

Reflections on Early History of Official Statistics and a Modest Proposal for Global Coordination

*Samuel Kotz*¹

After sketching some abstract highlight features of the early statistical activities in Europe and the UK (up to the nineteenth century) the author proposes taking steps towards globalization of basic statistical concepts which are not sufficiently coordinated across different cultures and environments.

Key words: Technological advances; unification; statistical pioneers; EUROSTAT.

1. Introduction

The statistical community is witnessing at present (at the beginning of the 21st century) an astonishing lack of coordination between many hundreds of statistical offices and agencies scattered throughout the world. While the unprecedented technological advances could and should serve as a driving force and a vehicle for improvement of the accuracy, timeliness, relevance and coordination of statistical data gathered on a day-to-day basis by various private and public institutions without an overall planning, the reality is such that the efforts of dedicated civil servants and researchers are largely wasted. Substantial funds assigned for these efforts to harness modern technology, utilizing advances in statistical methodology achieved during the last 50 years for enhancing international cooperation and stability by means of routine but scientifically oriented statistical activities, may be lost unless urgent and well-planned international measures are taken as soon as possible.

2. A Brief History

Official statistical activities are the senior branch of the statistical profession – one of the oldest professions in the world. Ancient history unequivocally shows that governments have needed facts and figures about their territories mainly to provide a basis for taxation and to assess military capabilities. The linear B tablets from Minoan Crete and Cnossus Pylos deciphered by M.G.F. Ventris (1922–1956) turn out to consist largely of statistical records. The driving motive behind originally centralized statistical activities throughout the ages – from the population census ordered by Augustus Caesar, Charlemagne’s capitularies, the massive Norman enterprise of the Domesday Book to the eighteenth-century collecting of Prussian numbers by “the statistical bureau for the Prussian state” – was the unshakable

¹ George Washington University, Department of Management, 1776 G Street, NW, Suite 110, Washington, DC 20052, U.S.A. Email: kotz@gwu.edu

conviction that statistics unravels and measures the current state, power and development of a kingdom or a province.

It was firmly believed in the early nineteenth century, the golden era of official statistics throughout Europe (and to a lesser extent in the U.S.A.), that “statecraft, namely, the practical application of political science, is a mere sham without a statistical foundation.” This quotation is from the works of Ernst Engel (1821–1896), who served from 1860 as the Director of the Royal Prussian Statistical Bureau for over twenty years, governing it with an iron hand and creating a battery of institutions – administrative, educational and communicational.

His British counterpart, William Farr (1807–1883), was, between 1837 and 1877, the compiler of abstracts in the office of the Registrar General. His most spectacular achievement was pinpointing, by the examination of the statistics *compiled by him*, the source of the 1866 cholera epidemic as the water of a particular London Water Company. In the words of Francis Galton (1822–1911) – Charles Darwin’s first cousin and the founder of the science of Biometrics – Farr represented “the poetical side of statistics” as a quotation from Farr’s writings vividly indicates:

How the people of England live is one of the most important questions that can be considered; and how – and of what causes, and at what ages – they die is scarcely of less account; for it is the complement of the primary question teaching men how to live a longer, healthier, and happier life. Armed with this golden bough we may enter the gloomy kingdom of the dead, whither have gone in twenty years nine thousand English children, fathers, mothers, sisters, brothers, sons. . .each having left memories not easily forgotten. Here fortunately for this inquiry, they appear divested of all colour, form, character, passion, and the infinite individualities of life; by abstraction they are reduced to mere units undergoing changes as purely physical as the setting stars of astronomy or the decomposing atoms of chemistry; and as in those sciences so in this, the analysis of the elementary facts observed in their various relations to time and place will shed new light on the more complicated phenomena of national life.

The history of statistics (which is unfortunately not well-known even among the most competent historians of science) abounds with names of “citizens” who dedicated part of their lives to the compilation of statistical data with a firm belief in the extreme importance of their activities. One of the most shining names is that of John Gaunt (1620–1674), a London draper (who proudly referred to himself as a citizen of London), the author and the publisher of a small book, *Natural and Political Observations on the Bills of Mortality*. These *Observations* represent the first, and an extremely competent, attempt to draw scientific conclusions from statistical data. The inscription in a copy at the Library of the Royal Statistical Society reads:

Captain John Gaunt of London merits the high honour of being the founder of Statistics. His *Natural and Political Observations on the Bills of Mortality*. . .first directed public attention to the important inferences that might be deduced from correct registers of births, deaths, and marriages.

Indeed, we owe him the first scientific estimates of population size, the concept of the life table and a pioneering attempt to draw a representative sample. A deeply religious person, he eventually became a Catholic, giving up his civil and military offices. (He died

of poverty on April 18th, 1674, of jaundice, and later in the year his widow was granted an annual pension of £4 by the Draper's Company.)

