



Towards a Political Economy of Statistics

Klaus Heine

Erasmus University of Rotterdam, the Netherlands – heine@law.eur.nl

Erich Oltmanns*

Federal Statistical Office, Germany – erich.oltmanns@destatis.de

Abstract

Discussions are being held in many countries on the future of official statistical data production. As a contribution to this discussion, this article examines some aspects of a “political economy of statistics”, focussing on “statistical operationalization”, which is seen here as a central challenge for data production in the field of economic and social activities.

Keywords: statistical offices; data production; public choice.

1. Introduction

A broad social consensus exists on the role of the official statistical infrastructure in a democratic society. In 1992 the United Nations stated (restated in 2013) in the *Resolution of the Fundamental Principles of Official Statistics*: “Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens’ entitlement to public information.”

However, national statistical offices (NSOs) in many countries are confronted with new challenges. NSOs have to handle significant budget reductions that force them to increase efficiency and/or to reduce the staff. Additionally, there is, no consensus regarding the extent to which the existing statistical infrastructure fulfils the requirements of the *Fundamental Principles of Official Statistics*, or on how the structure should be developed in order to optimise its effectiveness. In fact, official statistics are currently the focus of debate in many countries (see for instance Eltinge, Biemer, and Holmberg, 2013). For example, it is said that some of the data on offer are obsolete or even superfluous, yet at the same time there evolve new data requirements from developments in society and economy that are not fulfilled by official statistics. The great recession and the renaissance of social indicators “beyond GDP” may be well-known examples.

To discuss these issues, in this paper the *New Political Economy* is applied. The paper is structured as follows. First, a brief analysis of the current discussions regarding official statistics is provided. Then some basic aspects of *New Political Economy* are discussed. The next section takes a thorough look at the importance of *operationalization* within the framework of the production of statistical data. Then, the main section presents various fields of application for a political economy of statistics.

2 The ongoing debate on official statistics

Discussions are currently being held in a number of countries on the future shape of the statistical infrastructure. Notable contributions were being made, for example, by Dillman (1996), Fellegi (1996) and Sundgren (1996). More recently the debate has become revived in a special issue of the “Journal of Official Statistics”, where Penneck (2013, p. 187) summarizes the actual challenges for statistical infrastructure as follows: “The redesign of statistical systems to reflect an agency-wide strategic architecture is one of the most fundamental challenges a statistics agency faces. It is both necessary and difficult. It is necessary because it will yield the cost savings and quality improvements that funders and users demand. The development of methods, systems and processes to common standards under a unified architecture yields savings in development and maintenance costs, which are timely

given ever tightening budget constraints. At the same time common metadata standards, methods and approaches guarantee quality standards at a time when users become more demanding. However, although the agency-wide benefits of this approach are self-evident, statistical producers have found this approach hard to deliver.“

These discussions have, up until now, very often bandied about arguments along the lines of a “basic supply of statistics” or “continual flow of data” which the government alone could guarantee. Similarly, the organisational principles of official statistics - particularly the obligation to provide information and the principle of confidentiality - have been evoked in order to justify the involvement of the government in the field of statistical data production. Only little attention was paid on the whole to the fact that economic theory can also be employed to legitimate the public production of statistical data. In particular arguments from “market failure theory“ and “new political economy“ can be drawn. Market failure theory argues that statistical information has the characteristics of a public good and that the collection and storage of data implies economies of scale (see for example Reamer, 2010). As a result the market will neither produce a sufficient amount of statistical information (public good), nor will there be a competitive pricing of the statistical information. If one takes dynamic effects also into account, it becomes in addition doubtful how much innovation a private producer of statistical data will come up with. This is not to say that a public producer of statistical data will provide the information in an efficient way easily. But one can suspect that the public production of statistical data has some generic economic advantages against the private production.

So far it was assumed that the individuals involved in the production of statistical data behave rational in the sense that the private production of data suffers from market failure, but that government can cure that by public production. This implies that the political process behind the public production of statistical data runs without flaws. The political decision maker is assumed as a benevolent actor that does not decide in his own self-interest or strategically, but in the interest of society as a whole. However, that must not be the case. There is indeed a political process behind the production of statistical data that is driven by political interests (politicians), rent-seeking groups (industry) and bureaucracy (departmental red tape). From that political process it cannot be assumed that it will – being unconstrained – lead to desirable results. Thus in a political economy of statistics institutions and law play a decisive role, because they are the instruments with which the political process becomes constrained and, moreover, gets democratic legitimacy.

3 The production of statistical data and political economy

The approach of new political economy (public choice theory) is a branch of political economy theory which stresses the interdependence of policy and economy. It is based on the assumption that actors in the field of economic policy (politicians, economic policy advisers, bureaucrats, etc.) work on the principle of self-interest and do not necessarily strive to improve the welfare of citizens for which they are responsible. The normative purpose of the new political economy is to propose rules which are agreed on the level of a basic consensus and which should guide subsequent action within the ongoing policy-making process in order to place the greatest possible restraints on self-interested action by economic and political actors (Vanberg, 1994; Mueller, 2003).

