacted at consecutive points in time to curb that disease. Consider also that exponential smoothing formulas – indeed any form of smoothing – is of no help in locating the point in time beyond which the data of a time series should not be used at all in a forecast.

You recognized the importance of linking the data of a time series to the underlying historic events but prefer this to be an exception because of the likely costs. You stated “*clearly, this is only possible for very important time series*”. I see no reason for such a restriction; time series are produced because of the need to learn things about an economic or social situation that only the statistical overview produced by the time series can reveal. This kind of analysis obliges statisticians to get involved with the underlying reality of the social or economic forces presented in a time series instead of relying only on some mathematical formula divorced from socio-economic reality. That kind of analysis should be done before any further attempts are made to explore regularities and other mathematical properties of a time series<sup>5</sup>. The suggestion by Gilchrist “...an examination of past data to select a number of reasonable models. These are then examined from the point of view of goodness of fit not only to the whole data but also to smaller regions of the data” is a proof of what I criticized, namely limiting that examination to the algebraic numbers, studying only these e.g. for regularities or trends instead of “going outside of the data, in fact, outside of statistics, to explore what those economic and social facts are from which those data were produced’. Too much work? Too costly? But that is how every time series ought to be studied if any use is actually to be made of such an analysis instead of taking the comfortable way by ‘staying indoors’ and limiting the analysis just to manipulating the data with algebra.

With regard to the many typos and deplorable misprints I should mention that the production of this book was “outsourced” by the publisher to India. Although I was communicating daily with some English-named Indian employees during the publishing phase, my role was limited to paying attention to graphs, formulas and the clarity of expression. The typographical detail supposedly was taken care of by the publisher’s proofreaders.

In response to your complaint of “*insufficient linkages to the relevant literature*” I should point out that there was no discussion in the literature of the epistemology of social and economic ‘phenomena’ as being made perceptible through “real life objects”, their statistical counterparts, the “statistical counting units” that act like iron filings that allow the phenomenon ‘magnetism’ to become visible. There was and continues to be hardly any literature relating to the ideas in the book. In most of these matters I had to “fly solo”. As you obviously are a well-read rising scholar I challenge you to find publications that clarified these basic concepts along the lines discussed in chapters 1 – 4. I did quote Kendall because he questioned the relevance of the teaching of statistics for social scientists (p.19, 46, 51, 60), expressed doubts about the value of mathematical decomposition of time series (p. 83, 84) and on misuses of statistical inference (p.190, 204). Although he advocated change in teaching statistics he did not make concrete proposals why, what and how to change. I quoted Ehrenberg in chapter 3 because he misleadingly characterized aggregation as “data reduction” (p.6, 16). Kendall’s and Ehrenberg’s ideas of ‘change’ were different than the ideas for change I am advocating, particularly regarding their reluctance to sideline probability. This reduced importance of probability for socio-economic statistics is expressed in my book by postponing its discussion to chapter 8, as an expression of its subsidiary, less important role. I consider this as sufficient reason not to have referred to those publications of these authors’ that you thought were supporting my position.

I did appreciate your thoughtful review and hope that you may eventually join “*the author’s belief that ‘social and economic statistics, though numeric, is essentially quantified history of society, not a branch of mathematics’.*”

Othmar W. Winkler  
Professor emeritus  
Georgetown University

---

<sup>1</sup> As an example, Einstein summarized his research into the basics of the physical world in a few famously simple formulas!

<sup>2</sup> Take for example the announcement by the Washington Statistical Society, e-mail of November 28, 2012, of the lecture with the title “*Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations*” at the ‘American University’ in Washington DC.

---

<sup>3</sup> The development of the ideas presented in my book are traced by the profound Chinese insight quoted in the heading to chapter 5 as amplified by a personal update:

*“Thirty years ago (actually 70 years), before old monk (me, though married 58 years, and seven children) had studied Zen (statistics) he saw the mountains as mountains (economic data), waters as waters (social data). Later he came to know a good master (at the University of Vienna) and was first initiated into Zen (Statistical theory) he no longer saw mountains as mountains or waters as waters (through the optic of the theory of statistics). Now (in 2013, at age 90) he had got a resting place (retirement from Georgetown University), he again (finally) sees that mountains are only mountains and waters only waters (recognizing the true historic nature of socio-economic data as quite different from and in contrast to the essentially timeless, un-historic measurements in the sciences, that are treated as just algebraic numbers)”*

<sup>4</sup> The academic discussion of e.g. ‘linear aggregation’ had a mathematical approach, different from the epistemological approach in chapter 4

<sup>5</sup> The title of my book “Interpreting Economic and Social **Data**...” really should be understood as: “Interpreting **Economic and Social Situations with the help of Data**”, shifting attention from the statistical-mathematical detail to the socio-economic historic detail of the situation presented by the time series.