

ISI World Statistics Congress
International Statistics Institute
Invited Paper Session 27
organized by Professor Helmut Maier
**Is there economic order within the natural system,
and if yes how can we use it to measure and manage poverty?**

**Fundamental law of information and its impact on quality of statistics
in information society and knowledge - based economy**

by Jozef Olenski
Warsaw University, Faculty of Economics

Abstract

Worse information ousts better information. This fundamental law of information (FLI) in information society and in market driven knowledge based economy, is influencing all domains of social and economic life. In this paper the forms of impact of the FLI on social statistics and the consequences for quality of statistics are analyzed, with special reference to politically sensitive social surveys (e.g. measuring quality of life, poverty). Approaches and tools of controlling of the impact and consequences of the FLI on quality of data in social statistics are proposed.

Key words: *quality of information, social statistics, statistical standards, behavioural type*

1. Fundamental law of information

Fundamental law of information¹ (FLI) defines commonly observed phenomenon that **worse information ousts better information**. In information society, in knowledge – based, market - driven economy the processes of **ousting better information by worse information** are common, are interfering all domains of social, political and economic life, and influencing all information systems and processes on local, regional and global scale. Modern information technologies are strengthening the power of the FLI in total and global scale. Official statistics and statistical scientific surveys are also in the field of acting of the FLI. In statistics, including official statistics, the pressure of ousting better statistics by worse statistics are visible, under slogans of higher efficiency, reduction of costs of statistical production, outsourcing.

Main reason of the power of impact of the FLI on quality of statistics is that in information society and in knowledge - based economy information processes and systems are driven by the market. That means that information processes, which are and should remain purely semiotic social processes, are becoming - first of all - the processes of production of information, and are designed, managed and maintained like any other processes of production of goods and services in market - driven economy.

Unfortunately, the market - driven approach has been commonly adapted to statistics, especially to the systems of official statistics and to research projects sponsored by governments, international organizations and businesses. In all countries the economic efficiency of production processes of statistical data (*more data, cheaper, faster*), is the main criterion of quality of official statistical system for the governments and for other institutions, including international statistical organizations. The consequences of that approach are systematic cuts of budgets for statistical activities in almost all countries, the pressure of funders of statistical surveys (governments, international organizations, businesses) on cutting the resources for methodological works, shortening the time of research and analy-

¹ The fundamental law of information is defined and analyzed in: Olenski J., *Fundamental law of information in knowledge – based economy*, Warsaw University, Faculty of Economics, Warsaw 2011, (in print)

ses. Official statisticians and scientists realizing sponsored “projects” are working under strong and more growing pressure of budget limits and rigorous timeliness. The price of that pressure is – first of all - the quality of statistical information.

Very important phenomenon of information society and in knowledge – based economy is the absolute **information gap between** producers and users of information. The end - users are not able – as a rule – to evaluate the quality and pertinence of statistical information. Also sponsors, funders and organizations supervising statistical surveys are not able to evaluate the quality of surveys and data received. The only way of control of the quality of surveys and data is to check if the surveys meet scientific methodological standards approved *ex ante* and if the statisticians obey the rules of statistical ethics. However, the elaborating of methodology is an integral part of the surveys as the process of production, so there is no practical possibility of verification of the control of quality of information in statistical processes. The only “safeguard” of the quality of statistical information is the **professional ethics of statisticians** realizing surveys.

The power of acting of the FLI is especially strong and explicitly visible in social statistics. The reasons why official social statistics is so strongly “attacked” by the FLI are (a) political sensitivity of official social indicators, (b) conventionalism of international and national methodological standards in social statistics, (c) social indicators are interesting for common, non professional users; journalists publishing social indicators are “translating” them from professional language of statistical surveys to the common language of mass media, (d) social phenomena and related statistical indicators are specific for local and regional economies, social groups, civilization and cultural environment, it is not possible to elaborating relevant, adequate and - at the same time - universal statistical standards.

The objective of this paper is the analysis the forms of the impact of the FLI in social statistics, with special reference to quality of life and poverty measuring, and the evaluating of the consequences of this impact to the quality of statistical description and analysis of social phenomena. The exemplification of general methodological considerations is concentrated on the concepts of *quality of life* and *poverty*. Some proposals what to do to control the acting of FLI on social statistics and to protect social information against the processes of replacing better information by worse, often bad and misleading information, are also discussed.

