

A Quantitative Approach to the Russian Past: A Comment on “European Statistics, Russian Numbers and Social Dynamics, 1861–1914” by Alessandro Stanziani

Andrei Markevich and Ekaterina Zhuravskaya

Alessandro Stanziani’s article re-launches the discussion about the quality of Russian imperial statistics and the relevance of quantitative analysis for historical research at an important moment for Russia’s economic history, when a lot of new data are being compiled and used by scholars. Similar productive discussions took place at other critical junctions for the fields of history, economics, political science, and other social sciences. For example, Robert Fogel’s and Stanley Engerman’s “Time on the Cross” (1974) triggered a profound discussion of potential benefits and limitations of quantitative approach to studying the history of the United States.¹ The punch line of that discussion can be illustrated by the justification of the 1993 Nobel Prize in economics dedicated to Fogel “for having renewed research in economic history by applying economic theory and quantitative methods in order to explain economic and institutional change.” In the context of Russian history, similar discussions took place in the Soviet Union in the 1970s between Ivan Koval’chenko and Boris Litvak and then later in this journal in the 1990s.²

In the fields of economics and political science, today these debates turned into a consensus. The use of quantitative research based on historical data gave rise to many important general-interest contributions. A growing share of papers devoted to economic and political history in the leading economics and political science journals over the recent decades highlights that top economics and political science journals value the use of good-quality historical data over anecdotal evidence, provided that: 1) scholars make adequate corrections for systematic biases and measurement errors with available econometric techniques; 2) perform extensive robustness checks to using alternative data sources and ensuring that the patterns uncovered in the data are not driven by any particular subgroup of observations; and 3) provide

1. See, for instance: Robert William Fogel, “The Limits of Quantitative Methods in History,” *The American Historical Review* 80, No. 2 (April 1975): 329–50.

2. I.D. Koval’chenko and N.B. Selunskaja, “Massovyie istochniki i kolichestvennye metody v izuchennii agrarnoi istorii Rossii,” in I.D. Koval’chenko and V.A. Tishkov, *Kolichestvennye metody v sovetskoi i amerikanskoi istoriografii* (Moscow, 1983); B.G. Litvak, *Ocherki istochnikovedeniia massovoi dokumentatsii XIX–nachala XX vekov* (Moscow, 1979); Steven L. Hoch, “On Good Numbers and Bad: Malthus, Population Trends and Peasant Standard of Living in Late Imperial Russia,” *Slavic Review* 53, no. 1 (Spring 1994): 41–75; John Komlos, “On the Biological Standard of Living in Russia and the Soviet Union,” *Slavic Review* 58, no.1 (Spring 1999): 71–79; B. Mironov, “New Approaches to Old Problems: The Well-Being of the Population of Russia from 1821 to 1910 as Measured by Physical Stature,” *Slavic Review* 58, no.1 (Spring 1999): 1–26; Stephen G. Wheatcroft, “The Great Leap Upwards: Anthropometric Data and Indicators of Crises and Secular Change in Soviet Welfare Levels, 1880–1960,” *Slavic Review* 58, no. 1 (Spring 1999): 27–60.

convincing evidence for the mechanism behind the uncovered statistical patterns.³ Today cliometrics, or economic history research that extensively relies on statistical analysis, is an integral part of the field of modern economic history.

Russian economic history as a sub-field of the scholarly discipline of economic history has been following the same development trajectory over recent years. A large body of Russian historical statistics available from various 19th century sources facilitated the use of a quantitative approach to Russian economic history. As Stanziani correctly notes in his essay, the idea of “the identification of social ‘facts’ with quantities” got substantial prominence in 19th century Russia, which contributed to the richness and diversity of historical sources dating back to this epoch. These historical sources are available to us today and are used by quantitative economic historians. Stanziani questions the quality of these data, arguing that these statistical sources shed light on “the emerging role of economic knowledge in the political arena,” and provide a false picture of economic and social reality that they were originally designed to measure. He argues that statistics in the Russian empire appeared in a particular historical context, which affected both the quality and content of the historical data sources. As a consequence, he argues, one should “know who produced the source, when, how and why” before using any statistical sources and figures. In Stanziani’s view, the tendency to use these statistics by the quantitative economic historians raises a number of important concerns, which ultimately questions the validity of the whole approach. In Stanziani’s words, quantitative research using Russian imperial statistics considers “figures as data and not as sources to be put under historical scrutiny,” which ultimately results in “predetermined answers.”

