How Economic and Social Statistics became the Stepchildren of the Profession

Othmar W. Winkler
Professor emeritus
Georgetown University
winklero@georgetown.edu

"The past is a foreign country: they do things differently there"[1]

1. Introduction. A review of statistical journals, such as 'Amstat News', 'The American Statistician', JASA 'The Journal of the American Statistical Association'as well as a few European statistical journals, like 'Advances in Statistical Analysis' (AStA)[2] or 'Jahrbücher für Nationalökonomie und Statistik' showed that nearly all articles, book reviews, etc. deal with bio statistics, medical statistics, statistics in the sciences. There is hardly a mention of economic or social statistics except for an occasional article on a mathematical model for the economy. This leaves the impression that statistics is about measurements and measurement errors in the sciences, statistical inference and testing mathematical models. Occasionally socio economic statistical topics are presented with a sense that these topics are of minor interest, the "stepchildren of the profession". The situation has consequences for the requirements for statistical literacy in the social sciences and statistics in the natural sciences; they differ in important ways. Here an example of the latter:

"The years since 1964 have been characterized by the knowledge that there are fast Fourier transform algorithms ..… previously the data analyzed consisted almost totally of discrete or continuous real valued Time Series. Now the joint analysis of many series, such as the 625 records by the large Aperture Seismic Array in Montana has become common. (p.274) A variety of alternate estimates of the power spectrum have been receiving practical attention lately in the high-resolution estimate and the autoregressive estimation… the maximum entropy estimate. The usual estimate of a power spectrum is a quadratic function of the observations when the data have not been pre-filtered …. Another area… is the fitting of finite parameter models, especially when…an understandable non-stationarity is present" (p. 275)[3]

Looking at this quotation "how statistically literate do you feel?"You might answer 'yes' if you were a mathematical statistician. A social scientist's response may be less positive. Although the public in general has become more educated during the last centuries, statistics has become increasingly more 'scientific' that is, mathematically more demanding placing ever higher demands on the literacy of the user of statistical material.

[1] L.B. Hartley, English novelist, quoted in the article by Robert J. Samuelson titled "The economy's uncharted path", in 'The Washington Post', Monday, February 6, 2012. This appears to be an appropriate title for the printed version of my talk at JSM San Diego, 2012. Given the focus of this session on statistical literacy, this approach to the history of socio-economic statistics can be presented as one of changing literacy. Two sides here: on one side, the public in general has become more educated in the last 200 years of statistics. On the other side, statistics has become more 'scientific', in the hands of mathematical specialists, demanding a more mathematical education.

[2] See note # 16 on p. 6

[3] David R. Brillinger –Some History of Data Analysis of Time Series in the United States" p267-280 D.B,. Owen ed. "On the History of Statistics and Probability" Statistics: Textbooks and monographs, volume 17, Marcel Dekker, New York 1976

Let the observation by Solomon Kullback serve as a warning: "even professional statisticians have problems making sense of those economic equations"[4].

2. A startling Discovery about the term "Statistics". To my surprise, I discovered that none of the activities today considered the domain of statistics were called thus. Instead they were called measurements in astronomy and in engineering, information about diseases and mortality used for actuarial tables, inventories of the population for tax purposes and recruitment of conscripts, reports about agriculture, the harvest of different crops and the rural population were not called 'agricultural statistics'. Nor was the registration of marriages, birth and death in England called "vital statistics". Information on export and import was listed as "foreign trade", nor were inquiries into prices called 'price statistics'. Wm. Farr, for instance, became a "compiler of Abstracts" and was later promoted to "Superintendent of the Registrar-General's Office"[5] – really as the director of a bureau of statistics, yet the word 'Statistics' was not mentioned in this context. On the other hand the occasional publications of the late 18th or 19th century with titles like "Statistik" hardly contained any numbers, describing in anecdotal form, like a farmer's almanac, geographic and other noteworthy features of the state, but contained hardly any numbers. An exception is the term 'Census' which has been around since biblical times. Those whom today we unquestioningly consider to have been statisticians and doing statistical work, e.g. vital statistics, did not refer to their activities as 'statistics'. In fact, the terms "Statistics" and "Statistician" were not used until relatively recently in the 19th century.

The trap in dealing with history is the subconscious tendency to apply contemporary categories and understandings to events hundreds of years earlier when people had a very different understanding of their situations and using a different terminology. We ascribe without a second thought the term statistics to activities which those who were involved with them two or three hundred years ago did not think that they were performing "Statistics". Scientists like Laplace or Gauss, for example, considered themselves to be mathematicians working on measurement errors in astronomy, developing an elegant mathematical theory of human and instrumental errors. It did not occur to them to call what they did to be "Statistics". The problem arises when measurements in astronomy and the ensuing mathematical approaches to measurement errors were, without hesitation, called 'statistics' by historians looking back with today's vocabulary and concepts.