2. Forms of impact of the fundamental law of information on quality of statistics

Fundamental law of information is occurring in practice in many forms, in all statistical processes and systems. Main forms that should be identified and taken into account by all statisticians and scientists who are the stakeholders of statistical processes are following:

- Specificity of social phenomena *vs.* generalization of statistical description
- Subjectivity of social perception *vs.* objectivity of identification and measuring
- Conceptualization of social phenomena. Scientific *vs.* “intuitive” concepts.
- Quality of social phenomena *vs.* quantity of statistical measuring.
- Quantitative statistical approach *vs.* qualitative analytical approach
- Global standards *vs.* specific standards
- Fragmentarily statistical observation and measuring *vs.* holistic approach of social analyses
- Use of language of mass media *vs.* use of professional statistical language
- Use of social statistical indicators for political purposes *vs.* use of statistics for scientific purposes
- Conflict of ethics: statistics as science or pragmatic service

3. Specificity of social phenomena *vs.* generalization of statistical description

Social phenomena are related to concrete individuals, households, social groups, localities, regions, cultural and civilizational **behavioural types**. Social phenomena and processes exist in specific civilizational environment. They should be analyzed only in full cultural and civilization context. In the analyses of social processes model of economy on all levels: individuals, households, local societies, national economy, should be treated as the integral part of civilization environment. Methods of

identification, modelling, selecting and defining their important attributes and methods of measuring should be adjusted to the specificity of civilizational environment. Analyses and evaluation of social phenomena must be embedded in specific civilizational context. Otherwise both statistical measuring, indicators and statistical analyses may be confusing and misleading for the users not acquainted well with the reality of specific social environment. Especially ambiguous are **aggregated composite indicators**, “mixing” and arbitrarily weighting selected specific indicators. Good, commonly known example of such composite indicator in HDI (human development index), which only advantage is its simplicity (*o sancta simplicitas!*).

For example, social concept of the *quality of life* is very useful statistical concept only if it is measured as the matrix of indicators describing different specific attributes referring to the quality of life of an individual and household in specific civilizational and economic context. This context is different in atomized societies of huge industrialized agglomerations with developed institutions of social security and insurance, it is different in exploding "mega-agglomerations" in Asia or Africa, and will be completely different in small village societies in Switzerland or Tajikistan. Each specific attribute of the matrix of indicators measuring of the *quality of life* should be interpreted in these different environments and civilizational contexts.

One of important characteristics of civilizational context is the *model of a household* and its role in the life of individuals. For example, a main characteristic of the *quality of life* is both objective and subjective *social security* of an individual and of a household in all aspects: psychological, health-care, economic, environmental, external security, informational security etc. *Social security* as the attribute of the *quality of life* of an individual in socially atomized cities of industrialized or post – industrial economies depends very much on institutions and governments. In that institutional context official statistics can measure the *security* and a characteristic of the *quality of life* by measuring the resources used by social security institutions (social care, police, social actions of local governments, NGO's etc.), registered value of services provided by them to individuals and households and measuring subjective satisfaction of individuals and households from the activities of those institutions. In other societies, in which the security of individuals and households is based on direct close relations of families and neighbours, the absence of governmental institutions and NGO's in the life of families, households and local societies does not mean lower level of security as the attribute of the *quality of life*.

Analyzing and interpreting the matrix of variables characterizing the social *security* as an aspect of the *quality of life*, the user should see and understand the relations of complementarity between the elements of this matrix, e.g. subjective high level of security of an individual thanks to strong familial relations vs. in other case – thanks to well developed institutional social assistance networks organized by governments.

For example, statistical concept of *poverty* is strongly related with civilizational context of individuals, households and specificity of local entities of national economies. Statistical concept of *poverty* is described by the matrix of statistical variables describing economic situation of individuals, households and families. The relations of complementarity between those variables are crucial for proper interpretation and evaluation of the *level of poverty* of an individual, family or household. Low level of income achieved by an individual in monetary form for paid work may not be the symptom of poverty if that individual is a part of family providing satisfactory level of economic security for all its members, not necessarily living in the same dwelling. On the other hand an individual getting relatively higher monetary income may have negative income for disposal, may have very low level of objective and subjective security and should be considered as *poor*.

