What Opportunities Do New Technologies Bring About for Price Statistics?[1]

Vladimir Bessonov, *National Research University Higher School of Economics*

bessonov@hse.ru

The paper discusses new opportunities for Russian price statistics that present-day information and communication technologies bring about. The paper is a response to the study Isakov et al. (2021) dedicated to the effort of developing a toolset to build a price quotation database through automated internet data collection and construction of consumer price indexes based on it. Discussed are the potential implications of this activity for price statistics.

1. Introduction

Advances in information and telecommunication technologies have in recent years led to a sharp rise in online sales (or sales involving the internet), propelling this hitherto exotic retail channel to establish itself as a leader within a short space of time. The coronavirus that emerged in 2020 gave further impetus to this process. Moving forward, this sales channel is certain to gain further ground.

These technologies can also help set up a system of automated price quotation collection and other data collection to construct consumer price indexes (CPIs). The volumes of data immediately available are rising sharply as the costs of their collection are dramatically declining. This suggests that new technologies are set to destroy the age-old monopoly of national statistical services on building price statistics indicators of acceptable quality. Historically, these operations were based on costly nationwide collection of price quotations and involved huge numbers of experts.

Yet the implications of new technologies for price statistics are not only about breaking the statistical authority's monopoly on primary data acquisition. It is no

---

less important that new technologies open up new development opportunities for price statistics and inflation process analysis. All this calls for a discussion and thorough understanding.

The Isakov et al. paper (2021) is an example of a new-generation research paper in statistics, based on cutting-edge web-based big data technologies. The paper is focused on the development and description of a toolset to construct a database and price indexes on its platform. In the following sections, we discuss what substantive implications this activity may have for Russian consumer price statistics and what opportunities it opens up.

2. Building trust in Russian price statistics

The indicator estimates which can be obtained in the paper under discussion are not based on officially registered price quotations. This is essential because, among other things, official consumer price indexes tend to be viewed with criticism; they are said to underrate an increase in the cost of life, which is reflected in persistently higher household inflation expectations.[2] Should we challenge as much as we please or support official indexes, a judgement on how reliable these data are can only be achieved through a quantitative analysis based on a reliable dataset of price quotations obtained independently of the statistical authority.

The paper under discussion affords this opportunity. The authors' estimates are essentially aligned with official figures, with several minor discrepancies explained by the specifics of online sales in the time of coronavirus, differences in territorial coverage, and similar factors. Moving forward, as this sales channel evolves, the degree of alignment between alternative estimates and official data may become higher. The emergence of alternative estimates alongside deep insights into the problems of measuring inflation is expected with time to enhance trust in official data, which may in turn help anchor inflation expectations in Russia.

3. Promoting research

A key outcome of the paper is the creation of an extensive, well-structured and generally accessible database of price quotations. This is likely to carry several far-reaching implications. First and foremost, the emergence of a new reliable data source with low access costs may promote further research in measuring Russian inflation. Only scarce studies are currently available in this field. This is a paradoxical situation given that these aspects have received fairly extensive[3] coverage in foreign literature even for emerging, let alone advanced, economies.

---

[2] See, e. g. inFOM (2021).
[3] In this connection, we only refer to the ILO et al. (2020).

Importantly, price statistics is – compared to other branches of statistics – of critical importance thanks to price indexes playing the role of deflators, among others. When nominal values are deflated into real ones, the latter inherit the distortions imbedded in the price indexes. This is why having reliable price indexes is indispensable for obtaining real macroeconomic indicators of acceptable quality. Poor quality of these indicators is fraught with significant distortions of our view of ongoing economic processes.

We believe that the scarcity of available research is in no small measure caused by inaccessible data such research efforts demand or the high costs of access to them. The overall access issue is a decades-old problem of Russian statistics,[4] going back to Soviet times. The precedent of creating a well-structured and publicly available database may become an impactful factor to propel Rosstat experts to find, at last, a solution to the data access problem.[5] This would become the second key outcome of the paper under discussion.

4. Feedback for the statistical authority

The authors of alternative inflation estimates are forced to look into the nuances of building relevant indicators. In doing so, they help reveal the problems of official methodology, which holds the promise of their solution. In this connection, let us expand on just one issue related to building an official weekly CPI. According to the official methodology text,[6] official weekly indexes are built in accordance with an incorrect index formula. Growth rates of the aggregate weekly CPI are defined as the weighted arithmetic mean of individual index growth rates for the same period. To obtain an aggregate CPI estimate for longer periods (e.g. for a month), weekly CPI estimates are successively multiplied. At first sight, this seems like a totally innocent procedure. However, it is well known that such index formulas (they are known as Sauerbeck chain indexes) induce an ever-growing bias in cumulative price indexes. Even if no individual price indexes show a growth trend, the aggregate index may increase without limit. This is explained by the fact that indexes built on every chain linking step fail to satisfy the time reversal test.