3. The Historic Roots of "Statistics" H. Westergaard[6] distinguishes three separate, main branches leading to what today is called statistics. Each had a different start, different

[4] Taken from notes I took many years ago. Unfortunately I omitted the proper identification of that publication.

[5] Westergaard, op.cit. p. 137.

[6] The early activities in England and France. Achenwall, then Schlözer who was his successor, (1735-1809) in Germany at the University of Göttingen, tried to define the term "statistics" presenting the nebulous conception of the "Staatsmerkwürdigkeiten" in Achenwall's work. (p.9). Schlözer Achenwall's successor "tried.. to give an absolutely clear definition of the term "statistics" entangled.. in the nebulous conception of the "Staatsmerkwürdigkeiten. (his Theorie der Statistik, 1804) recommended exact figures instead of vague expressions such as 'flourishing manufactures' but was not .. successful. "Such expressions found in contemporary books on "Staatenkunde". Meusel in 1804 published a "Lehrbuch der Statistik" actually calling that area statistics, but contained few numbers". Harald Westergaard, "Contributions to the History of Statistics", Agathon Press, Inc. New York, NY 1968

methods and different purposes. I am inclined to add the 'probability calculus' as a fourth branch.

3.1 "Staatenkunde or **"Staatswissenschaft**". Translated meaning "Knowledge of countries" and "Science of the comparative description of the noteworthy features of different countries". These were titles of courses offered at German Universities. It really is state-istics which morphed into statistics. It probably is the oldest ancestor of the word 'statistics' that originally was used to describe the physical-geographic features of a state (country), its people, its administration, financial matters and other economic and social aspects that today are covered as Socio-Economic Statistics[7]. The concern with taxes and the recruitment of soldiers led to attempts at counting its population. Eventually, under fiscal pressure, somewhat arbitrary samples began to be taken as estimates based on the limited available data. This broad development evolved in European countries, especially in France, Italy, Germany and the Scandinavian countries. Except for India, under the British influence, no comparable developments seem to have happened in other parts of the world

At that time professors in German universities scorned the clerks in charge of tabulating matters of interest to the state, referring to them as "Tabellenknechte"[8]. They used this derogatory term to ridicule those who compiled the numeric information in long-hand with goose quill and inkwell into lists, the 'Tabellen', as well as to mock the attempts to quantify those academicians' theoretical concepts of economics and society.

3.2 Political Arithmetic. A contemporary statistician would feel more familiar with the so-called 'Political Arithmetic' which originated in England in the 17th century relying on lists of registered births and deaths, questions of mortality and other problems of vital statistics, eventually also including economic matters. Later the expression "Political Arithmetic" grew obsolete, and the name statistics, which the German professors had used for their description of states, was adopted in its place. Mathematical thinking entered with concerns for public welfare when John Graunt studied diseases and mortality in London publishing in 1662 his "Natural and Political Observations upon the Bills of Mortality" based on the death certificates of some reliable precincts of London. Using an estimated family size of 8 persons, he produced early estimates of the population of London and also for England as a whole. His work was a forerunner of statistical sampling, but as already mentioned, at that time was not yet called "Statistics".

Another sophisticated approach to estimating the population without a complete count was done in France by Laplace 1786 and 1812. His concern with variability and precision of his estimates led to an early form of sampling theory inspired by his occupation with the calculus of probability as the mathematics of gambling. Eventually probability sampling became acceptable. Later Neyman, Pearson and Gosset developed the theory for small samples. The work of these and earlier mathematicians became a central part of

[7]" "Statistics" was originally applied to the comparative description of states (German: Staatenkunde) This can be traced as far back as Aristotle… later cultivated by Italian and other authors…it reached its culmination in German universities of the 17th and 18th century…the root of the word "statistics" in the Italian word 'il stato' and 'statista' a man who dealt with the affairs of the state. "Statistics" would mean a collection of facts … . of interest to a statesman whether in the form of numerical observations or not." p.2, H. Westergaard, op. cit.

[8]The word 'Knecht', though related to the English 'knight' in this context was not the synonym of the high riding knights of King Arthur's court, but the lowly uneducated farmhands on typical Austro-German farms, the 'Bauernhöfe'. The closest to 'Knecht' would be the American "cowboy".

the theory of statistics, leading to mathematical statistics:[9] The German Economist Karl Gustav Adolf Knies introduced the term 'Statistik' and Statistiker ('statistician') declaring in 1850 that the two separate branches concerned with public administration, 'Staatenkunde' and 'Political Arithmetic' should be united under this common name.

