Some Future Challenges for Czech Official Statistics

Edvard Outrata[1] | *Former President of the CZSO, Prague, Czech Republic*

Abstract

At its centenary, Czech official statistics is facing several challenges. Some are well known, others are emerging in a new and rapidly changing world. The Czech Statistical Office and other parts of Czech official statistics have a good record of coping and are well prepared. The time ahead, however, may be harder than the past. The acceleration of change in the world at large emphasises several challenges in front of official statistics, see relevance and education, as well as a number of technical developments like big data and changing means of communication. The main issue, however, will probably be the need to protect the independence of official statistics from vested interests in a world where automatisms replace specific decisions and where statistics are used as controls.

Keywords	JEL code
Czech official statistics, challenges, dangers	*Z00*

INTRODUCTION

We are celebrating a century of Czechoslovak official statistics these days, as well as 120 years of official statistics in the Kingdom of Bohemia. While the development of statistics in what is now the Czech Republic has, of course, a longer tradition both in the contributions of other institutions and in the context of imperial statistics in Vienna, these two dates are salient, because they mark the institutionalisation of official statistics as an autonomous entity in our country´s civil service and start a tradition of professionalism and trustworthiness that its present avatar, the Czech Statistical Office, builds upon.

The growth and professional development of official statistics in the 19th and 20th centuries were spurred by the self-confidence that enlightenment brought to the state and its civil service, namely that it is both right and possible to govern by taking decisions on the basis of verified facts (as opposed to hunches and feelings). This was later called "evidence-based decision-making". The concentration of official statistics in separate institutions created, in turn, further demand for the development of statistical theory and methods, by whose application statistics grew in quantity and quality. The modern state now requires a comprehensive description of the country and its people in all relevant aspects. It uses this picture and its dynamics constantly and routinely for the growing number of decisions it must take in the process of governing in a multiplying number of areas and in expanding detail. Official statistics, supplying data to evidence-based decision-making, has become a foundation stone of the process of governing.

As a result, advanced statistical offices are not run-of-the-mill civil service institutions. Specialised statistical work requires high proficiency both in up-to-date statistical theory and in the management

---

[1] Member of the Czech Statistical Council. E-mail: e.outrata@centrum.cz.

of country-wide activities. More than other state institutions, its effectiveness requires a high level of credibility both in protecting data provided by respondents in confidence, and in publishing true and reliable results. This credibility is required not just as a moral principle, but particularly, also as a condition of its proper functioning. Respondents could not be trusted to provide true information, if they did not trust the institution; nor would statistics from such data be of any use.

Czech official statistics has managed to maintain a high professional standard throughout its history. The fact that this standard was re-established early on after the fall of communism goes to the credit of the founders of Czech official statistics and their teams before the second world war, as well as to the professionalism of the staff and some of the leadership of the institution, who did not allow the ethos of the profession to disappear under the pressures of fifty years of totalitarian rule.

1 ACCELERATION OF CHANGE

The world, whose accurate description is the objective of official statistics, has always been changing. The rapidity of these changes has been continuously increasing at least since the period of enlightenment, during which the idea of beneficial change was promoted, based on growing knowledge and verified information. Official statistics was thus both an object and one of many subjects of this change and has been developing in the process; and keeping pace with these changes has always been a necessity and an important challenge. However, since about 30 years ago the number, rapidity and profoundness of changes in the world have advanced to the point where new types of challenges have developed and official statistics must cope with them. I see four categories of such changes, each posing distinct new issues, namely technological change. changes in human perception, shifting relevance, and changes in the process of governance.

Progress in information technology will probably continue to revolutionise all aspects of official statistics, data collection, data processing and dissemination. It will continue to spur methodological development (for example in the use and processing of big data).

As things stand at this point in time, in my opinion the following challenges facing the Czech Statistical Office and official statistics in general seem to merit consideration.

