DOI: 10.30680/ECO0131-7652-2019-10-8-34

Generating State Statistics: A View from «Below»

MOLYARENKO, O.A., Cand. Sci. (Sociology), National Research University Higher School of Economics, Moscow

Abstract. Official statistical information is regularly publicly criticized for the incorrect representation of social, economic, and other processes and phenomena. The purpose of this study is to analyze the quality and causes of distorting quantitative data accumulated by the state officials at lower levels – direct executors possessing information about the actual methods of collecting primary and administrative data. More than 270 in-depth interviews with state and municipal officials, heads of municipalities, state employees, and municipal institutions were conducted in 13 regions of the Russian Federation. At this level, three basic groups of restrictions on the official statistics were revealed: 1) excessive simplification of reality by the state optics, 2) systematic unintentional shifts in collection and capture of indicators, 3) intentional falsification of statistical indicators.
Keywords: official statistics; departmental statistics; distortion of statistics; falsification of statistics

Introduction

The problems of quality, accuracy, timeliness and fullness of official statistical data become real from the moment of its inception and first use as an information basis for management decision making. The study of special literature on this problem (in more detail see [Molyarenko, 2016]) shows that the requirement to understand and remove methodological, conceptual, and other restrictions of the official statistics is rather acute for many governments. Special attention is paid to it in the developed countries. Examples of discussing 'sensitive' issues[1] were studies of the methodological audit of official statistics and its further reform in China [Holz, 2003], meticulous analysis of methodological restrictions in various statistical tools in European countries [Stiglitz et al., 2016].

The basic problems of official statistics are in a way timeless for Russia –both the imperial and Soviet systems of the public collection

[1] One may assume that reasons for ignoring problems in this field may not be explained solely by required large-scale costs to reform the system of statistics. Some managers do not wish to accept the reality of lack of 'objective knowledge' as this may implicitly signify the state's weakness.

of quantitative data featured large-scale doubling of indicators, incompatibility of techniques (e.g. county, province, and central statistics – before the revolution; inter-sectoral incompatibility – in the soviet time), politicization of statistical results and other dysfunctions (more on this [Molyarenko, 2016]).

Opinions voiced in the mass media[2], our field interviews with local officials and other citizens testify that a growing consensus of the public opinion perceives government programs and policies as largely inefficient. Thus, teachers and doctors complain that the salaries have not reached the level planned by the May 2012 decrees. It is noted that the statistical growth was purely formal – by transfer of stimulating bonuses to the base part of salaries, through an increase of workload (teaching hours, number of patients) per wage and reduction of "extra" medical or teaching personnel. In such a way, the real salary of some doctors and teachers with the same or higher workload went down. At the same time, heads of hospitals, school directors, and corresponding regional ministers report the successful execution of the May decrees[3]. It seems that in such conditions the subject of quality of governmental statistics, as well as the question of accuracy and entirety of the informative basis as a foundation for development and implementation of management decisions (not excluding the potential of methodological audit of statistical activity and coverage of methodological and conceptual restrictions of the official statistics), become very urgent indeed.

---

[2] *Kotlyar P.* Correct or fiddle with: what is happening to the scientists' salaries // Gazeta.tu 24.04.2019. URL: https://www.gazeta.ru/science/2019/04/24_a_12319891.shtml; *Lomskaya T.* The Russians started to die more often from rare diseases and unknown reasons Vedomosti. 11.03.2018. URL: https://www.vedomosti.ru/economics/articles/2018/03/12/753159-rossiyane-chasche-umirat?utm_source=vk&utm_campaign=share&utm_medium=social&utm_content=753159-rossiyane-chasche-umirat; Arms twisted as forensic anatomists are made to change diagnoses for indices // Meditsinskaya Rossia. 17.08.2018. URL: https://medrussia.org/20446-patologoanatomam-vivorachivayut-ruki; Experts evidenced "gross exaggeration" in punishment for drugs in Russia // RBK. 28.09.2017. URL: https://www.rbc.ru/society/28/09/2017/59ccbc489a79473ad83cd991 (the date of access to all sources: 03.06.2019).

[3] We note that in point of fact the regional authorities are forced to falsify data– raising salaries has become for them a not financed necessity from the federal level, none of the federal subjects had means to do it. See, e.g.: *Togayeva L., Zubarevich N.* We observe an emergency and lying management system // Republic. 17.02.2014. URL: https://republic.ru/posts/l/1057459. How Russia tries to catch up to the May decrees of Putin // DW. 16.03.2018. URL: https://www.dw.com/ru/как-россия-пытается-угнаться-за-майскими-указами-владимира-путина/a-43001633 et al. (the date of access to all sources: 29.05.2019).

Methods of research

A research that is restricted by standard statistical tools is not capable to resolve set questions – such methods would have kept us in conceptual and methodological blinkers leaving space just for establishing technical failures and proposing possible variants for their elimination. This is the reason why we employed the methods of in-depth interviews and overt observation. The use of high-quality field methods to verify official statistics was repeatedly and expectedly criticized by statisticians, economists, demographers. However, one should not overlook the concept of 'newcomers' in science – quite regularly substantive breakthroughs are made by scientists that resort to unconventional ways of cutting through subject areas (applying methods of one science to study another). In this context and evaluating the obtained results we believe that using quality sociological and anthropological tool sets to unveil restrictions in the official statistical information was validated.

Information was collected in the course of field trips. The empirical base of research is described below.