Information plays an important role both at a constitutional level and within the ongoing policy-making process. Its significance lies in the fact that actors in the political and economic system (politicians, voters, etc.) are not, as a general rule, fully informed. They therefore request information, although their purpose is neither to be fully informed nor to maximise the welfare of the public, but to seek individual gain.

In this respect, the statistical infrastructure has an important role to play. Information of relevance to society is provided in the form of statistical data and is judged by different stakeholders to be of varying significance depending on the benefits they derive from it. Voters, for example, are primarily interested in key economic data which are conveyed to them mainly via the mass media, whereas politicians look for a significantly broader range of data. This information is not just taken directly from the supply on offer from the statistical infrastructure, but is often obtained via economic policy advisors, who themselves request information from the statistical infrastructure and/or are themselves

actively engaged in the production of data (Heine and Mause, 2004).

Statistical data can thus be characterised as a *product* that plays a crucial role in the political process. At the constitutional level, when a basic consensus is sought, the production of statistical data is already an issue: it is first of all necessary to agree on objectives relating to the statistical infrastructure. As mentioned in the introduction, there is general agreement on this point. Detailed discussions must then be held on what form these objectives should take with, for example, decisions being required on those areas of data production in which activity of the government is justified and those in which private production is sufficient and legitimate. When it comes to official statistics, questions also arise regarding the way in which data production is organised. Finally, a basic consensus needs to be reached on the general conditions governing the behaviour of the bodies which have been selected to make up the statistical infrastructure (rules on the transmission of data, etc.).

Hence the purpose of a political economy of statistics should be to disseminate proposals on achieving a coherent statistical infrastructure that best serves the interests of society. Thereby the specific characteristics of statistical data as a produced good are crucial. As a matter of fact some of the key characteristics of this product are closely linked with statistical operationalization.

4 Statistical Operationalization

The concept of operationalization is usually derived from Bridgman (1927) whose method is called operationalism.¹ Within the framework of economics and social sciences it means that the concepts developed and used by theory cannot be observed and measured in a direct manner. Instead, it is necessary to use variables that can be empirically measured. And the challenge is to construct empirical measurable variables that correspond at the best to the concepts of theory. In principle this touches upon the foundational problem of model specification in statistics, when a theoretical model from social sciences is empirically applied.

Statistical operationalization can be portrayed as three interrelated structural elements. The first, which can be described as the *horizontal aspect of operationalization in a broadest sense*, catches the variability of a theoretical concept with regard to time and content. This means that a specific operationalization may be useful for handling one problem, but may be irrelevant in other cases. Similarly, when a suitable specification is found it does not last forever, but can at some point become obsolete by economic and social progress and associated changes in models and/or social values.

The second structural element can be described as the *horizontal aspect in a narrower sense*: Here the idea of operationalization is based on the assumption that there is a fundamental gulf between the theoretical concept and measurable variable which cannot be bridged. Therefore a statistical specification cannot be viewed as a snapshot of reality, but it involves the creative construction of models to record reality which then triggers a learning process towards a better approximation of the underlying model (Pearson 1955). And according to this (constructivist) view, even for a given theoretical concept there exists a variety of solutions to the problem of operationalization.

The structure is completed with a component which can be called the *vertical aspect of operationalization*. If it is assumed that the construction of numerical concepts in statistics is dictated by theory, then the theory building is of great importance, because theory provides the framework within which statistical operationalization takes place. This leads to the idea that the body of economic and social theories consists of a hierarchical structure from a simple “basic model“ via “variants of the basic model“ up to “targeted theories“ for the analysis of specific issues. Yet this model structure has to be applied logically consistent to the numerical concepts of statistics. This means that problems of operationalization do not only exist horizontally between competing theories on the same level, but also vertically at widely divergent levels of theory. In addition, the numerical concepts for specific levels of theory can be understood as specific cases of numerical concepts from general levels of theory.

It should be noted that the *vertical aspect of operationalization* involves a conceptual shift with regard to data production. It emphasises the *functional* against the *institutional* perspective. In other words,

¹ See also <http://en.wikipedia.org/wiki/Operationalization>.

attention is given not to the gulf between data producers and data users, which has hitherto marked the discussions on the shape of the statistical infrastructure, but on data production as a coherent multistage process which is targeted at specific (research or policy) goals. This perspective will be taken for discussing the production of statistical data from the viewpoint of economic policy.