1.1 Relevance

One of the eternal challenges of official statistics is to provide its users the information they really require in a changing world. This problem is compounded by the need to maintain time series of indicators, which poses a limit on methodological improvement and serves as a conservative break on change. The need to harmonise statistics over large areas in order to allow horizontal comparisons and supranational statistical time series has been with us for a long time, but the need has grown more urgent with the integration of the European market. Reducing a statistic to only those components which can be harmonised can be dangerous, because it can limit the information the statistic provides. An example was the exclusion of housing prices from the harmonised CPI, which was necessary because of the differences in taxation schemes and living style patterns in different nations, but which reduced the relevance of the CPI as a practical measure of inflation for some important purposes.

The body of official statistics is fully anchored in the industrial age and is still largely calibrated to serve the needs of this age. Its product is at its most reliable where the measured phenomenon can be counted like beans. This is largely the case in population statistics and in statistics describing primary and secondary industries (or, more precisely, the legal market in primary and secondary industry products). It is also these areas where sampling methods are most effective in saving resources while maintaining a minimal and well measured loss of precision.

The product of services is much harder to measure in this fashion, so it often has to be roughly estimated. While this was acceptable in the past, the growing importance of the tertiary sector is making

this situation awkward and may start to affect the reliability of some statistics (like the GDP). Our efforts in combining rough estimates in sectors that cannot be precisely measured with detailed precision in traditional sectors may sometimes seem similar to those of the man in the old joke, who was searching for a lost penny not in the dark where he had lost it, but under the lamppost where he could see much better. In a similar way (only more so), the unofficial economy and production that does not reach the market (like housework) can only be estimated very roughly.

More important than these and similar well known and recognised problems in the body of official statistics are new problems of relevance which will crop up as society develops. Official statistics will be required to measure new features of life that were not relevant before. The recent expansion of data production and availability may make this possible, but at the same time, opens the question of the intrinsic value of data and its dynamics. Information has become a valuable asset in its own right, so methods should be developed that measure its state and dynamics, as is the case with other assets. Answers to these requirements will probably lead to further demands for new statistics. Developments in this area may lead to new important challenges for official statistics.

More down to earth is the fate of the gross domestic product. As is the case with other popular, and therefore sensitive, indicators (CPI, unemployment), GDP is being interpreted and used in a much broader sense than its methodology justifies. In popular understanding it has all but replaced the gross national income as the ultimate measure of economic performance and is even generally used as a measure of wealth, prosperity and general public happiness. The misappropriation of GDP for these purposes is, of course, the result of a growing call for an indicator that would somehow quantify the vague feeling people have of some country or some time period being happier and nicer than another one. This feeling does not necessarily correspond to measures of economic performance, so the use of such measures (usually GDP) is clearly wrong. Attempts to define another, true, measure of happiness which could then be used to compare not only countries horizontally, but also country dynamics, have had only limited success. Whether statistics can ultimately satisfy this demand at all, is uncertain, even doubtful, but the demand for it is always with us and will remain a challenge for official statistics. It is an important challenge, if only to allow the GDP to drop its transcendental burden and return exclusively to its true and important, purely economic, sense.

The greatest number of challenges that confront official statistics in the area of relevance address, however, its ability to satisfy demands from users whose needs reflect changes in the world around them and who need to adjust their function to these new facts. The form and scope of these is impossible to predict. Yet, these requirements will grow with time and will demand constant re-evaluation of the body of statistics that are being produced. In order to be prepared to cope with these challenges, official statistics must continue to develop its profession and to open itself to ideas and initiatives from users and academia.

1.2 Education

The Czech Statistical Office has maintained close and fruitful links with the Faculty of Informatics and Statistics of the University of Economics and with other relevant university faculties, thus keeping up to date with statistical theory and knowledge, providing academic institutions with practical know-how and participating in the training of young statisticians. Its challenge in the area of education, however, extends further.

Official statistics has always been confronted with how many interpretations of its important results are counter-intuitive. Some of these apparent paradoxes are classical and appear as a warning in statistical textbooks; there are also several good monographs devoted to this phenomenon. The untrained human brain seems to be relatively bad at understanding the contribution of probability theory, and so untrained intuition leads often to mistaken conclusions.