In-depth interviews with 181 people from 83 municipalities. In-depth interviews in the Lipetsk, Kaluga, Vologda, Tyumen regions, and republic of Komi were part of *a specialized project on the quality of official (aggregated by Rosstat, federal organs of the executive branch, regional authorities, municipalities) statistical information* in 2013–2014[4]. The main groups of responders: state and municipal employees (municipal administrations, Rosstat departments, federal tax service inspections, departments of Directorate of the Federal Migration Service and Ministry of Interior, departments of Rosregister, Land Cadaster Chamber, federal and regional treasury, heads of municipalities, employees of state and municipal offices (centers of employment, social protection, the pension fund of Russia, municipal schools)).

In-depth interviews with 84 responders in 29 municipalities in Pskov and Omsk regions were conducted in 2016–2017 as part of *the project on the nonformalized (ownerless) communal property*[5]*, which due to legal non-existence is not always recognized by*

[4] The project is carried out with financing from the Fund of social research support "Khamovniki".

[5] This project is also implemented with financing from the Fund for support of sociological research 'Khamovniki'.

the official statistics. The principal groups of responders: state and municipal employees (municipal administrations, regional authorities by profile – housing and utility services, property relations), heads of municipalities, state office employees (forestry departments).

Additionally, in 2014–2019 as part of student research trips of the chair of the local self-government of National Research University Higher School of Economics, there was selective verification through in-depth interview with state and municipal employees thanks to the university-wide program "Opening Russia anew" (URL: https://foi.hse.ru/openrussia). Additional questions about the fullness, quality, and accuracy of official statistical information in the context of municipal units were asked to the employees of state and municipal establishments in the Ivanovo, Smolensk, and Tver regions, Krasnodar territory, Nizhniy Novgorod, and Kalinin regions.

Cross comparison of the content of scientific publications[6], interviews, and observations helped form three sources of reasons for lack of informational value (low quality, inaccuracy, incompleteness) of official statistics (three basic restriction groups) that were discovered on the level of data collections and their entry into the state system and document flow. *The first source:* inevitable but often excessive simplification, 'flattened' by the state notion about socio-economic processes for the sake of their registration. Second: unintentional systematic deviation of real techniques of collection of primary data and their processing from formal, and statistic distortion due to specific structure of the administrative-territorial system of data collection, in other words, technological and technical distortions. *Third:* intentional falsification of indicators for the purpose of improved quality of effective performance of power organs and other misrepresentation of socio-psychological origin arising for reasons that statistics can to be a neutral tool.

Let us consider these sources in detail.

[6] This required a vast disciplinary variety– the most important substantive aspects were drawn from the works of statisticians, economists, anthropologists, demographers, sociologists, geographers, philosophers, experts on state and municipal governance.

Simplification of reality

The standardized tools developed by the state for the registration of socio-economic and other processes and events leads to an oversimplified image of the real life that often fails to capture phenomena that are fundamental for the socio-economic stability of actual municipalities. In a broad sense, this problem was considered by J. Scott [Scott, 1998]. With multiple examples – scientific forestry, the establishment of a single system of weights and measures, city planning, and others – he shows how the state attempts to simplify, make visible and reportable objects of its management, thus, paying attention only to meaningful features for the subject of management, while 'meaningless' – reductive – often turn out to be fundamental for maintaining ecological, economic, social, and other systems.

The emasculation of contents through quantitative tools was time and again pointed out by economists. Thus, L. Mises explained that quantitative problems in the sphere of human activity may be perceived only through comprehension (which requires qualitative methods). He criticized purely qualitative methods in economics and their supporters for not taking into account the lack of functional connections between variables [Mises, 2005].

The official statistics by means of unified for the whole country methods registers similar and approximately the same aspects and traits of events (only those of them that the state is interested in) for aggregation at federal and regional levels. At the same time, specific, unique features of municipalities and local communities that determine their socio-economic position there are outside the scope of the administrative view[7] [Molyarenko, 2016].

It should be noted that even the said uniformity of indicators is conditional as all cases require abstraction of specific details:

"Behind every figure, there are many qualifications of various kinds that may not be expressed in figures. <...> If every figure were

[7] As there is no statistical data for living on dachas or leaving for seasonal work (i.e. the state has no tools for recording pendulum, seasonal or other recurring movements of people over the territory of the country that has an impact on the quality of contingent figures, e.g. estimated numbers of the population in municipalities and regions), there are no estimates of people's income from fishing activity (berries, mushrooms, wild fruit, etc. and their sale) although in many municipalities these types of activity are the main source of income (not salaries, pensions or other that form part of officially recorded income).

followed by qualifications, the picture would have been different" (a municipal administration official of an urban district in the Kaluga region).

In other words, the centralized methodology in various ways does not take into consideration the local specifics. Meanwhile, the methodologies keep on changing (often without a transition period to ensure compatibility), and for some indicators there are no instructions on what to consider at all (e.g. the number of people that go in for sport). Problems of metadata, close nature, and variability of methodology have been addresses i.a. by V.A. Bessonov. He describes building a long time series of macro-economic indicators as "a result of trying to measure a changing object with a changing standard" [Bessonov, 2015].

This being said, it would be a mistake and idealization to presume the existence of a single system of indicators. In fact, the information goes 'upward' through a range of territorial and departmental networks. This is to say that every government agency has its own 'optics' unlike any other. Thus, there were many cases when the data of the regional office of Federal tax service and Rosstat on salaries and other economic indicators did not match:

"We have questions with Rosstat. For example, the rise in salaries. The statistics say: "This month salaries rose – by this much". <…> We call the firms: Where? Did you give a rise? – No. – A rise? – No. <…> We always have such kind of problems – nothing matches. We tell [to Rosstat]: "Where do you get the data берете"?» (from an interview with a head of Federal tax inspection in a municipal district of Lipetsk region).