Another important name on the early British statistical horizon is that of Sir John Sinclair (1754–1835) – a Scottish agricultural reformer – who was the author-editor of the stupendous 21-volume *Statistical Accounts of Scotland*. This work was the result of compendious answers to detailed questionnaires – the respondents being the ministers of the 938 parishes of the Church of Scotland.

The Anglicized version of the German term *Statistik* was introduced into the English language by Sinclair in this monumental work, whose aim was to ascertain the “quantum of happiness” enjoyed by the inhabitants of the country.

It is, however, in Pastor Johann Peter Süssmilch's (1707–1767) treatise, *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung desselben erwiesen* (2 vols. Berlin), which appeared in three editions (1741, 1747 and posthumously in (1775–1776)), that we see the beginning of the German (continental) statistical school. The two volumes (with Royal Approbation incorporated in the second edition) were unprecedented (some 50 years before J. Sinclair's Account) collections of facts combining church registers and mortality statistics. His perceptive comments on population management were among others in which the marriage rate and the age of marriage were seen to depend upon the availability of farmland. This in turn was held to fix fecundity. He predicted fluctuating birth rates. As a population grows, land is less valuable, marriage is delayed, and the birth rate drops. But in due course, there is a shortage of labor and land is more available, so the marriage age decreases and the birth rate climbs.

Süssmilch was one of the first players in the long line of “biopoliticians” whose goal was to provide “comprehensive measures, statistical assessments (the biopolitical pole) and interventions aimed at the entire social body or a group as a whole” (quoted from Michel Foucault's *The History of Sexuality* 1980, p. 138). His work was a predecessor of Malthusian debate, one of the most famous pieces of biopolitics which has special relevance nowadays to the Third World population dilemma. Süssmilch's volumes triggered the first tables on immigration, emigration, nationality, and race in the list of categories of things counted by the above-mentioned Royal Prussian Statistical Bureau. (This list, by the way, comprised, in its heyday some seven pages.) This overzealousness may perhaps be understandable in the wake of the devastation caused by the Seven Years' War (1750–1757).

It should be pointed out that the Swedish statistician Per Wargentin (1717–1783), Süssmilch's contemporary, should be given credit for the advanced state and remarkable achievements of Swedish statistics in the eighteenth century, and his data was used by Süssmilch – especially in the later edition of his *Magnum Opus*.

During the brief lull in Parisian anglophilia, the Napoleonic cavalcade strengthened the bonds between the French and German schools of statistics and in 1805 de Ferrière's famous *Archives Statistiques* published lengthy passages from Süssmilch's volumes for the first time in France.

Secret des finances de France découvert et départi en trois livres, which appeared in 1581, should be viewed as the first “statistical” publication. We are on more sound ground

when we consider the circulars of 1664 and 1697 dealing with overall statistical surveys from one administrative area as the beginning of French statistics.

The Golden Age of Regional French statistics coincides with the turbulent French Revolution (year IV (1794–1804)).

The Minister Jean-Antoine Chaptal and Francois de Neufchateau were instrumental (together with members of the Society of Statistics and the Society for the Encouragement of Industry) in producing the masterpiece which was the great *Statistique générale et particuliere de la France*, in 1804, followed by *Theoric élémentaire*, written by D.F. Donnat, a scholar-minister, in 1805. The treatise concludes with the following assertion, which is often viewed as a mixture of sociological subtlety and political naivety:

The culture of statistics cannot fail to have a good influence on the future peace of Europe, he wrote, to demonstrate that those princes best informed of the demographic or economic forces would not embark on wars; in this way statistics was a diplomatic and military instrument. As far as internal politice were concerned, the author contrived to detail its utility to private individuals; it would enlighten them as to the need for order, correct wrong opinions on social equality, protect society from hotheads and innovators, speak tellingly of the complexity of reality, be of assistance to traders and speculators. In short, statistics could be put to good use in society.

The drama of statistics in France in the beginning of the 19th century concerns the only French mathematician with statistical experience during that time. The well-to-do French statistical buffs were not aware of, much less familiar with, Laplace's (1749–1827) work on probability theory; and those who were, saw in it little practical implication. On the other hand, French mathematicians could glean little of theoretical interest in the fact-grubbing of statistics (with the exception of demography, perhaps). E. Duvillare was totally ignored when he pleaded the need to apply more sophisticated mathematical techniques in statistical tabulation and analysis.

In fact, for a long time, even in England, the mathematical development of life-tables for actuarial purposes had remained insulated from broader statistical concerns, while in post-Napoleonic Germany the implications for rank-ordering of the error theory of mathematical astronomers were only to be picked up decades later via England by Francis Galton. On the continent, only during the Restoration, with Cournot and Quetelet (1796–1874), was the theoretical framework for the analysis of numerical data to become an essential component of statistics. Hence, until the 1830s–40s, the collection, classification and interpretation of data, whether in France or elsewhere, did not rely on advanced mathematical techniques (with the exception of demography).

Its practitioners became opaquely aware of the dangers of too crude a utilization of averages only in the early 20th century.

3. A Modest Proposal

The problems that are faced by the official statisticians at the beginning of the 21st century can perhaps be appreciated by quoting some of the leading personalities.