5. The statistical infrastructure from the perspective of economic policy

The idea of statistical operationalization suggests that statistical data should be produced on the basis of a scientific theory in combination with a research or policy goal. Official data production, on the other hand, is characterised by the fact that the applied numerical concepts are determined by legislative or legal means, which assures their legitimacy. They are basically the outcome of political agreements between stakeholders from the worlds of politics, economics and law.

At very general levels of theory it can be assumed that these stakeholders' ideas and intentions on determining numerical concepts will generally correspond and cause no bigger problems. However, if the level of specialisation of the underlying basic theories or the stakeholders' interest in the findings increases, it becomes also likely that the numerical concepts are determined more in accordance with the ideas of those stakeholders who exercise the greatest influence in the decision-making process. This can explain the dissatisfaction with the statistical infrastructure expressed quite often by actors from the field of economics and other sciences. They generally belong to “lower order“ interest groups. Hence the data available on the basis of the numerical concepts selected do not - or only partially - satisfy their requirements. That is, the numerical concepts at hand are only adequate for a limited subset of specific theories.

This raises the issue of improving the relevant decision-making mechanisms. At first sight, it looks as if there was a need for mechanisms which give disadvantaged interest groups greater influence to decide on the statistical infrastructure (right of veto, blocking minorities, etc.). Within the paradigm of the new political economy, however, it is doubtful whether this would lead to a satisfying qualitative improvement in the official data available. Given that (self-interested) actors will find a means of extending their influence (by forming coalitions, swapping votes, etc.), this problem would *ceteris paribus* not be resolved, but merely displaced or would pop up in more sophisticated ways.

At this point, it is possible to bring the *vertical aspect of operationalization* into the debate. As a starting point it is safe to say that official data should service the demands of the widest possible range of users; it is tax payers' money that funds the production of official data. Therefore, the policy-making process should not be targeted at the satisfaction of specific demands concerning the statistical infrastructure (which correspond with middle or upper levels of theory), but to create a stock of data at very general levels of theory, which all data users can tap into on a vertical plane to satisfy their specific thirst for knowledge. In other words, a robust statistical infrastructure is needed from that then specific numerical concepts can be derived. These more specific concepts correspond to the specific needs of higher levels of theory and thus serve also the interests of specific data users. The advantage of such a conceptualization of data provision is apparent. While a less malleable statistical infrastructure will always give advantage to those interest groups which have the factual power to determine the official statistical infrastructure, a concept of statistical infrastructure that is focused on general levels of theory cannot be captured so easily by interest groups. Insofar such a concept of official statistical infrastructure can also claim democratic legitimacy.

This is not the place to discuss what physical shape such a stock of data would take, although the data would need to be much more disaggregated, which would mean that the production of official data would even actually need to be expanded. The supply of such data would also require a different emphasis, with priority no longer being given to the organisation of individual surveys, but to re-designing the statistical infrastructure as an interlocking system of surveys. Nor is it easy to answer the question which political decision-making mechanisms would be best suited for a data supply of this type. On the one hand, there is certainly a case for political consensus among the stakeholders and interest groups involved. But it is also important not to overlook the implied danger, when decisions about the official statistical infrastructure are cast into the political arena. This danger would not just arise because of the wishes of users whose voices had not been heard in previous rounds of the

decision-making process, but also because it might be worthwhile to express other desires which had hitherto been covered by private data suppliers. As a result rent-seeking coalitions might emerge, in which via logrolling the production of official statistical data becomes enlarged beyond its socially efficient amount (Mueller 2003, pp. 104). Obviously the political decision-making process needs checks and balances. For this reason, a more principal consensus on the form which the data supply should take has to be found. It is this consideration which will be examined in greater detail in the following section.

The question which institution is best suited to playing this regulatory role can be answered by using two criteria. The first concerns the independence from interests of individual user groups. If regulatory power is assigned to a specific user group, this again leads to the situation portrayed above, in which certain user groups are less well served by the design of the statistical infrastructure than other groups. The second criterion is technical competence in the production of statistical data. If emphasis is given to the systemic nature of the statistical infrastructure, the relevance of regulatory power does not just lie in identifying users' needs and assessing the importance of these to society, but also in achieving synergies in the production of data.

It thus becomes clear that statistical offices should be the first to be considered for this role. They have the technical competence to adapt users' wishes to the statistical infrastructure and disseminate proposals for the design of the system. This comprises - in terms of the postulated systemic approach - that the role of the regulatory body is not just to organise primary surveys but also to exert influence on the data products used in the context of secondary surveys. Admittedly, they do not generally fulfil the first criterion of independence. In many countries NSOs report to a ministry, the prime minister or the president. Countries with an NSO that is not an agency of the government are seldom.² It could, however, be countered that the potential problem concerning independence is defused to a great extent by the fact that official data producers in democratic countries are bound by strict political neutrality. In so saying, dependence here is not primarily viewed as the danger of possible partisanship or even the manipulation of data in favour of certain interest groups, but can also lie in the choice of questions used as the basis of surveys. In this sense, the problem could undoubtedly be addressed in a more appropriate fashion, if official statistics were given a similar status to that of central banks.