"I had a case when I brought to our ministry data from a tax inspection. Things we collect every year. The amount of balance sheet profit, income… If we cannot strike the balance we take it from them. I come to Kaluga and say. Kaluga takes their data from statistics and they have a much low profit. I say "the profit we have is 3 times higher in the city. This is an authorized data from the taxman". – "No, our statistics is this" (from an interview with the head of the economic department in a city district administration of Kaluga region).

The situation is further aggravated by the fact that instead of averaging and reciprocal checking of data we have some

administrative organs cut off from collecting or using data. In 2006–2007 we had the rights of Rosstat on 'correct' information confirmed. By 2010 its virtual monopoly on statistical data for municipal records was approved. For example, in the case of a diverging figure on municipal population numbers between Rosstat and a municipal administration, it is the Rosstat data that is to be considered correct (and used for calculating inter-budget transfers). This way of resolving problems (enforcing 'elimination' of interdepartmental indicator divergence) has already been applied in our history. Thus, in 1930, growing 'tired' of permanently discrepant figures from the Central statistical department and the State planning committee, the country's leadership liquidated the CSD as a separate organ and subordinated it to the SPC.

Absolute priority given to statistical data and quantitative tools of reality registration is a standard trap for managers and scientists. According to multiple observations, when working out methodologies and constructions their creators carefully formulate existing and potential ontological, epistemic, and technical restrictions. However, as those tools are used by practitioners they are idealized, restrictions for their application are consciously or unwittingly 'erased', ignored (perhaps due to an urge for illusory knowing-all and full information awareness). As a rule, this causes serious errors and crises due to the accumulated critical mass of processes and phenomena concealed from the state and scientists.

It is this very context that is used to interpret the causes of the economic crisis of 2008–2009 by a group of economists led by the Nobel laureate J. Stiglitz [Stiglitz et al., 2016]. The scientists point out that Simon Kuznets and pioneers of the concept of the Gross Domestic Product (GDP, an indicator developed in the 1930-s) were very careful and kept on reminding themselves and people around them about the limitations of employed tools. The consequent economic management practice idealized and made a fetish of this indicator – Stiglitz and his colleagues criticize European politicians and managers who practically conflated GDP with people's living standards.

Such simplifications are representative of whole scientific schools. It appears justified that the Keynesianism presumes (and presumed) a much more transparent present, future, and ability to be computed of the surrounding world than Keynes himself believed. He, in

fact, blasted the classic economic school precisely for its illusory transparency and computability [Keynes, 1998]. Being rather a proponent of the ontological view of uncertainty[8] (although there are also statements concerning epistemological restrictions) he is critical of the followers of the epistemological approach and those implementing an excessive number of complex tools calculate the future for creating an illusion of predictability and stability[9], as in such a case *the complexity and uncertainty of reality are replaced by a complex of modes for its perception.*

G. Akerlof and R. Shiller believe that the theory of Keynes had been eviscerated, especially in the 1960–1970-s from the irrational components of human behavior as well as many statements of philosophical nature concerning the ontological opacity of the world and irreducibility of socio-economic processes to numerical indicators [Akerlof, Shiller, 2010]. To be fair, though, we should note that Keynes held that under *stable conditions* theories supporting the possibility of knowing the world and mathematical predictability of the future are relatively effective.

Keynes' convictions about the uncertainty of the surrounding world and the illusion of being able to figure it out mainly shaped as a response to the great depression in the USA. Similarly, the decidedly hard criticism of absolute priority given to numerical indexes by managers and economists [Stiglitz et al., 2016], I believe may be attributed to the global economic crisis (that started 2007). In other words, the conceptual and methodological restrictions of numerical tools of registering reality are things that, regrettably,

[8] The **ontological** representation is based on the recognition of the essential impossibility to understand the world due to permanent changes in the socio-economic environment, the lack of strict cause-effect relationships. This approach assumes no capacity to reduce the level of uncertainty and the main emphasis goes, as a rule, to neutralize its consequences. The proponents of alternative, **epistemological** (gnosiological, cognitive), representation of uncertainty consider it possible to understand the world despite the fact that no economic agent (including the state) is capable to collect all the dispersed information and hence is forced to select imperfect information. The asymmetry and incompleteness of information are the causes and errors of forecasting. To reduce the level of uncertainty, such an approach employs methods for increasing the volume and improving the quality of available information [Terzi, 2010].

It is assumed that the Russian government used the gnosiological representation and that is why it is extensively яexpanding the number of indicators and reports, while the vast gaps of vision arise from ontological opacity of socio-economic processes got the managers.

[9] It is ironic in this contest that Keynesianism introduced the use of aggregated values in economics leading i.a. to the development of quantitative methods.

economists 'recall' no sooner than when the illusion of absolute knowledge and control of socio-economic processes leads to a disaster.

Systematic unintended bias in collection and registration of data

If the first group of errors stems from the impossibility to register the full multidimensional reality through the statistical data collected by formal methods and failure to capture some key processes due to the absence of tools for their registration, the second one reflects the discrepancy of actual methods of data collection and processing with the formal ones. The point is that even the indicators under investigation are registered with significant errors. The main causes of distortion are: 1) excessive cuts of staff in regional sections of public organs (including those responsible for collecting primary and administrative data) за сбор первичных и административных данных) при сохранении преимущественно прежних методов работы; 2) redundancy and duplication of collected data with missing analysis and feedback; 3) suspension of local self-government organs from data collection; 4) serious inter-departmental information barriers. At the same time, no methodology audit of statistical activity is carried out in Russia at all (there is only its slight resemblance for census while none is done in term of methodological audit for current data collection) although this activity is crucial for the state, which is interested in the valid informational basis for its decisions [Holz, 2003].