The conclusion from the considerations above is that any statistical analysis of social phenomena should be referred to specific behavioural types of individuals, households, families and local societies embedded in their civilizational context.

4. Subjectivity of social perception vs. objectivity of identification and measuring

Social statistics has not elaborated general methodological platform, which in economic statistics is realized by the system of national accounts (SNA). It seems that similar general platform for

social phenomena and processes is neither possible nor necessary. Social phenomena and processes should be observed, analyzed and evaluated by statistics from two aspects: subjective and objective. Those two aspects **should be closely interrelated** in all statistical surveys. Separated, isolated identification, observation and measuring of subjective perception of social phenomena without parallel observation of their objective (hard) aspects may lead to misunderstanding and analytical mistakes, as well as – *vice versa* – the analysis and evaluation of social phenomena on the basis of objective attributes and hard data only is showing only one “side of the coin”.

In social statistics the measuring of subjective perceptions and, from the other side, objective values of the same or equivalent attributes of the same classes of social entities (e.g. individuals, households, families, social groups etc.) and processes (e.g. participation in culture or political life, access to health services, intellectual capital etc.) should be integrated in one survey or in the set of interrelated, methodologically integrated surveys.

5. Conceptualization of social phenomena. Scientific vs. “intuitive” concepts.

Conceptualization of social phenomena, processes and entities is one of most difficult practical task for official statisticians. Social sciences define models social phenomena, processes and entities in different ways specific for each domain of social sciences. Official statistics responsible for collection of statistical data for many users (economists, sociologists, politologists, businesses, administration, politicians, etc.) is expected to conceptualize the surveys taking into account the needs and interests of wide scope of different users. In practice it is not possible to elaborate concepts, definitions, statistical units, and statistical indicators acceptable and satisfactory for so wide range of users.

Social statisticians involved in official statistics have elaborated **generally** accepted concepts and definitions of basic statistical social entities and basic statistical concepts describing most common social phenomena and processes. Official statisticians have also elaborated classifications and nomenclatures used as metadata standards in social statistics. Those concepts and definitions have – as a rule – good scientific basis. Most commonly used concepts and definitions, classifications and nomenclatures are introduced on international and national levels as mandatory standards or optional recommendations. Professional end – users and scientists accept those standards (although they may not satisfy their all specific needs).

The “conflict of interest” appears between some groups of important users that are not acquainted enough with statistical methodology, but expect that they are entitled to using statistical data according to their professional and ethical principles and interpreting statistics in their own professional language. Those users belong to the following social groups: (a) politicians, (b) journalists and commentators in mass media, (c) high level officers in public administration, (d) businessmen, (e) social activists in NGO’s. Most of them are using intuitive understanding of concepts, which lexically are identical or similar to scientific statistical concepts. Basic concepts of social statistics and quantitative statistical data are used by those users, often in the same time and place in different meaning and in different social contexts. The objectives of using scientific statistical indicators for political purposes may be different from the objectives of using social statistics for scientific purposes.

Statistical concepts that have recently made real political career are *poverty* and *poverty line*. The interpretation of official statistical data characterizing the *level of poverty*, by number and structure of people and households *living above or below poverty line* etc. depends on the objectives of concrete users: who and for what purpose is using those indicators. Sometimes the same statistical numbers and terms are understood and interpreted in different way, and - what is more difficult for identification – are interpreted in the same way for different social environments and contexts, or without precise identifying and defining those contexts precisely or explicitly. In mass media, in political discussions and administrative decision – making processes the intuitive, **subjective narration** with statistical terms and numbers inside, **is replacing scientific statistical information**.

Official statisticians and scientists involved in conducting statistical surveys and data production should pay more attention and effort to elaborating metadata relevant for precise description of social entities, phenomena and processes. Standard metadata help to control the quality of methodology of surveys and data produced. Official statisticians should be more active in disseminating those

metadata as standards not only for statistical purposes, but also for other areas of social and economic activities.