3.3 Measurement in Science. Apart from and independently of these administrative activities that later were called statistics, evolved an entire body of ingenious solutions to problems of measurement in the sciences. It started in astronomy. The German mathematician Gauss developed the theory of "measurement error" while working on improving the measurement of the transit time of stars[10]. Independently of this Laplace developed the first test of a statistical hypothesis while trying to determine whether comets are members of the solar system". These developments at that time were not called 'statistics'. The errors of measurement led to a general theory of their distribution and variance, leading Gauss to formulate his well-known 'normal distribution' also appropriately called the 'error curve'. From this evolved later problems of measurement in other branches of the physical and natural sciences, in engineering, in industrial quality control and acceptance sampling, as well as in industrial and agricultural experimentation with controlled inputs.

In his "The History of Statistics –The Measurement of Uncertainty Before 1900"[11] Stigler dealt expertly with astronomy and pure mathematics. He hardly dealt with Socio-Economic Statistics. What today goes as mathematical statistics belongs to the mathematical theories of measurement, of probability, set theory, and topology but not to state-istics as the description of the State. "Legendre's principle of least squares had nearly come of age as the basic tool of mathematical statistics"[12]. Neither Gauss nor Laplace would have considered their work as belonging to state-istics as the description of the state. That was Stigler's unquestioning use of the term 'Statistics', looking back at their work with the ideas of 1986. As already mentioned, early in the 20th century R. Fisher, Neyman, Goset and Yates opened new path with the designs of agricultural experiments, that included randomness, as well as research in medical and biological statistical work leading to ever more sophisticated inferences. Probability began to take over the field. Eventually a theoretical structure evolved as mathematical statistics. Meanwhile descriptive statistics in the socio-economic domain pretty much is disappearing from the teaching of statistics at all levels, as if it hardly belonged to statistics any longer. Calling Statistics' the scientific measurements in astronomy, biology, physics and other branches of science has contributed to the mistaken believe that the aggregate numbers obtained from anecdotal communication by people responding to surveys, are as reliable as the precise, instrument-supported individual measurements in the natural sciences. It is worth repeating that those kinds of statistical data are very different. Statistics in the sciences has little in common with the original meaning of statistics as 'Staatswissenschaft', the description of noteworthy features of the

[9] Wei Ching Chang "Theory and Sampling Practice" pp 299-315 in The History of Statistics and Probability, ed. D.B. Owen, Marcel Dekker, New York, Basel, 1976

[10] "Statistics in Astronomy in the United States" Elizabeth L. Scott. p 319 in: On the History of Statistics and probability, D.B. Owen ed. Marcel Dekker 1976.That article was the hindsight of an astronomer who also was a statistician.

[11] The History of Statistics-The Measurement of Uncertainty Before 1900" Stephen M. Stigler, The Belknap Press of Harvard University Press, Cambridge, Mass. and London, England, 1986

[12] Stigler, op. in cit. chapter 'The Gauss-Laplace Synthesis' p.149

state. I find it puzzling that astronomers, physicist, biologists, engineers and mathematicians adopted for their work the term "Statistics"[13]. R. Fisher went even further by referring to summary measures, like the arithmetic mean, as "a Statistic".

3.4 The Calculus of Probability. "But besides these two lines of evolution *(Staatenkunde and Political Arithmetic)* the calculus of probability deserves attention from…the results of games of dice this system reached a very high development. Laplace's 'Théorie analytíque des probabilités' (1812). For a long while … the calculus of probabilities had less influence on statistics than might be expected … confining … to abstract theories which had little or nothing to do with reality".[14] Laplace undertook in 1786 and 1812 estimates of the population of France instead of a complete census. He devised an early form of random sampling of the population, inspired by his longstanding occupation with the calculus of probability, as the mathematics of gambling[15]. His concern with variability, precision and the probability of his estimates led to an early form of sampling theory. Although today this and the work of later mathematicians Neyman, Pearson, Gossett and Yates is without question considered a pillar of the theory of statistics, at Laspeyre's time it was not yet considered to be "statistics", let alone mathematical statistics[16]. Among the sub-fields of mathematics the calculus of probability is of particular interest for the theory of Statistics. The concern with and the development of probability originated with the problems of how to win in games of chance, and became of interest in actuarial matters, tables of survival, tontines and life insurance. The number of applications grew starting with issues of the reliability of measurements in astronomy, expanding to biology, engineering, industrial quality control, acceptance sampling, large sample theory and later the theory of small samples. More recently set theory and the mathematics of computer programming languages joined the calculus of probability. Probability also became involved with 'Political Arithmetic'. Further developments that relied on probability were in decision theory and decision making under uncertainty. The mathematical theories of probability, of set theory, and of a variety of numeric algorithms became mathematical statistics[17]. Each of these mathematical fields continued their own developments apart from mathematical statistics.