While scientists and managers may be aware of the errors of the first group as official methodologies are regularly published by Rosstat and may be analyzed for their restrictions, the distortions caused by deviation of actual data collection methods from the formal ones are quite seldom considered or discussed. Not only for the pointed out reason of the lack of methodological audit but also because the executives (rank and file specialists) do not wish to speak publicly about their failure to observe the formal techniques even if those kinds of bias stem from objective circumstances that require correcting the tools themselves.

Collection and registration of primary and administrative data is, thus, a "black box" – neither managers of regional or higher levels nor scientists or research staff, who use official or sectoral statistics

in their work (with rare exceptions) do not attempt to understand how this data is really collected and what their limitations are.

As already has been said, an important factor is the administrative-territorial structure of the sector that captures performance (the geographical and transport projection of power organs). This explains, e.g. territorial distortions of official unemployment. The services and centers of employment keep records on the basis of applications. Persons registered as unemployed must come twice a month in person to the employment center in order to confirm their status. The unemployment benefit in 2016 was 850–4900 Rubles a month (since 2019–1500–8000 Rub.). That is why those not living close by or within approximate transport vicinity would prefer not to register in the case their transportation or other expenses exceed or are comparable with the benefit sum itself. Those that live in the vicinity, however, may register even if they are not currently looking for a job or have informal sources of income but would not mind getting an additional benefit. With other things being equal the result is over-registration of official unemployed in administrative centers of towns and districts and under-registration on the outskirts.

"– Do you have high hidden unemployment?

– Yes. They used to come here. And it was so awkward. They came to our unemployment center and the unemployment rate was higher. They stopped it – the rate of unemployment in the district plummeted. Everyone was pleased.

– Nice ways to combat unemployment…

– *Yes. They sued several times – there was nothing we could do. But then we referred to the law that the employment center does not have to go out. We are remote and that's it.*

– *Let people travel 200 km.*

– *It is 350 rules to *** [the district center]. Twice a month to visit. Both ways make this 1400. Food to buy. An unemployed gets 900 Rubles. What is then the point of going twice a month? It is as bad as that.* (from an interview of the head of a rural settlement in the Vologda region).

In other words, if the territory does not have an executioner that registers some indicators, according to the official data there are no phenomena.

Similar processes take place in the activity of law enforcement agencies. As a rule, their staff is stationed in administrative centers of municipalities. As part of the reform of the Ministry of internal affairs in 2010–2011, reductions touched mostly the low ranks [Shklyaruk et al., 2015] – officers of rural and urban police departments and local police officers. According to our observations, before 2010 an average municipal district had 15 local police officers, now – three.

Considering the required volumes of reporting, local officers often do not have time to visit administrative centers of all rural settlements and some places remain unattended during the whole service period of such an officer.

Besides the territorial factor, there is also the one that concerns the limited nature of human capacities. If (a hypothetical case) all officers work 'to the pointy of exhaustion' but the crime rate meanwhile rises 10 times, the same officers objectively would not be capable to register this: open corresponding cases, register files, fill them with data and start necessary investigations. Therefore, according to the crime statistics, no one will be able to notice a significant rise in criminal cases as no one will be there to record an indicator spike (intentional falsification will be considered in the next section, now we are talking about *unwitting* distortions).

Thus, the official statistics of various processes directly depends on the territorial structure of power organs and available executives for its recording.

Meanwhile, a detailed analysis of so-called 'optimization' processes in the territorial structures of national and municipal power organs proves that any type of headcount reduction leads to a reduced headcount of lower strata – operators dealing directly with the public, businesses, etc., i.e. those connected with reality, not 'paperwork'.

The cases of reduction in Rosstat, the Ministry of internal affairs, the Federal Tax service, municipalities have some common features. The starting point everywhere is the affirmation of excessive numbers of civil servants and non-optimal cost of their operation that calls for the need to cut staff and raise the efficiency of territorial departments. There are many cases of bypassing such measures, e.g. there is a simple method of outstaffing – transferring personnel to subordinate authorities. Statistically, the number of officials goes down, the

employees lose the status of civil servants but otherwise, nothing changes. If a reduction is inevitable, there are multiple reasons for it to concern the rank and file personnel in the first place – the already mentioned officers of internal affairs, tax inspectors[10], employees of the district and municipal departments of Rosstat[11], and other organs. Moreover, the higher executive load on executives leads to lower quality of information they produce and deliver to the management system. The loss of control and information is captured by the superiors who try to improve the situation by introducing new forms of reporting that create more work apart from their direct responsibilities and data collection.

One often sees attempts to compensate for personnel cuts at low levels with more staff at the regional level. The new branches introduce new forms of reporting for low-rank personnel. Thus, the management system is getting out of touch with reality due to the disappearing base embodied by the executives on the ground (see about this [Molyarenko, 2016]).

The overload of low-level executives, efforts to save costs of their operation, the lack of well-tuned mechanisms of interdepartmental exchange of primary and administrative data often lead to the impossibility of thorough compliance with formal methodologies. Thus, the officers of Rosstat often refer to a widespread problem of poor coverage of individual entrepreneurs by the comprehensive and selective census. The territorial departments of state statistics send down a list, which is incomplete and not up to date, with missing postal and legal

[10] Approximately since 2010, inspections have expanded. Earlier, one tax inspection serviced one municipal district, now it serves 3–5.