Rosstat experts had to deal with this problem full-scale between 1992 and 1993, when constructing industrial producer price indexes (PPIs). The aggregate PPI, built by aggregating individual indexes that pointed to price growth in late 1992 in relation to late 1991, showed a 33.8-fold rise in producer prices. Yet, the estimate obtained by successive multiplication of aggregate CPIs built by aggregating individual indexes, showing month-end price growth relative to month-end prices, showed a 63.5-fold increase in producer prices, that is,

[4] For further details, see papers by Bessonov (2015, 2017).
[5] Based on the contents of Rosstat Development Strategy, no solution of this problem is planned. See Rosstat (2019), Voprosy Statistiki (2019).
[6] See Rosstat (2015).

up by 88%.[7] The situation was aggravated by the fact that both estimates were official and published concurrently on the same pages of the official publications. This problem for post-Soviet economies was studied in Lequiller and Zeischang (1994), where the authors explained in simple language that Sauerbeck chain indexes should not be used. Over a course of many years, Rosstat experts remembered but eventually must have forgotten that.

This bias may be behind Rosstat's non-publication of a weekly CPI time series. What Rosstat publishes only is the past week's CPIs, with a significant rounding error, as well as a time series of individual indexes, which are aggregated into weekly CPIs. Without knowledge of weights, users are unable to reproduce the aggregate index. As for Rosstat, they use weekly CPIs only for the current month; at month-end, the estimate for price growth since the start of the month is replaced by the official CPI estimate for the month. It is hard to imagine the authors of alternative estimates being engaged in a senseless exercise of this kind. They would inevitably build and publish a long series of weekly CPIs, without any rounding 'tricks'. This may result in the upward bias becoming evident. It is likely that the upward bias of alternative weekly estimates relative to official monthly CPIs shown in the chart in Isakov et al. (2021) is caused by this reason, among others.

Should the authors of alternative estimates pioneer the use of a correct methodology (which is based on the use of the weighted geometrical mean or at least on individual indexes relative to the end of the previous year – similar to a monthly CPI[8]), the statistical authority's experts would be incentivised to correct their methodology accordingly. Should the authors of alternative estimates of weekly indicators launch their publication as a long time series with acceptable precision, this would propel their colleagues at Rosstat to change their data publication policy – otherwise their official data would risk becoming of secondary importance. Official statisticians, whether they like it or not, would be involved in the discussion of these issues. All this is to stimulate the statistical authority to form a feedback mechanism. This would be the third vital consequence of the paper. With no alternative data available, some evident problems could be neglected. Once the capabilities to generate new estimates have come to the fore, this monopoly could also be broken.

5. New possibilities for analysing inflation

Another important consequence the effort under discussion may entail is associated with the emergence of new opportunities for inflation process analysis. The following example suggests itself.

The effort of building alternative estimates of weekly inflation gained traction in the time of the corona crisis in 2020, which is hardly a coincidence. On the one hand, in a time of crisis – given that crises are short-lived in nature – a sharp rise

---

[7] See Bessonov (2003) for details.
[8] See Rosstat (2014).

is seen in demand for high-frequency data. On the other hand, measuring economic activity in crisis times involves additional difficulties, which results in less precise estimates.

Incidentally, the corona crisis was accompanied by change in consumer spending patterns, as lockdowns inflicted a steep decline in the consumption of many types of services and a marked decline in the consumption of non-food products. Yet, considering that consumer price indexes are built on the system of weights put in place before the crisis,[9] these changes in spending patterns are not recognised in the index construction process. They could have been subsequently taken into account but the principles of consumer price statistics in Russia, as well as in other countries, rule out revisions to released CPI data. This is to say that official consumer price statistics come out in clean copies without drafts. The practice is explained by the fact that CPI data are used for indexing pensions, social and other benefits – and revisions would therefore call into question previously made indexations.

As for alternative estimates, they are free of such constraints. The analysts calculating these estimates have every right to revise previously made estimates, and in doing so they may be able to quantify the distorting impact of the coronavirus. Such adjustments could be also considered by the analysts in respect to official indexes. Importantly, the nature of the impact that changes in consumer spending patterns is making in the course of the corona crisis is a challenging subject as there are non-economic reasons behind these changes. It is worthy of note that from a technical perspective preliminary studies by the authors of the paper are quite fit to conduct this analysis, while its costs are negligible.

Another acute problem that came to the fore in the time of the corona crisis is the temporary non-availability of certain products and services to consumers. When such products/services are included in the basket based on which CPI estimates are made, a need for supplements to definitions (imputation) of relevant price quotations arises. Whatever the algorithm of this imputation, it generates the uncertainty of price index changes in a time of the corona crisis, with the nature of those a priori unclear. The problem cannot be solved within the confines of the official methodology. As regards alternative estimates, the problem is solved by revising retrospective index estimates through the use of the consumer basket that actually took shape over the course of the corona crisis: this basket strips out positions that need of price quotation imputation. Crucially, the technology in use also enables collection of information essential to obtain a system of weights.