6. Legal framework vs. real life

Many social processes are regulated by laws. Because of that, many concepts, classifications and nomenclatures as well as many statistical concepts and indicators are originally invented and defined not by scientists and statisticians, but by lawyers and politicians in early stages of the processes of legislation and editing the texts of legal acts. Each legal act has its “statistical component”, in which statistical concepts and data used as regulators are specified. In practice official statisticians and social scientists are not invited – as a rule – to the legislation process from the beginning. Sometimes statisticians and scientists are involved, but in final phases of editing the laws, when concepts, definitions and indicators are already defined and “heavily negotiated” by politicians, social activists, lobbyists and public administration. Even if statisticians are participating in legislation processes from very beginning, they are rather expected to answer the question **how quickly** statistics would be able to produce demanded data, and they are not asked to proposing concepts, definitions and related statistical variables, algorithms of computing the variables, methods of interpretation of meta-information and numerical data.

It seems that the duty of official statisticians (see UN FP) is the explaining and convincing the leading stakeholders of legislative processes that the role of statistics should be the role of important partners in elaborating good regulations of social and economic processes, supported by relevant statistical information.

7. Quality vs. quantity of statistical measuring. Quantitative statistical approach vs. qualitative analytical approach

Statistical description of social phenomena and processes is complete only if quantitative statistical indicators and variables are associated with qualitative information on specific social entities, on their social, political, economic and legal context. Official statisticians have elaborated standard metadata and methods of measuring and computing quantitative variables. Many of them are recommended by international and supranational statistical organization as standards. However quantitative component of social statistical information is up to now the domain of creativity of statisticians, researchers and users.

It seems that best practices of standardization of qualitative components of statistical information elaborated by economic statisticians should be carefully analyzed by social statisticians and creatively adopted for the needs and for the benefit of social statistics.

The above suggestion is also relevant to statistical analyses of social processes. However in social analyses we may observe the difference between analytical discipline of econometricians and humanistic freedom on invention of sociologists, social psychologists and politologists.

8. Global standards vs. specific standards. Problems of comparability in social statistics.

As it was mentioned above, economic statistics has elaborated internationally accepted conceptual standards and common methodological platform of the SNA and related economic satellite accounts. In most of the domains of social statistics the process of international harmonization and comparability of social statistical indicators is in early phases of discussion. Some positive effects have been achieved in the field common both for economic and social statistics, for example in labour market statistics, in the statistics of economic aspects of activities of social entities (e.g. LFS, household budgets, expenditures and quantitative data on standardized social services provided by governments and NGO's etc).

There are expressed the expectations from the part of politicians and scientists that official statistics shall produce comparable statistical data enabling international and regional comparisons of social phenomena and processes. Politicians need data enabling the to compare the levels of social development, quality of live, poverty, of different social groups in different regions, cities, countries. Social

scientists are looking also for statistical data for international comparisons of dynamic social processes. Some level of comparability has been achieved thanks to international standard classifications, nomenclatures, concepts and indicators for basic social phenomena, that have been elaborated and recommended by UNSC, OECD, Eurostat and other organization. Many countries use those standards directly, or – more often - as the bases for national classifications, nomenclatures and concepts. Many official statistical institutes produce social data following those standards.

However, as it was mentioned, any social statistical information should consist of two interrelated sides: (1) results of statistical surveys, and (2) descriptive structured information on social, economic, ecological and political environment of social entities. Achieving international comparability of statistical variables (numbers) is only one part of statistical information needed for social statistical analyses. It is also needed the comparable, harmonized information on the environment of social entities and processes.

For example, the *poverty line* computed in the same way for all regions or for many countries may have different values of indicators for different regions and countries. E.g. standard definition of an *employed person* (see standard definition of unemployment of the LFS – one hour of paid job per week !) may not be adequate for the situation of most of European countries. Proper interpretation of data on the quality of life should be embedded in real social context. For *quality of life* in one region decisive may be ecological situation (e.g. for a region of ecological catastrophe), in other region – access to equivalently paid job (e.g. for a region of very high structural unemployment), in other region or country – proper level of safety of citizens and effective system of justice, in other region – access to health care services or education, or political freedom.

The integrity and comparability of numerical values of indicators is dependent on the comparability of social and economic environment. Bringing social data to comparability on international or trans-regional scale can not be limited to the verification of the integrity of methodology of surveys, and homogeneity of algorithms of computing the variables.