[11] Since 1991 the staff of Rosstat has been cut by 40%. In most territories the departments of state statistics have been eliminated as independent entities with their own functions – as a rule, the state statistics in a municipality is represented by several employees that are part of the staff of the territorial organ of state statistics responsible for collecting data and entering them into electronic forms. Prior to 2005–2007 statistic departments in districts and cities were independent legal entities that were engaged in analytical work. These analytics was compared and organs of local self-government and in a 'corrected' form sent 'upwards'. The respondents suppose that centralization of analytics led to lower quality of statistics in municipal establishments. At the same time, experts point out that aggregating information for sending 'upwards' is one of the most serious problems of contemporary Russian statistics as much of it gets lost generating new inquiries for the same indicators at lower levels. A simultaneous solution for the task of inter-departmental reconciliation and excessive aggregation could be comparison and checking of primary and administrative data collected about individuals and legal entities by various organs of power but it is difficult because of confidentiality (the laws 149-ФЗ, 152-ФЗ, 282-ФЗ, more about it later).

addresses, expired and incorrect information. We do not refer here to the shadow economy; this concerns only the registered firms. The low quality of this list is caused among other things by the separate registration of businesses in statistical and tax organs in regions without 'a single window'. The rate of response from individual entrepreneurs is quite low (a small technical detail is that in order to save on costs questionnaires are sent out by simple, not registered mail. So, many do not answer knowing that they may always claim not having received the letter. The dates of the exercise are calculated within an ideal methodological scheme "letters sent – responses provided – data entered into the system" (the problem was described in detail by Rosstat employees). Thus, the formal methodology may not be implemented and samples are essentially displaced. Such difficulties and displacements in actual data collection are actively discussed by executives themselves but never conceptualized as a problem on the level of departmental agencies and researchers working with statistics.

In this sense, the most valuable thing is the methodological audit carried out for mass surveys by operatives from the Russian Presidential Academy of National Economy and Public Administration [Rogozin et al., 2016]. It reveals that (field workers) low ranked operatives need to meet very high demands and norms of output (a number of filled-in forms) with a short completion period for fieldwork and low pay for interviews. Under such conditions, the formal results of work are sometimes improved by tricks – sometimes working with accessible responders, not randomly selected or partially filling forms themselves. This factor falls into the category of unintended falsifications as the field interviewers have no interest to manufacture survey results, they just try in the current conditions to fulfill their work so as to avoid punishment for missing deadlines and low productivity.

There were similar situations when formal methodologies were not followed due to overload, low pay, and specific labor incentives (bonuses are awarded to those meeting the quota or covering their area faster than others) at the popular censuses (2002 and 2010). Thus, in order to speed up their work many census takers who did not find people at home copied their names and other information from housing records so as to avoid second visits and miss deadlines.

Systematic unintentional distortions appear as part of the information on housing and utilities infrastructure. After reorganizing and eliminating soviet style enterprises and institutions we had lots of various types of infrastructure written off from balance-sheets such as water pipelines, sewage systems, auto roads, electric and heating networks, etc[12]. Those were supposed to become the ownership of municipalities. However, 'municipalization' of the former soviet social utilities infrastructure did not go fully according to plans due to some financial, technical, and legal reasons. According to formal norms, municipalities were to take gradual inventory and process documentation for such ownerless property but they were too busy with their current affairs and had no staff or finance to do the necessary job. That is why the actual recognition of undocumented infrastructure takes place only when it breaks down and requires repair[13]. This makes a major part of the infrastructure invisible to the state.

Thus, due to several objective restrictions, the actual techniques of data collection and registration digress from the formal ones. Significant discrepancies in territorial distribution of symptoms arise from the lack of necessary personnel, specific forms of authority structures, working conditions of operatives that register and represent this or that process in the official statistics.

Intentional falsification of statistical data

Statistics is not a politically neutral instrument as it is used for evaluating the efficiency of government programs and policies, appraising the operation of authorities and separate officials, determination of required volumes of financing (calculation of inter-budget transfers)[14]. Thus, the results of population census essentially

[12] See the groundwork for the project of 'Khamovmniki' Fund – "A constructive role of informal relations in the system of state and municipal management", URL: http://khamovniky.ru/konstruktivnaya-rol-neformalnykh-otnosheniy-v-sisteme-gosudarstvennogo-i-munitsipalnogo-upravleniya655b

[13] According to an interview with municipal employees in all of the municipalities visited within the project on the ownerless property.

[14] It should be noted that all of the described projects are not exclusively Russian. E.g. China is dealing with all of the same problems, see: China will fight hard against falsification of statistics // Trudд. 29.05.2019. URL: http://www.trud.ru/article/29–05–2019/1376326_kitaj_budet_vesti_zhestkuju_borbu_s_moshennichestvom_so_statistikoj.html (дата обращения: 30.05.2019).

prove or disprove the efficiency of national projects and policies [Pyankova, 2012].

The demographic data are subject to oppositely directed invasions and distortions. Thus, according to the interviewed municipal and Rosstat employees, prior to 2010 when the organs of local self-government in rural settlements were responsible for collecting the data of population numbers and submitting it to the statistical organs they generally inflated the data conflating the permanent inhabitants with temporary ones. Nowadays, many rural and municipal officials complain about a significant reduction by Rosstat of the permanent inhabitants' numbers. For example, rotation workers are not included (this brings down the income of a municipality from inter-budget transfers, which are calculated on a per capita basis) [Molyarenko, 2016].

"– In other municipalities they told us that during the census Rosstat brought down the real number of population. Did you suffer because of it?