As Yves Franchet, at the time Director General at EUROSTAT, put it so clearly at the September 1996 conference in Washington, D.C.:

In our information society, statistics become a product like any other and have to compete with all sorts of information coming from various sources. Timeliness tends to become the most important issue, very often at the expense of accuracy, reliability, and relevance. Official statistics will keep their market if *they* are timely and, above all, more relevant and more reliable than other sources.

J. Steven Landefeld (Director, Bureau of Economic Analysis, U.S. Department of Commerce) echoes these concerns in stating that, “accuracy, timeliness and relevance are the most important characteristics of good economic statistics. Although these characteristics have always been important, in today’s “information age” there seems to be special premium on them.”

In an excellent thought-provoking paper presented at the same conference on Quality of European Business Statistics, Photis Nanopoulos, at the time Director of EUROSTAT, deals frankly and touches openly with the problems related to providing the European Union with a high quality statistical information service. He points out:

Quality of the products and of the service to the users has become of increasing concern within National Statistical Institutes, which in the absence of any international standardization have developed their own standards (U.S. Bureau of the Census, Statistics Canada and Statistics Sweden are forerunners); others are developing Total Quality Management to improve the efficiency of their activities.

Further on, Nanopoulos notes:

The theoretical framework defined at European level has to be translated into the real world within each Member State. This process of interfacing between the general European framework and the specific national context presents many difficulties, due to the different cultural and linguistic environments, the administrative systems and the organizational structures of enterprises. It is thus important to observe how the general European framework is translated at national level.

To remedy the situation the author proposes to develop an alternative quantitative definition(s) which may be more flexible for adaptation within various specific national contexts, including Asian countries. His tentative conclusion is that the popular Total Quality Management approach is – to put it bluntly – an Emperor without clothes as far as statistical data is concerned, and that other avenues ought to be explored for solving the problem of defining unambiguously what is meant by the quality of *statistical data*.

New definitions of the basic factors such as timeliness, quality, relevance and coordination of statistical data from a modern global unifying perspective ought to be formalized and eventually be submitted to the appropriate authorities for universal adaptation. Needless to say that appropriately substantiated and universally accepted definitions of timeliness, based on the so-called algorithmic methodology, represent an important contribution to security and international cooperation. To the best of the author’s knowledge, the concerned researchers at Statistics Sweden are at the forefront of these problems, and at the cutting edge of these activities.

The danger facing the statistical profession *especially* in this advanced technological period, when data collection and initial data analysis are becoming almost routine operations, and the temptation to carry out more elaborate data gathering (subject to

budget constraints!) is almost irresistible, was remarkably well pointed out by Francis Galton (some 100 years ago!) and his warning should constantly be kept in mind:

So much labor is being applied to Anthropometric observation in various American Colleges, that we may appropriately consider whether it is employed in the best possible direction, and to what really valuable results it is likely to lead. It is human frailty to which statisticians are eminently liable, to look upon means as ends. They learn to take keen pleasure in the mere accumulation of neatly tabulated figures, carefully added and averaged, quite irrespectively of any use to which those figures can be applied. They are like moneymakers, who spend their lives in piling up wealth for the pure pleasure of doing so, as if wealth were an end in itself, and not a mere instrument for making life more full, more useful, and more bright.

(F. Galton (1891) "Useful Anthropometry" in Proceedings of the American Association for Advancement of Physical Education, 6, 51.)

Statistics is and will continue to be relevant when it strictly keeps to the lofty goals of increasing the "quantum of happiness" and making life more full and more peaceful.

4. References

- DeLeeuw, F. (1990). The Reliability of U.S. Gross National Product. *Journal of Business and Economic Statistics*, 8, 191–203.
- Duncan, J. and Cross, A. (1993). *Statistics for the 21st Century*. Dun and Bradstreet.
- Hacking, I. (1990). *The Taming of Chance. Ideas in Context: 17*. Cambridge University Press.
- Hilts, V.L. (1970). William Farr (1807–1883) and the Human Unit. *Victorian Studies*, 14, 143–150.
- Kenessy, Z. (1994). *The Accounts of Nations*: Amsterdam Press.
- Kenessy, Z. (1995). International Statistical Organization. *The Future of Statistics, An International Perspective*. Voorburg: International Statistical Institute, 103–146.
- Kotz, S. and Johnson, N.L. (1997). *Biographies of Leading Statistical Personalities*. NY: John Wiley and Sons.
- Perrot, J.C. and Weolf, S.J. (1984). *State and Statistics in France 1889–1915*. Chur, Switzerland: Harwood Publishers.
- Porter, T.M. (1986). *The Rise of Statistical Thinking 1820–1900*: Princeton University Press.
- Ryten, J. and McMechan, J. (1995). Globalizing the Economic Statistics System. CES, R.23 Add.1, UN Document, March 28.
- Stiegler, S.M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900*: Cambridge University Press.

Received June 2004

Revised May 2005