Bringing social statistical information to comparability is the conceptual process of constructing models of social processes adequate to the specificity of the country, integrated analysis of objective and subjective indicators and finally the interpretation of basic concepts within the frames of social, economic, political and ecological environment in which societies are living.

10. Fragmentation of statistical observation and measuring vs. holistic approach of social analyses

Social statistics is a complex of a large number of separate problems - oriented surveys. In statistical practice one may distinguish three main approaches to defining the problem and methodology of survey: (1) collection of available data on some classes of social entities, processes or phenomena (e.g. household budgets), (2) verification or falsification of pre-defined thesis or theses (e.g. victimisation of women), (3) combined approach i.e. collecting of available data selected from the point of view of verification or falsification of some theses (e.g. LFS, EU-SILC). Common feature of all these approaches is the *thesis - driven fragmentation* of observations of statistical entities and the thesis – driven pre-selection of observed and measured attributes.

Social surveys are designed and conducted autonomously. The idea of developing an integrated system of social surveys is up to now a slogan and an objective that has not been achieved in practice. It seems that the reason of unsuccessful of building the *integrated system of social surveys* is the qualitative variety of statistical entities and the lack of common conceptual platform for different social phenomena. In economic statistics there is one basic class of statistical entities: an establishment. Economic statistical variables are the attributes of establishments and enterprises as economic entities. Economic statistics had also developed one integrated conceptual platform for survey: the SNA. To the contrary, in social statistics many qualitatively different classes of entities are identified, the defining of the class of entity often depends on the objective or thesis of a survey. Up to now common methodological frame for driving samples is rather a futuristic idea, not the reality.

The consequence of the lack of common methodological platform and problem-driven or thesis – driven approaches typical form sociological surveys is that particular social surveys are collecting

fragmented information on some selected aspect of specific class of social entity. Statistical data collected in one *problem – driven* or *thesis – driven* survey can not be integrated with other survey oriented to describing and analyzing other problems or to falsifying other theses.

In official statistics many efforts have been made to developing more integrated methodological frameworks for domain - oriented complex of statistical surveys. E.g. in many countries the LFS frame and samples are driven to enabling the collecting of data on economic activity of households and at the same time to collect information (so called modules) on qualitatively different topics. For example, some statistical offices use the LFS survey as the “carrier” for other subject - matter modules like the survey on ICT in households or on participation in culture etc.). The LFS frame occurred to be an effective tool for driving samples for other modules by adapting so called *behavioural types approach*. For example, one *behavioural type* of a household is the household that is one socio – economic unit, characterized with “typical” attributes:

- Number of members of household, its age and sex structure
- Education, professions, skills,
- Health status, disability (in a household),
- Economic activity (active members of households, number of inactive members, reasons and structure of inactive members), main sources of income,
- Employment records,
- Localization, access to labour markets (local, regional).

The *behavioural type approach* seems to be the promising direction on the way to developing more integrated system of social surveys. However the specificity of society and variety of phenomena of social life, as well as the specificities of methodologies of sociology, politology, social psychology, ethnography etc. and measuring subjective attributes is determining the limits of integration of social statistics and of the comparability of social data.

11. Use of language of mass media vs. use of professional statistical language

Important for quality, integrity and objectivity of statistics is the propensity of “VIP” end – users of statistics (politicians, managers, businessmen, experts commenting statistics in mass media) to presenting statistical information in “more understandable” language, i.e. in the language of mass media and in colloquial language. The semantics of those languages does not contain precise scientific concepts elaborated by statistics.

The danger of *oversimplification* of statistical information in the process of “translation” of the outputs of statistical surveys represented originally in professional statistical language, to the language of mass media or to colloquial language, is real. Often it leads to misunderstanding and to the misusing of statistical data. It is especially dangerous for the credibility of statistical agencies and statistics as science in those areas of social statistics that are describing and analyzing the issues politically sensitive, common for all members of the society, like *conditions of life, poverty, safety* etc.

Statistical literacy and respecting the principles of statistical professional ethics by the stakeholders involved in dissemination of social statistics to the public is necessary to protect credibility of statistics, statisticians and statistical agencies.