– Yes. According to the records that we keep – we have 930 people, according to the census it is 720…

– Was it much in terms of financing?

– Yes. All of our established standards are on a per capita basis and we had cuts. About one million was cut at once. And then, some more.

– Did you try to protest about the results?

– Of course, we did, *** [a neighboring] rural settlement even sued. But nothing helped… Right after the census, I remember, we had a big meeting in the regional government of all the heads of our […] region. Everybody spoke about their concerns. They told us … our [region] – is one of the best in terms of the census and that is all. Our objections are not taken into account.

– When Rosstat held the census, did they cooperate with you?

– You see, the thing is… If they told us to count all there are. We would have taken the same data from the house records, right? Everybody who lives in the house and that is all. Somehow, we now believe it was done on purpose. They said we should do it exactly like this. Go and record only those who now is at home. And what if somebody is away at work or for some other reason? We have, for example, many young families, who work for a week in Cherepovets, they come home only on the weekend.

– So, even if a person was at work they were not counted?

– *Yes, yes, yes. Such were the instructions. Now we know better. So, when another census happens… But, there will probably be no settlement here any longer…"* (from an interview with the head of a settlement in the Vologda region).

At the same time, according to demographers, during the 2010 census, major cities (in order not to lose the status of 'millionaire') and the national republics (to boost the volume of federal transfers) purposefully overstated the population numbers[15].

A more detailed consideration of reasons and consequences of intentional falsification of statistical data may be found in the paper devoted to the problems of criminal statistics [Shklyaruk et al., 2015]. Forming key statistical figures takes place completely inside the system of law enforcement agencies, which are consequently appraised on the basis of the same criteria. The authors distinguish three groups of reasons for unreliable crime statistics: 1) the latency of crime, 2) selection and filtering and 3) direct falsification of registered data. The first group (as concerns the intentional distortions) arises from the fact that a number of crimes (those that will surely scale down the crime detection rates) are consciously concealed. They are not registered by law enforcement officers by not accepting an application form or by not opening a criminal case. The second group of misrepresentation arises due to the selection of a more 'convenient' qualification for a crime that has to be registered. Thus, if a crime seems to be unsolvable it should be registered as minimally 'serious'. And if there is a 'trusted' accused from a vulnerable social category (a vagrant, unemployed, etc.) and there is a high certainty to 'resolve' the case, it may be qualified as a maximum weighty clause (in order to raise the solvency rate of particularly serious crimes). The third group is related to the direct falsification of data in statistical cards that are used for the preparation of departmental statistics.

The research into manipulation with data, which are aggregated and transmitted for evaluating the efficiency of operation of local self-government organs and regional authorities demonstrates that negative changes are not covered sufficiently while positive are

[15] *Mkrtchan N. Dead souls* // Forbes. 30.03.2011. URL: http://www.forbes.ru/ekonomika-column/65636-mertvye-dushi (the date of access: 27.05.2019).

blown out of proportion [Kalgin, Eliseenko, 2015]. A.S. Kalgin distinguishes two basic strategies of manipulating figures "careful" and "daring" [Kalgin, 2016]. The first targets avoiding the attention of higher authorities and that is why the bureaucrats 'normalize' figures by flattening the curve and submitting minimal changes in values. The second one is based on the regular boosting of figures that put the work of the authorities and particular officials in a positive light. In such a case the most vulnerable are the 'honest' employees who record real figures as they look most depressing in the general statistical picture. Thereby, in the course of time, even the initially responsible public servants start manipulating statistics in order not to look 'the worst' [Kalgin, Eliseenko, 2015].

Manipulations in the area of medical statistics are well-known. Thus, if the federal authorities set the task to reduce the mortality rate for a certain range of causes, the operators start classifying deaths into other causes including deaths due to unknown causes (see the paper by V.V. Yumaguzin and M.V. Vinnik in the same issue of the journal[16]). From an observations diary of a student participating in a field expedition from the National Research University Higher School of Economics guided by O.A. Molyarenko to the Nizhegorodskaya region:

"According to a medical statistician, the most frequent cause of death is a cardio-vascular disease (cerebral hemorrhage) and oncology. Many young people die from cirrhosis of the liver, hepatitis, and anemia. As for 'inaccurate' data she said that their department submitted real data to Nizhniy Novgorod ('we do not want to adjust figures'), but afterward it happened that they received orders and were forced to change data. <…> For example, there is a plan to reduce deaths from cardio-vascular diseases. If a person dies of a stroke but he had diabetes, the cause of death will be registered as diabetes. Instead of diabetes, it may be asthma. <…> From a conversation about an ambulance we learned that the hospital has six emergency vehicles but only three of them go out to answer

[16] See more about the same: [Krasheninnikova, 2011]; *Lomskaya TT.* Russians started to die more often from rare diseases and unknown reasons // Vedomosti. 11.03.2018. URL: https://www.vedomosti.ru/economics/articles/2018/03/12/753159-rossiyane-chasche-umirat; Pathologists' hands are twisted to change diagnoses for the sake of indices // Meditsinskaya Rossia. 17.08.2018. URL: https://medrussia.org/20446-patologoanatomam-vivorachivayut-ruki The date of access: 20.05.2019).

calls. The ambulance has to arrive in less than 20 minutes. The district is large, some places may take over an hour to reach. The department submitted real data until they received an order [bring the figure to the regulation norm]. "Moscow requires the figures and we submit them".