It is worth noting that Stanziani’s view is not shared by a substantial part of scholars, who affirm that Russia’s imperial statistics may be used to address questions in Russian economic history rather than just being useful for the analysis of the history of ideas.⁴ Our aim here, however, is to address these issues from the perspective of quantitative economic history. In truth, serious quantitative economic history research pays considerable attention to the quality of data and sources employed in the analysis and tackles potential biases emerging from poor quality of the data head on. In this essay, we attempt to clarify this point and illustrate how the quantitative approach to history deals with the data issues using a number of examples about the recent developments in the field as well as our own recent contributions to Russian economic history.

3. Ran Abramitzky, “Economics and the Modern Economic Historian,” *National Bureau of Economic Research*, NBER Working paper 21636 (2015), at www.nber.org/papers/w21636 (last accessed January 20, 2017).

4. Koval’chenko and Selunskaja, “Massovyie istochniki i kolichestvennyie metody”; Paul Gregory, *Before Command: An Economic History of Russia from Emancipation to the First Five-Year Plan* (Princeton, NJ, 1994); A. M. Anfimov and A.P. Korelin (Eds.) *Rossiiia 1913 god. Statistiko-dokumentalnyi spravocnik* (St. Petersburg, 1995), among many others.

Data Quality Affects the Nature of Research Questions

An important feature of Russian imperial statistics is the presence of carefully crafted, detailed documentation of the procedures used for data collection and for the construction of quantitative indicators. This documentation allows addressing the question of data quality for the purpose of each particular study. Modern quantitative economic historians who aspire to make an important contribution to the field do not have “blind confidence” in each and every figure extracted from a historical source. Rather, all serious scholars are well aware of both potential measurement errors and possible systematic biases in the original data stemming, for example, from the incentives on the ground to misreport, or from various types of selections into the sample. In contrast to Stanziani’s claim, these limitations of the data do not make historical statistics inappropriate for use in any analysis, but instead affect the set of potentially addressable questions. For example, consider the debate on Russian grain production figures. Since Dmitrii Ivantsov, many scholars have argued that official grain-yield figures were severely underestimated because of the tendency to underreport production by local agents.⁵ The literature offers estimates of the magnitude of this underestimation in a range between ten and nineteen percent. These magnitudes are substantial and have first order implications for answering the questions about the absolute level of agricultural production in each particular past year. However, for the periods when the data-gathering procedure did not change, and to the extent that the incentives of local agents to underreport were similar across different parts of the empire (which was likely to be the case), the official figures, despite the severe downward bias in them, can and should be used to answer questions about both the dynamics and the geographical distribution of grain production in the late Russian empire.

The Multitude of Alternatives and Unrelated Data Sources Give Rise to the Possibility of Data Crosschecks

Stanziani uses an example of *zemstvo* surveys to illustrate the unreliable nature of 19th century statistics. Indeed, there are important reasons to believe that these data have not been collected as a result of random sample selection. The mass of different data sources, collected through different means using different methodologies, independent from each other (even if by politically engaged actors), creates the possibility for a crosscheck. In contrast to Yulii Ianson and Alexander Chuprov, to whose writings from the 1870s and 1880s Stanziani refers, we could use the results of the 1897 Imperial census, which allows placing the *zemstvo* surveys in a broader context. In addition, multiple statistical volumes published by the Central Statistical Committee of the Ministry of Internal Affairs since the 1860s provides information that allows checking the validity of *zemstvo* statistics, as they have broad geographical coverage (normally covering at least the so-called fifty European provinces

5. Dmitrii N. Ivantsov, *K kritike russkoi urozhainoi statistiki (opyt analiza nekotorykh ofitsialnykh i zemskikh tekushchikh dannykh)* (Petrograd, 1915).

of the Russian empire). Checks of different sources is a norm of quantitative economic history. Obviously, a use of a number of alternative statistical sources for crosschecking their consistency one against another is not a new technique. This “old trick” had been widely used by many generations of scholars who worked with historical statistics. We stand on the shoulders of these scholars. For example, the quality of figures from the annual reports of governors in the Russian empire were a subject of scrutiny. For example, in 1974, Alexander Nifontov showed that the dynamics of grain yields across the provinces of the Russian empire is highly correlated with the dynamics of the yields from the annual reports of the Ministry of State Property.⁶ Such comparisons are an important source of validating the information collected by different authorities.