While the requirement to counteract these misunderstandings has always been with us, this need has become more imperative in the last few decades. This is so because of at least three factors. Firstly, statistics

have probably become more widely used in general political argumentation, thus giving more chance to mistaken or mendacious misinterpretations. Secondly, methodological development and harmonisation of statistics have led in many areas (unemployment, inflation) to the parallel publication of indicators describing the same phenomenon using different methodologies. While this is, of course, a positive development allowing to choose the most fitting model best suited to each immediate requirement, the parallel publication of figures describing the same phenomenon but differing in value is often confusing for the general reader, who might not understand the methodological differences. And thirdly, the general spread of data and information enabled by the recent advancement in information technology has given space to organised misinformation and mendacity, which can be supported by deliberate statistical falsifications constructed to feign scientific evidence that does not exist in reality.

The only reliable protection against toxic impacts of false information is raising the standard of general education in the relevant fields. This is clearly a major challenge for the powers that be in democratic countries in their fight against "fake news" and general misinformation. Official statistics should address this challenge as its own, and in cooperation with schools and other educational institutions look for ways of teaching the general public the basics of statistics. Reducing ignorance is the only way to counter the twin dangers of general misbelief in published statistics, and facile belief in all figures that appear to be "scientific".

The record of the Czech Statistical Office in the area of educating the public and popularising official statistics is actually very good. Apart from its approved statistical programme, it publishes many studies and monographs of topical subjects and, by far not least, publishes an attractive monthly magazine for the general public, which has won awards for its superb quality among similar publications by Czech public offices. General media have also helped by publishing some methodological detail of statistics they use. (While these details are often tedious, their regular appearance reinforces a general understanding of the dangers of misinterpretation.) All that acknowledged, there is still a long way to go before the immunity of the public against the misleading use of statistical data can be taken for granted.

1.3 Data protection and record linkage

Record linkage between surveys, and particularly between surveys and administrative data, even where technically and methodologically possible, was often rejected in the past century, because of its implications for data security. Official statistics and other public offices that collected information for official use adhered to strict promises required by specific laws not to allow the use of the data they had collected for any other purposes than those defined in the laws. This guarantee was seen as an important barrier to deliberate or accidental misuse of collected information. It also established full responsibility for this protection in each case with one public office only. Toward the end of the 20th century, with the growing amount of information collected by government offices, it became obvious that this arrangement is in many instances very wasteful, because it often required the same information to be submitted in different format by the same respondents to several government institutions. That led to legitimate protests by respondents and the public at large, who correctly believed that information once given to the government should be made available to all legitimate users by the government, which then should not come around asking for it again with another form under the heading of another department. The public position on this issue, while somewhat schizophrenic in the minds of official statisticians ("don´t ask me for the same information a second time, but never allow it to be used for any other purpose!") was very strong and, obviously, fully justifiable at a time of continuously increasing response burden.

In the end the obvious solution was found based on the arrangement where official statistics received in principle access to all government data while guaranteeing protection at least at the same level as provided by the other parties. The legal objections were also overcome due to the fact that protection of data in the Czech Statistical Office was already at least as strict as elsewhere and was seen to be so.

Recent technological development allows to build on the availability of data and to search for the application of big data methods to extract further information from data that already exist in the public or government domains. The development and application of such methods in the statistical programme is an interesting new methodological challenge for official statistics in the future.

1.4 Big data

New technological advances have been expanding the amounts of information that is available in the retail and production processes. This raises the possibility in the future for official statistics to get this information directly from the original source. If successful, this approach should eliminate the costs of reporting and thus do away with much of the response burden. The design and implementation of such systems is a major challenge for future official statistics and could in due course radically change the whole pattern of data collection.

The challenge of big data lies no longer in its technical aspect, but rather in overcoming problems of cooperation between official statistics and private business. It raises the problems of data protection to a new level and may bring along methodological concerns where it will need to combine information from different types of sources.

Czech official statistics has already started to address these issues in the collection of price data in retail, where the advantages are most obvious and where the data protection issue is least important. Broader applications will be much harder and constitute a major challenge for the future.