The heads of rural and urban communities in the Central federal district in an interview told about some of the 'forced' data manipulations in the area of housing and housing-utilities infrastructure. If a municipality declares the housing facilities to be decrepit or in bad repair, it must allot funds for their repair or construct new houses for rehousing. However, due to budget shortage municipalities cannot finance such programs while a high percentage of decrepit and dangerous houses and failure to take measures calls for sanctions on behalf of the regional authorities. That is why many municipalities (or rather inter-departmental commissions founded by the local power organs) prefer simply not to acknowledge such houses decrepit or in bad repair.

The cases described here show that intentional data manipulation takes place at the very least in medical, criminal, municipal, housing, and demographic statistics. In our opinion, if such falsification is not revealed, the image of the object gets further from the truth while the authorities have the illusion of improving the situation and full control over all processes since statistically they react positively to their decision making.

The outlook

We observed local success in staving off problems of the second and third groups by means of inter-departmental reconciliation of databases, merging all information on individuals and companies. Improvement of information databases sometimes takes place as a byproduct of achieving other targets. In 2016 a governor from outside came to the Kaliningrad region together with his team[17]. At the moment, the region had established local elites loyal to the former management. So, the new leadership faced the task of bringing the region under control, making it transparent and understandable. So, besides starting the reform of the municipal structure and

[17] The case is based on the interview with three regional public servants of the minister or vice-minister level from various departments and municipal officers from four of the separate municipalities of the Kaliningrad region.

replacing the heads of municipal organs, the new management began establishing the 'digital authoritarianism' (an expression by one of the surveyed). In particular, an innovation was proposed to improve targeted social assistance – a citizen applying for any kind of social assistance had to give consent for the regional ministry of social policy to obtain and process their personal data[18]. This allowed a single information system to collect the data about one's employment, salary, marital status, real estate (apartment, house, land), auto vehicles, bank accounts of the applicants. This requirement on its own allowed obtaining a detailed database for at least one-fifth of the regional population as almost 200 thousand people receive social benefits (the number of those applying for it was greater)[19]. Apart from that, the social sphere statistics became more accurate – the number of officially unemployed went down by several thousand (as a result of the exclusion of those having sources of income).

However, this format is applied in the cases of declarative principle, in other cases, information is reconciled, if it happens at all, informally on the level of field operators who exchange information about actual individuals or firms for confirmation of data in departmental databases (which is against the law).

Integration or mutual checks of primary and administrative data may turn out to be a promising mechanism that would allow reducing distortions arising from the lack of staff and failure to follow the techniques of data collection (problems of the second group). However, their massive (centralized by the state) legal application is not likely in the near future. The fact of the matter is that the state statistics had been moving in this direction until the mid- 2000-s. E.g. Rosstat and the Federal Tax Service together developed initiatives that would allow by means of data exchange to improve the quality of statistics, reduce the load on respondents

[18] According to a recently added amendment, an applicant has to bring a part of the documents himself: URL: https://www.newkaliningrad.ru/news/briefs/community/22955428-alikhanov-o-proverke-schetov-poluchateley-lgot-my-peresmotreli-podkhody.html (the date of access: 20.05.2019).

[19] URL: https://www.newkaliningrad.ru/news/community/21847485-eshche-bolshiy-khaos-chto-ne-tak-s-sotsialnym-kodeksom-kaliningradskoy-oblasti.html (the date of access: 20.05.2019).

and make some surveys cheaper. But a number of laws[20] adopted in 2006–2007 led to inter-departmental information barriers[21] and established the Rosstat monopoly on 'correct' data [Molyarenko, 2016].

According to our observations, the problem of such barriers existing is accompanied by reluctance of the authorities to lose their own databases and switch to a common one as this would mean a reduction of financing, staffing and information resources of the department, in other words, lower administrative weight [Kordonsky, 2006] and contradiction to the tasks of a maximizing bureaucrat [Niskanen, 2004]. Such concerns provoke not only anti-unification discourse but also real resistance to introduced common databases. For this or other reasons, initiatives to introduce unified databases lead to partly filled and methodologically opaque information products. The most skeptical of my respondents and colleagues suppose that 'multiverse' statistics is simply advantageous to public servants as every indicator has a number of values in various bases and depending on circumstances one may choose which one to use. Anyway, in the near future one may hardly expect a higher quality of official statistics as a result of convergence or cross-checking departmental information systems.

Conclusion

Borrowing and expanding the example of J. Scott about scientific forestry we'll describe the problems through a metaphor: the number and height of trees tell us nothing about the quality of forest and vitality of the forest ecosystem (problems of the first group). Meanwhile, the forestry service is 'over-optimized' in a way that every forester is responsible for several thousand hectares of forest, and 39 out of 40 hours of his working week he has to spend not on the essential work but on writing reports for a dozen superior authorities, statistical and controlling organs (which do not reconcile the data they get among themselves). That is why

[20] Federal law from 27 July 2006 № 152-ФЗ "On personal data", a federal law from 29 July 2006 № 149-ФЗ "On information, information technologies, and data protection", from 29 November 2007 № 282–ФЗ "On official statistical accounting and the system of state statistics in the Russian Federation".

[21] It is prohibited to exchange data on individuals and legal entities without their written consent.

for objective reasons he may not follow formal methodologies that require, say, a daily beat of the forest and he is forced to collect information by other methods and base his case on intuition rather than calculation[22] (problems of the second group). If this system evaluates the efficiency of the forester's work and his salary level on the basis of how many trees his territory has, it is only logical to expect shenanigans and data falsification (problems of the third group).