12. Conflict of ethics and FLI

Last but not least the form of acting the FLI in social statistics refers also to professional ethics. The stakeholders of statistical processes in social statistics are following their own professional ethics or the lack of ethical rules in using and managing information. The differences between statistical ethics represented by official statisticians following the Fundamental Principles of Official Statistics, economists representing some economic theories, sociologists interested in proving some ideas, are generating ethical conflicts. Extremely important is the conflict created by the difference between professional ethics of journalists (if such ethics still exists in practice), ethics of different groups of interests represented by end – users.

In the process of designing surveys, defining the objectives of surveys, elaborating the methodology, processing, interpreting and disseminating the data, the statisticians taking the responsibility for the surveys should know real ethical approaches of all classes of stakeholders and should envisage potential impacts of *ethical gaps* of different groups of stakeholders on the quality of methods, data and the use of data. The impact of the FLI in the sphere of ethics is extremely strong, because there are very limited possibilities of control of ethical behaviour of stakeholders of statistical processes.

The way to protect the quality of statistical information is the dissemination of the knowledge on the ISI Deontological Code, UN Fundamental Principles of Official Statistics, other codes of professional ethics and codes of best practices of national statistical societies and international statistical organizations.

13. What statisticians can do to protect social statistics against the impact of the FLI in market - driven knowledge - based economy?

The impact of the FLI on social statistics is extremely strong. Because of the specificity of subject matter areas of social statistics, methodological differences of sciences involved in social surveys and the specificity and wide scope of users of output data, the forms of influence of the FLI on social statistical surveys is stronger than in economic statistics.

Main tools of control of impact of the FLI on quality of data are:

- a) Statistical ethics. Subordination of other professional ethical rules of stakeholders to the principles of ISI deontological code, UN Fundamental principles and statistical codes of ethics of countries and of statistical associations.
- b) Developing common meta-information standards for describing social entities, phenomena and processes (classifications, nomenclatures, concepts and definitions).
- c) Developing standard behavioural typologies of social entities, processes and phenomena.
- d) Integrating statistical surveys on the basis of common multi – purpose statistical frames and using specific surveys as the “vehicles” for several topic oriented or problem – oriented modules (e.g. LFS or households budgets may be used as “vehicles” for modules).
- e) Developing common methodological principles harmonizing and building methodological compromise of sciences involved in social statistics (economy, sociology, social psychology, politology etc.)
- f) Reacting on misuse of statistical data for non – statistical purposes.
- g) Control of language in which products of social statistics are disseminated to the public and to external users.

The authority of *statistics as general methodological science* may help to elaborate effective tools of control of the FLI on social statistics.

14. Bibliography

- 1) ISI Declaration on Statistical Ethics, Amsterdam 1885
- 2) UN Fundamental Principles of Official Statistics, Geneva 1992
- 3) Olenski J., *Fundamental law of information in knowledge – based economy*, Warsaw University, Faculty of Economics, Warsaw 2011 (in Polish, in print)
- 4) Olenski J., *The citizens` right to be provided with information and the duties of democratic state in modern IT environment in a democratic state in the light of the UN fundamental principles of official statistics and the ISI declaration on statistical ethics*, Invited paper for the Special Committee on Statistical Ethics of the International Statistical Institute, 53rd Session of the Statistical Institute in Seoul 20-28 August 2001.
- 5) Olenski J., *The framework of standards in social statistics - with special reference to transition countries in the ECE region - non-candidates to the European Union*, Economic Commission for Europe, Statistics Division. Invited paper. Geneva, June 2001

- 6) Olenski J., *The citizens` right to information and the duties of democratic state in global information society in the light of the Fundamental principles of official statistics and the ISI declaration on statistical ethics*, w: Proceedings of the ECIS 2001 - the 9th European Conference on Information Systems, 26 - 28 June 2001, Bled, Slovenia.
- 7) Olenski J., *Information crises in new economy*, Proceedings of the International Interdisciplinary Conference Crises and Sustainability, Warsaw, 9 – 10 October 2004
- 8) Olenski J., *Law of information quality in information society and in market – driven, globalised knowledge – based economy*, Proceedings of the International Conference Informatics 2005, Bratislava – Vienna 2005.