1.5 Register quality

Czech official statistics have long been challenged by the inadequate quality of registers maintained by other government offices. While this is basically a technical issue, it is very frustrating that efforts to minimise response burden and perform record linkages, which have finally become legal after much additional effort, are hampered by register quality and, generally, by difficulty in building common teams with other government institutions to make the registers fit for running important surveys. The prime example here is the decennial census which, as it turns out, again cannot be fully run from existing popular registers because of incompatibility or even poor quality, and so will again require the additional effort and cost that this engenders from both official statistics staff and the general population. The census is, of course, not the only instance. A great challenge for Czech official statistics remains, therefore, to facilitate and implement a working system of reliable registers together with other government departments.

1.6 Automatisms and changes in the decision-making process

There are several factors in the present world that cause the decision-making process to be more complicated than it used to be only a few decades ago. Firstly, amounts of data available to the decision makers and relevant to the issue at hand have increased significantly. Secondly, the number and sophistication of vested interests have multiplied, increasing the complexity of decision-making processes they try to influence. Thirdly, globalisation on the one hand and democracy on the other, while opening the world to many new endeavours and possibilities, require in many areas more rules and regulations, which in turn complicate the decision-making process. Fourthly, technological progress keeps extending these endeavours and possibilities, leading to further regulations in newly invented areas. Fifthly, the growing propensity to litigation seems to make matters worse yet. All these factors combine to make deciding much more onerous than before, both in the number of decisions required, and in the complexity of each of them.

In this situation it is reasonable to reduce the number of necessary decisions by grouping some of them, designing a rule that covers a series of decisions and allowing the rule to apply automatically, instead of deciding each time separately. Such a general decision can, of course, be changed once the situation changes (or exceptions may be approved), but such a change is elaborate and time consuming, so will often be avoided or postponed where and when possible.

This idea is actually very old and serves its purpose (avoidance of lengthy negotiations and litigation where the pattern of arguments is the same and where a general rule can be negotiated in the first place). One of the oldest and established examples is the indexing of payments to some measure of inflation, e.g. of rents or pensions to the consumer price index. Necessary negotiations no longer occur every year but are performed once for a longer period and then followed automatically. The parties to the negotiation effectively delegate their annual decision to an indicator produced according to a pre-defined methodology (often not fully understood by the parties but trusted to be neutral to their interests). The indicator is usually produced by official statistics.

This approach is becoming very common in the developed world and particularly in the European Union in many areas (e.g. the Maastricht criteria). Thus, the product of official statistics is in many cases becoming the arbiter of conflicting interests, rather than only a measure of some societal phenomenon. Almost by stealth, the character of the statistical product is being transformed.

The danger of this change has been recognised in the 1970´s, when Goodhart´s Law was formulated by several authors at about the same time. It may best be expressed as follows: *"an indicator used as a control ceases to be a reliable statistic"*. This had already been obvious to managers in the planned economy of communist countries, who found that the "gross material product" indicator suffered unexplainable distortions once it became the main control of factory performance used for the allocation of quarterly salary premiums. (Without the experience of communist economy, it took the authors of Goodhart´s Law a bit longer to formulate it.) The reason why Goodhart´s Law applies is clear: once the indicator has become a control, vested interests will find a way to satisfy the formal definition of the indicator with the least necessary cost. This is nearly always possible, because nearly all statistics are designed as models on the assumption of neutrality. In other words, once there is a strong enough incentive to find ways around the control, such a way will be found. As a result, the statistic might lose much of its relevance.

Official statistics tries to defend these indicators by systematic improvement of their definition (as can be seen in the history of GDP) or by inventing new indicators when old ones are compromised, all with mixed and diminishing success. The deterioration of the relevance of the affected statistics due to Goodhart´s Law is a growing challenge for official statistics in the future.

The greater danger following from the replacement of individual decisions by automatisms, however, is a shift in the object of manipulation by vested interests. Today in most instances, the pressure of lobby groups is still being applied to the actual decision makers, politicians or civil servants. This did not change with the early instances of automatisms, as (e.g.) in our example above, nobody would try to influence the method of computing the CPI instead of putting pressure on the parties to the general indexing rule. This, however, seems to be changing with the multiplication of fixed automatisms today. The more such rules are being established where the statistic itself decides money flows or other advantages without the direct participation of the original decision makers, the more practical it will be for vested interests to invest in manipulating the statistic-cum-control directly. While in some instances this pressure may result in simply circumventing the control by satisfying the indicator by other means (thereby "only" compromising the relevance of the indicator), in other cases it may be directed at the definition of the indicator. Official statistics will probably be the target of illicit pressure of this kind more often than in the past. Standing up to this pressure will be, in my opinion, the greatest challenge official statistics will meet in the near future.