6. Methodology development

New opportunities arise for advances in the methodology for building consumer price indexes, as another outcome of the paper. In this way, the toolset

---

[9] See Rosstat (2014).

proposed in the paper under study enables to tackle the problem of hedonic index construction. Although the constant quality index has been a subject of research for quite some time,[10] there are extremely scarce studies into how changes in product/service quality affect Russian price indexes. One could consider, as a palliative solution, applying to Russian consumer price indexes quality change adjustments obtained for overseas markets. This technique assumes, however, that the response of the Russian consumer market in this respect is essentially similar to that of a national market for which quality change adjustments are estimated. This assumption is far from evident and needs to be verified – and constructing hedonic indexes is indispensable here.

Up until now, the construction of hedonic indexes for Russia has been contained by unavailable data. The information on price quotations Rosstat collects is not sufficient for this. Beyond the data collected, for each product or service there should be also data on its consumer properties. Consequently, constructing hedonic models requires product/service-specific consumer properties data. Yet again, to understand which exactly consumer properties data need to be collected, one must build hedonic models. This is essentially the chicken or the egg dilemma. The technology behind data collection in Isakov et al. (2021) helps us break this vicious circle. Furthermore, there are keen insights into the technical aspect of a solution to this problem.

7. Conclusion

The activity under discussion – creating consumer price databases and constructing alternative consumer price indexes based on them – will be of great value if it gains momentum as part of the statistical authority's operations. If this is the case, first, this activity would not cease to exist in a while. Second, the developers could switch their attention to solutions to other essential problems in Russian price statistics – those that Rosstat experts have yet to find time for (in the spirit of the principle 'Petrol yours, ideas ours'). It is important to realise that the interaction like this would be beneficial for each and every party, and the statistical authority, when embarking on this interaction, is set to find a cornucopia – rather than a Pandora's box.

References

Bessonov, V. (2003). *Vvedeniye v analiz rossiyskoy makroekonomicheskoy dinamiki perekhodnogo perioda* [Introduction to the Analysis of Russian Macroeconomic Dynamics of the Transition Period]. Moscow: The Institute for Economic Problems of the Transition Period. [In Russian].

---

[10] See Triplett (2006).

Bessonov, V. (2015). What Will the Modern Russian Statistics Have in Store for History? *Voprosy Ekonomiki,* 1, pp. 125–146. [In Russian]. https://doi.org/10.32609/0042-8736-2015-1-125-146

Bessonov, V. (2017). What the Russian Information and Statistical System Should Be Like? *Voprosy Statistiki,* 4, pp. 22–37. [In Russian].

ILO, IMF, OECD, Eurostat, United Nations and World Bank (2020). *Consumer Price Index Manual: Concepts and Methods.* Geneva: International Labour Office.

inFOM (2021). *Izmereniye Inflyatsionnykh Ozhidaniy i Potrebitel'skikh Nastroyeniy na Osnove Oprosov Naseleniya* [Survey-Based Assessment of Inflation Expectations and Consumer Sentiment]. inFOM Analytical Report, January 2021 (Wave 95). Moscow: inFOM. [In Russian].

Isakov, A., Latypov, R., Repin, A., Postolit, E., Evseev, A. and Sinelnikova-Muryleva, E. (2021). Hard Numbers: Open Consumer Price Database. *Russian Journal of Money and Finance,* 80(1), pp. 104–119. https://doi.org/10.31477/rjmf.202101.104

Lequiller, F. I. and Zeischang, K. D. (1994). Drift in Producer Price Indices for the Former Soviet Union Countries. *IMF Staff Papers,* 41(3), pp. 526–532.

Rosstat (2014). *Ofitsial'naya Statisticheskaya Metodologiya Organizatsii Statisticheskogo Nablyudeniya za Potrebitel'skimi Tsenami na Tovary i Uslugi i Raschet Indeksov Potrebitel'skikh Tsen* [The Official Statistical Methodology for Statistical Monitoring of Consumer Prices for Products and Services and Calculating Consumer Price Indexes]. Moscow: Rosstat. [In Russian].

Rosstat (2015). *Ofitsial'naya Statisticheskaya Metodologiya po Opredeleniyu Yezhenedel'noy Otsenki Indeksa Potrebitel'skikh Tsen* [The Official Statistical Methodology for Estimating Weekly Consumer Price Indexes]. Moscow: Rosstat. [In Russian].

Rosstat (2019). Strategy for Development of Rosstat for the Year 2024 (Draft). *Voprosy Statistiki,* 26(4), pp. 3–24 [In Russian].

Triplett, J. (2006). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products.* OECD.

Voprosy Statistiki (2019). Discussion on the Strategy for Development of Rosstat for the Year 2024: Expert Commentaries. *Voprosy statistiki,* 26(4), pp. 25–31. [In Russian].