New Compiled Historical Data Sets Open to Use by All Scholars Allow better Scrutiny of the Data and Research Based On It

One recent well-pronounced tendency in the field of Russian quantitative economic and social history is the emergence of new complex datasets, which combine a number of original historical sources and are made available to the scholarly community for further use in quantitative analysis. New computing as well as data-storage and data-sharing technologies decrease the costs of constructing, maintaining, and sharing historical datasets, leading to a substantial increase in the number of online data sources. Public databases of Russian historical statistics created not for the purposes of addressing a particular research question but rather for the benefit of the whole research community working on a particular historical period are very useful for validating and cross-checking different sources. The most prominent examples are the Demoscope dataset, The Electronic Repository of Russian Historical Statistics, and the Moscow State University “dynamics” data set.⁷ These projects illustrate well that quantitative economic historians do care about the quality of statistics: they not only digitize and simplify access to historical data, but also guide users on the reported numbers. For example, the Electronic Repository of Russian Historical Statistics, compiled by one of the authors of this essay, explicitly reviews the corresponding historical literature for each subset of figures. In addition to reporting the data, this online statistical archive reports analytical notes, which explicitly discusses the origins and quality of the available sources as well as gives specific reasons for the selection of the sources for this dataset.⁸ Importantly, the new technologies and free access to new databases provide the possibility for the scholarly community to discuss the quality of each data series, which leads to better understanding of the

6. A.S. Nifontov, *Zernovoe proizvodstvo Rossii vo vtoroi polovine XIX veka* (Moscow, 1974).

7. See “Historical Appendix to *Demoskop Weekly*,” at www.demoscope.ru/weekly/pril.php (last accessed 11/8/2016); Gijs Kessler and Andrei Markevich, “Electronic Repository for Russian Historical Statistics,” at <https://ristat.org> (last accessed 11/8/2016); L. I. Borodkin, “Dinamika Ekonomicheskogo i Sotsial'nogo Razvitiia Rossii v XIX – Nachale XX vv.” at www.hist.msu.ru/Dynamics/index.html (last accessed 11/8/2016).

8. Kessler and Markevich, “Electronic Repository for Russian Historical Statistics.”

potential biases in the data. The online datasets also help non-specialists to navigate among a variety of historical figures.

Robustness Checks and Exploring the Mechanism are Norms in the Field

A necessary condition for a serious contribution to quantitative research using historical statistics is exploring the robustness of the obtained findings. One of the reasons for this requirement is the recognition of the potential presence of mistakes in the original data. Thus, it is important to demonstrate that results do not depend on specific figures or subsets of figures extracted from particular sources, especially those about which there are some data-quality concerns. In our paper on the effects of the abolition of serfdom in particular, which Stanziani uses as one of the recent examples of research applying quantitative methods to Russian economic history, we employ a battery of sensitivity checks to explore the robustness of our results about the positive impact of the abolition of serfdom on the economic development of an average Russian province. For example, we test whether our results depend on the inclusion of grain yield figures from different sources that used potentially different methods for data collection. In particular, we verify that our results are robust by excluding from the dataset all figures from the volumes published by the Central Statistical Committee, mentioned above.⁹

Furthermore, uncovering a statistical relationship between different historical series is considered not enough to claim an important contribution to the field, as researchers need to demonstrate the mechanism behind the uncovered relationship. They need to explain the reasons for this pattern and present empirical evidence in support of these claims.

Econometric Theory Helps to Understand the Effects of Measurement Errors in the Data and Correct for Systematic Biases

The issues of measurement errors in the data and sample selection are at the center of the validity of any statistical analysis. Other fields of applied research, particularly in economics, developed a set of techniques that can be used to cope with these issues for any application including the application to statistical analysis of historical data.

An important classical result of econometric theory is that a measurement error, which is not systematically related to the phenomenon under study, makes it less likely for a researcher to uncover a statistically significant pattern even when this pattern exists in reality. Thus, in many historical contexts, the existence of a statistical pattern in the data indicates a very strong underlying relationship if one believes that the corresponding historical statistics are full of measurement errors. These relationships need to be understood and interpreted.

9. Andrei Markevich and Ekaterina Zhuravskaya, "The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire," *SSRN working paper* (2017), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2514964 (last accessed January 28, 2017).

Systematic biases in the reported data, which are related to the question under study, may indeed produce erroneous results of a statistical analysis if not handled properly. For example, the issue of sample selection, which Stanziani illustrates in application to zemstvo surveys, is a very important concern for an empirical researcher and may lead to wrong conclusions if scholars do not recognize and address it. Econometric techniques that are now standard for any applied work in economics and political science, however, allow the correction of these selection biases in many historical applications. In particular, directly accounting for the selection criteria in the statistical analysis helps eliminate the results' potential bias due to the selection. In cases when it is not possible to fully account for the selection criteria, econometric theory guides us on how to correct for the biases by finding a source of variation in the explanatory variable that is unrelated to the studied outcome or its unobservable determinants (including mismeasurements). For example, the way in which serfdom statistics were collected and recorded in the mid-19th century may have been related to the economic development of specific geographical areas, which may cast doubt on the interpretation of a statistical relationship between serfdom and economic development across geographical units as causal. Yet, if one uses the variation in serfdom unrelated to methods of data collection (for example, relying on the fact that monastic lands nationalized by Catherine the Great in 1764 were not distributed to the gentry throughout the next century), one can uncover the true relationship between serfdom and economic development.¹⁰ The use of exogenous sources of variation to address potential biases in statistical analysis is the cornerstone of modern econometrics.