Thus, judging by the look from 'below', the official statistics loses its function of reference, there is a progressive gap between the indicators accumulated by the subjects of statistical record-keeping and the image of reality formed on their basis. The 'dead zones' (processes, objects, their aspects and features that are invisible for managers) are more or less found in all spheres of society and government – in the evaluation of economic activity, social status, and living standards of the population, in the recording of delinquency and crime, in the registration of communal property. In the meantime, our field research allows us to conclude that inconsistency at the stage of collecting (recording and registration) administrative and primary data leads to essential incompatibility of many statistical indicators time- and territory-wise: the majority of intertemporal and inter-territorial comparisons are methodologically incorrect due to the highest heterogeneity of real methodologies of obtaining and processing data.

The closest to reality and relatively compatible are the indicators of the budget sector itself (the number of state and municipal offices, their employees, the number of students in the state and municipal educational establishments, the number of beds in hospitals, the data

[22] *"There is a problem, I think, that no one in the city has full information about all the indicators we are asked to submit. Starting with the population and up to revenues and investment... we don't even have information of calculating land tax for its payers ... the only data we have is based on estimates. We make them by adding to the previous figures. We have worked here for 30 years. We call them "out of thin air"... earlier, in the 90-s – beginning of 2000-s we argued with the ministries, trying to clarify the figures. Now the Ministry of economics accepts the figures if it's on par with the regional average. Nobody can make evaluations – not us, not them...*

We want to wait till retirement and leave this madhouse ... you can hardly sleep at night, you just keep thinking about making things up... It is a creative process. Right. It is not just processing information. This would have been easier. Unfortunately, there is no information, one has to supplement it. What you make is what you have". (from an interview with a municipal employee of an urban district of Kaluga region).

on services rendered by the state, etc.), i.e. the features of actions and property of the state itself[23], and less – the information on the non-public sector of the economy, about demographic and social processes.

However, the problem does not lie in the fact that the statistical image of the object is not ideal (for a variety of restrictions it may not be such in principle) but rather in that the image of the object replaces the object itself, while the methodological restrictions (being a kind of a bridge between the 'map' and the 'territory' – statistical and relatively objective reality) are often cast aside by managers and researchers as unnecessary and cumbersome for work. The popularization of such conceptual, methodological, technical, and other types of restrictions is important in as much as it can correct calculations and draw well-warranted boundaries between what we know and what we don't know.

References

Akerlof, G.A. & Shiller, R.J. (2010). *Animal Spirits. How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*. Moscow: Alpina Business Books. (In Russ.).

Bessonov, V.A. (2015). What will the modern Russian Statistics have in Store for history? *Voprosy Ekonomiki.* No. 1. Pp. 125–146. (In Russ.).

Holz, C.A. (2003). «Fast, clear and accurate»: How reliable are Chinese output and economic growth statistics? *The China Quarterly.* No. 173. Pp. 122–163.

Kalgin, A. (2016). Implementation of performance management in regional government in Russia: Evidence of data manipulation. *Public Management Review.* Vol. 18, No 1. Pp. 110–138.

Kalgin, A. & Eliseenko, V. (2015). Why manipulate? Performance measurement and data manipulation in regional and local government. Working paper WP8/2015/05, Moscow: Higher School of Economics Publ. House.

Keynes, J.M. (1998). The general Theory of Employment. In: *Istoki*. Issue 3. Moscow. SU HSE. (In Russ.).

Kordonsky, S.G. (2006). *Power Markets. Administrative Markets of the USSR and Russia*. Moscow: OGI. (In Russ.).

Krasheninnikova, Yu.A. (2011). Medical Statistics as the Way to Legitimize the Distribution of Resources in the Russian Health Care System. *Public Administration Issues.* No. 4. Pp. 28–42. (In Russ.).

Mises, L. (2005). *Human Action: A Treatise on Economics*. Chelyabinsk: Sotsium. (In Russ.).

Molyarenko, O.A. (2016). *Formal and informal relations in information support of local self-government (sociological analysis)*. PhD Thesis, Khabarovsk. (In Russ.).

[23] There may be intentional falsifications here but due to specific treatment of internal processes in the state sector, most of them are quite difficult to carry out.

Niskanen, W.A. (2004). The peculiar economics of bureaucracy. *Vekhi ekonomicheskoj mysli. Vol. 4; Ekonomika blagosostoyaniya i obshchestvennyj vybor.* Pp. 477–493. (In Russ.).

Pyankova, A.I. (2012). Should the participation in the census be mandatory? *Voprosy Statistiki.* No. 10. Pp. 22–34. (In Russ.).

Rogozin, D.M., Kartavtsev, V.V., Galiyeva, N.I., Vyugovskaya, E.V. (2016). *Methodological Audit of a Mass Survey*. Moscow: Delo Publ. (In Russ.).

Scott, J.C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.

Shklyaruk, M., Skugarevsky, D., Dmitriyeva, A., Skifsky, I. & Begtin, I. (2015). *Criminal Statistics: Mechanisms of Formation, Causes of Distortion, and Ways to Reform*. Research Report. Saint-Petersburg, Moscow: Norma, Tsentr nezavisimykh sotsialnykh issledovany i obrazovaniya. (In Russ.).

Stigliz, J., Sen, A. & Fitoussi J.-P. (2010). *Mis-measuring Our Lives: Why GDP Doesn't Add Up*. New York: The New Press.

Terzi, A. (2010). Keynes's uncertainty is not about white or black swans. *Journal of Post Keynesian Economics*. Vol. 32, No. 4. Pp. 559–565.

For citation: Molyarenko, O.A. (2019). Generating State Statistics: a View from «Below». *ECO.* No. 10. Pp. 8-34. (In Russ.). DOI: 10.30680/ECO0131-7652-2019-10-8-34.