---

[13]To illustrate this point imagine an interview of a contemporary **S**tatistician with **G**auss. It may have gone something like this: **S:** Congratulations, your work has become the cornerstone of modern statistics. You are one of its founders. **G**: (looking surprised) Are you sure you are talking to the right person? I don't know what you are talking about. No, I had nothing to do with what you called description of the state, or state-stics" or 'political arithmetic. I am a mathematician working on problems of an astronomer, who asked me for help in an interesting problem he encountered with the measurements of the transit of stars. But thanks for the honor.

[14] Harald Westergaard, op.cit. p.3

[15] Stigler, op. in cit. chapter 'The Gauss-Laplace Synthesis'.

[16]From: Wei Ching Chang "Theory and Sampling Practice" pp 299-315 in On the History of Statistics and Probability, ed. D.B. Owen, Statistics Textbooks and Monographs, vol.17, Marcel Dekker, New York Basel, 1976

[17]After mathematical statistics became established as a separate branch of mathematics, an Association of Mathematical Statistics was formed that began issuing the journal "Annals of Mathematical Statistics". In 1973 that journal was renamed "Annals of Statistics", its new name indicating the trend to identify all of Statistics and Statistical Theory with 'mathematical statistics'. There was a parallel change of name in the journal 'Allgemeines Statistisches Archiv'.

4. The Status of Economic and Social Statistics. These four major root-areas of today's statistics, although interrelated, followed their own developments. Graphically one could trace these developments as parallel 'zones' reaching from the distant past into the present. Despite being overlapping on their margins, each one of these areas pursues its own goals and methods, and has its own experts and scholars. In that graph the trend of mathematical statistic would appear as an expanding shadow eventually covering all applied areas of statistics as witnessed by the articles accepted in statistical journals and the kind of books reviewed there.

At present statisticians are focused on developments in the medical and bio sciences while paying scant attention to socio-economic statistics. The growing financial interest in medical and drug research steers future statisticians in the direction of the growing need for well trained knowledgeable personnel in those fields. It has been a concern of ASA and the teaching in university departments. The recent developments in the economy would lead one to expect a greater interest in socio-economic statistics by departments of economics and schools of business. This is not happening, however, because they teach in academic departments statistics for the sciences and mathematical statistics but hardly any statistics for the social sciences. The reason for this trend is the longstanding misperception that social and economic data supposedly are 'measurements' of the same kind as those in the sciences.

In his review of a book about Keynes, Oscar Sheynin, a noted mathematician, wrote: "Michael E. Brady, (*the author of the book under review*) holds that concerning the use of mathematics in economics Keynes objected to the particular misuse of certain methods rather than to the general use of quantitative methods. Among his arguments is a quotation from Keynes who declared that '*mathematical reasoning now appears as an aid in its symbolic rather than its numerical character'*. O. Sheynin also noted that "Keynes' general approach is indirectly supported by the failure to improve political forecasts, or to help to explain past political events, by straightforward applications of game theory . . . claiming that Keynes anticipated some modern conclusions according to which **statistical analysis cannot be applied in economics just as in natural sciences**[18]"*(my highlighting)*

5. Conclusions. This paper explores the historic roots of the current neglect of socio-economic statistics, tracing the development of how probabilistic thinking began to dominate leading to the mistaken identification of all statistics with 'mathematical statistics' with the consequence that the mathematical formulas and models have indeed enhanced the reputation of statistics but lead to a stagnation of Socio-economic statistics proper that have been pushed aside. This happened despite the fact that economic and social statistics are the only areas of application that deserves to be called "statistics" – the quantitative knowledge of the state.

To satisfy the current literacy requirements the books listed on the website of the Statistical Literacy Project < StatLit.Org > can be recommended, particularly the book

---

While maintaining its signature initials AStA, it was renamed "Advances in Statistical Analysis" with a similar implication that all discussion about statistics really was about mathematical statistics. It was also the beginning of the disregard of economic and social statistics

[18]On p 10 O.Sheynin book review of the publication by Brady, Michael Emmett: J.M. Keynes' position on the general applicability of mathematical, logical and statistical methods in economic and social science. Synthese 76, 1-24 (1988) Reviewed in *Zentralblatt MATH 647.90020* in his electronic publication about The History of Mathematical Statistics

"Interpreting Economic and Social Data - A Foundation of Descriptive Statistics"[19] which is well suited to enhance literacy in socio- economic statistics

[19] More information is on the internet <www.StatLit.Org/Authors-Academic/Winkler> The book is available on amazon and also can be downloaded on the internet http://www.springerlink.com/books