If Historical Measures Do Not Exist, Reconstructed Proxies Are Useful, Provided that the Assumptions behind them Are Clear

Modern economic historians often manipulate data to reconstruct proxies for historical phenomena, for which data in the original sources do not exist. This reconstruction uses other available variables that relate to the phenomenon in question under some assumptions. Skeptics raise concerns about these reconstructions; however, the crucial question here is the validity of the specific assumptions that are behind each calculation rather than the validity of all such exercises. For example, we construct a proxy for the implementation of land reforms in each province in European Russia that followed the abolition of serfdom, that is, a variable that reflects the number of former serfs who started buyout contracts in each year during the period between 1862 and 1882.¹¹ For this, we use the 1877 cross-section on the number of peasants, who had not initiated the buyout operation by that time, and the redemption payments statistics, which reports the sums that peasants were supposed to pay

10. Johannes C. Buggle and Steven Nafziger, "Long-Run Consequences of Labor Coercion: Evidence from Russian Serfdom," (unpublished paper, 2015) available at http://www.johannesbuggle.com/docs/serfdom_1.4.pdf; Markevich and Zhuravskaya, "The Economic Effects of the Abolition of Serfdom."

11. Markevich and Zhuravskaya, "The Economic Effects of the Abolition of Serfdom."

each year in redemption by province.¹² Stanziani objects to this approach, arguing that “redemption amounts were seldom paid,” that they were a result of local negotiations, and that they were revised after the 1880s. Local negotiations led to differences in land prices on the basis of which redemption amounts were calculated. The facts that Stanziani presents are historically correct. However, two out of three of his criticisms are irrelevant for our procedure: it does not matter whether the initial amounts were paid or not because we do not use this information. Similarly, later revisions are irrelevant because we reconstruct the dynamics of buyout operations before 1882. As for the third criticism about local negotiations, we take this into account in two ways in our analysis. First, specific rules governed the negotiation process in each province, so that there was less variation in negotiation outcomes within provinces rather than among them. We use province-level data on redemptions to reconstruct the dynamics of buyout operations by province. Second, and most importantly, we use an exogenous source of variation in the progress of land reform for the analysis, which stems from the incentives of landlords to speed up the negotiation process related to the level of their pre-emancipation gentry’s indebtedness. Arguably, this variation is unrelated to any potential biases in the measurement of land reform implementation imbedded in our re-constructed series. Under this identifying assumption, we can use econometric techniques referred to in the previous section to uncover the true unbiased historical relationship from these data.

Modern Quantitative Economic History Inherently Uses the Positive Rather than Normative Approach to History

In his essay, Stanziani expresses concern about what he calls the “normative approach” of modern economic history. In particular, he is worried that “economic historians and economists tell us which institutions limited the economic growth of Russia and which reforms should have been adopted.” Indeed, modern economic history explores the link between institutions and economic development, an interesting and important question from both historical and economic perspectives. For example, in our work on serfdom, we compare development paths of Russian provinces with different levels of enserfment before and after its abolition. If serfdom affected economic development negatively, we could expect that provinces with higher ratio of the number serfs to the total rural population would develop faster after the abolition of serfdom. This is exactly what we find. Such an approach helps to isolate the impact of serfdom on economics development from other potential factors. This approach, however, does not prescribe anything. In contrast, it focuses on presenting what development could look like under the counterfactual scenario that is informative for an evaluation of the actual scenario. Such counterfactual benchmarks are useful in order to understand the role of serfdom in Russia’s economic development—the question that inspired many generations of students of Russian history.

12. I.I. Vil’son, *Vykupnye za zemli platezhi krest’ian-sobstvennikov byvshikh pomeshchichikh* (St Petersburg, 1878).

The Plural of Anecdote is Data, After All

Every argument that can be made about potential data problems in systematic data analysis within quantitative economic history research is just as relevant, or more so, to historical research that makes any generalizations without the use of quantitative methods. There are two important differences between relying and not relying on historical statistics for historical research. First, every rule has exceptions and relying only on anecdotal evidence does not give any sense of confidence in the answer (namely, answering the question of whether presented evidence is a rule or an exception). In contrast, statistical analysis explicitly gives us the confidence to explain any uncovered relationship. Second, non-quantitative history is actually more prone to the problem of selection: in contrast to using quantitative methods, the use of historical anecdotes does not allow correcting for potential errors and/or biases.