2 RESPONSES

The Czech Statistical Office as the main representative of official statistics in the Czech Republic is mostly well prepared to meet the challenges of today and, hopefully, of the near future. Its long tradition of reliable service over more than a century has built a position of trust in the nation, which bodes well for the future.

Some of the challenges listed above have been recognised for some time and are being addressed already, albeit with mixed success. Many of them require more cooperation with other parts of government or public institutions. This is probably the area where improvement is most urgent. The full use of existing

administrative data in the system for statistical purposes is still being hampered by incompatibilities of different kinds (particularly of registers), which are incomprehensible in the age of sharing information on the internet, social networks, blockchain systems and big data methods. Unfortunately, this problem cannot be resolved by official statistics alone, but official statistics must persevere in its efforts to overcome the "Chinese walls" between departments.

The problem of relevance in this rapidly changing world is common to the global statistical community and must mostly be pursued together with official statistics in the European Union and other developed nations. The Czech Statistical Office is already an integral part of this work.

While in some areas the Czech Statistical Office is a leader in popularising the function and role of official statistics, much still remains to be done in educating the general public to understand statistics, their strengths and limitations. In this time of ubiquitous proliferation of false information and fake news, the ability of the public to discern and reject misinformation disguised as serious statistics may become very important. The Czech Statistical Office should become a prime mover in providing this education.

By definition, the challenges that we are least prepared for are those we did not anticipate. In the rapidly developing world we live in, one must expect new challenges to appear regularly. It is therefore imperative that the Czech Statistical Office should actively monitor not only developments that have reached the point of implementation, but rather that it should be aware of developing ideas as soon as they emerge. In this area the Czech Statistical Office maintains fruitful contacts with statistical faculties at universities and some other statistical institutions. However, many new developments relevant for the future direction of statistics appear outside the statistical community in various governmental or private research and development bodies as well as other faculties of academia. To prepare for unexpected challenges, official statistics should extend their contacts with such institutions.

Most important and yet not very visible is the defence of the professional independence of official statistics and the quality and unbiased nature of its product in this period of growing pressures of vested interests. The Czech Statistical Office is relatively well prepared for this continuing challenge by legislation, professionality of its staff and leadership, its relatively elevated position in the government hierarchy, its past history and particularly by the trust of the people it serves. Whatever the future brings, this asset is the most important of all.

CONCLUSION

At its centenary Czech official statistics are well prepared for the times ahead of us, as far as we can see today. For future challenges awaiting it in the times ahead, it can build on its tradition and history and on the professional staff it has collected and trained. We may expect it to be as successful in its second century as it was in the first.

While maintaining the optimism this conclusion gives us reasons for, it is in my opinion important for official statistics to be aware of the growing danger to its essential and cherished independence caused by the practice of using statistics as controls in complex rules-based structures in the modern world. This is the more so, because the plethora of professional and technical challenges that need to be addressed, important in their own right, may obscure this greater jeopardy in the minds of our statistical community.

References

GOODHART, C. A. E. *Monetary Theory and Practice. The U.K. Experience.* London: Palgrave, 1984.

GOODHART, CH. Problems of Monetary Management. The U.K. Experience. In: COURAKIS, A. S. eds. *Inflation, Depression, and Economic Policy in the West*, 1981, pp. 111–146.

OUTRATA, E. Influence of Governance Issues on the Quality of Official Statistics. *Proceedings 59th ISI World Statistics Congress 25–30 August 2013*, Hong Kong, 2015 (Session IPS071), pp. 669–674.

STRATHERN, M. *Improving Ratings: Audit in the British University System.* European Review 5, 1997, pp. 305–321.