The final issue to be investigated is the extent to which the production of statistical data can be justified as a responsibility of the government. The point of departure here will be to contemplate whether statistical data production should not actually be organised on a competitive basis, thus providing a strong incentive to find an appropriate solution to the problem of operationalization which benefits the whole of society.

One specific point to consider in this context is the issue of innovation in the production of statistical data. In addition to a number of other fields, such as the development of new fields of enquiry, the design (type) of survey or the processing and dissemination of anonymised microdata, operationalization is also an area in which a potential for innovation exists. The introduction of some kind of competition between data producers might therefore be helpful to lift innovation and to provide numerical concepts that match the needs of users at the level of more specific theories.

The issue of innovation and provision of tailor-made solutions for specific users is, however, not the only factor that has to be taken into account. Market failure theory informs us that data from the statistical infrastructure can be ascribed the characteristics of public goods, because statistical data can be consumed on a non-competing basis. Whilst the exclusion principle can be applied, it is generally agreed that data from the statistical infrastructure should be available to all groups of society and can thus be regarded as public goods *per definitionem*. Given the expectation that some parties will behave as free-riders and not contribute to the public good, it is therefore likely that if data provision were purely private there would be under-investment in the statistical infrastructure.³

² A remarkable exception is Mongolia. The NSO of Mongolia is an independent agency under the supervision of the Parliament. See <http://unstats.un.org/unsd/dnss/cp/searchcp.aspx>.

³ Aspects of *public choice* and the economics of information may also lead to the possibility of over-investment. But this issue cannot be discussed here.

6. Conclusion

The preceding discussion has shown that data users' assessment of data offered by the statistical infrastructure is closely related to the policy-making process which leads to this infrastructure. The institutional independence of statistical offices is seen as an important prerequisite for improving the efficiency of official data production and to equilibrate the horizontal and vertical dimension of the statistical operationalization problem. In addition, the peril of under or over-investment into the statistical infrastructure can provide further justification for official data production.

This has provided the initial starting point for a political economy of statistics. Further research should, however, take the discussions to a higher level of sophistication since, for example, this paper did not look into the fact that within the paradigm of the new political economy (public choice) official data producers also work in their own self-interest. On the one hand, these data producers are organised in the form of government authorities; on the other hand they supply socially-relevant groups with statistical information and thus supply a service which has a inherent scientific basis. Their interest should not therefore simply be portrayed as that of “budget-maximising bureaucrats” (Niskanen, 1968) nor of economic policy advisers (Heine and Mause, 2004). There is, however, a need to identify the personal interest of official data producers for the purposes of the research object propagated here.

References

- Brigdman (1927). *The Logic of Modern Physics*, New York.
- Dillman, D. A. (1996). “Why Innovation is Difficult in Government Surveys,” *Journal of Official Statistics*, 12 113–124.
- Eltinge, J. L., Biemer P. P., and Holmberg A. (2013). “A Potential Framework for Integration of Architecture and Methodology to Improve Statistical Production Systems,” *Journal of Official Statistics*, 29 113–124.
- Fellegi, I. P. (1996). “Characteristics of an Effective Statistical System,” *International Statistical Review*, 64, 165–187.
- Heine, K., and Mause, K. (2004), “Policy Advice as an Investment Problem,” *Kyklos*, 57, 403–428.
- Hirshleifer, J. (1971), “The Private and Social Value of Information and the Reward to Incentive Activity,” *American Economic Review*, 61, 561–574.
- Hirshleifer, J., and Riley, J. G. (1992), *The Analytics of Uncertainty*, Cambridge.
- Mueller, D. C. (2003), *Public Choice III*, Cambridge.
- Niskanen, W. A. (1968), “The Peculiar Economics of Bureaucracy,” *American Economic Review*, 58, 293–305.
- Pearson, E. S. (1955), “Statistical Concepts in the Relation to Reality,” *Journal of the Royal Statistical Society, Series B*, 20, 204–207.
- Penneck, St. (2013), “Discussion on Systems and Architectures for High-Quality Statistics Production,” *Journal of Official Statistics*, 29, 187–192.
- Reamer, A. (2010), *Putting America to Work: The Essential Role of Federal Labor Market*, Working Paper, Brookings Institution, Metropolitan Policy Program.
- Sundgren, B. (1996), “Making Statistical Data More Available,” *International Statistical Review*, 64, 23–38.
- United Nations (1992/2013), *The Fundamental Principles of Official Statistics in the Region of the Economic Commission for Europe*, New York.
- Vanberg, V. (1994), *Rules and Choices in Economics*, London and New York: Routledge.