

THE RESULTS OF THE 1939 SOVIET CENSUS: TWO PROBLEMS OF ADEQUACY

MARK TOLTS

The article examines the adequacy of contemporary estimates of the total population of the Soviet Union based on the 1939 census. To do so, it analyzes the instructions for filling in the census form. Comparison of the better worded 1959 census instructions with the poorly worded instructions of the 1939 census shows that the latter created possibilities for double counting of the population. These findings confirm the validity of the lowest estimate of the total population of the USSR based on the 1939 census, given by the famous Russian demographer Andrei G. Volkov, which stood at only 167.6 million people. The impact of the inter-republic reallocation of prisoners' census forms was also estimated. For the entire population of Russia these estimates do not, for most indicators, change the picture previously known from the official census results. On the other hand, for Ukraine and especially Kazakhstan, the recalculations produced noticeable changes, in some cases resulting in significant corrections of the composition of the pre-war population.

Key words: population, census, Soviet Union, Russia, Ukraine, Kazakhstan.

Three decades have passed since the complete declassification of the materials of the 1939 Soviet census, yet attention to its results remains unabated. After all, only fragmentary results of the 1937 census that preceded it, rejected by the country's leadership, have survived. The results of this census have not been fully processed, and its scanty materials can only serve as an auxiliary tool for analysis. Thus, the 1939 census remains the only detailed source showing the size and structure of the population of the USSR and its parts on the eve of World War II after the gigantic upheavals of the 1930s. It is its data that are used in estimating losses as a result of the 1932-1933 famine and of the war in 1941-1945. And yet, they continue to be the subject of controversy. The most controversial question remains that of the total population of the country and its regions. This issue is the main focus of researchers. In contrast, the complex problem of the adequacy of structural indicators based on the census is still very poorly understood. This article is devoted to some important aspects of both of these problems. In the first section, when analyzing the general results of the 1939 census, special attention will be paid to the problematic provisions of the instructions for filling in the census form, which, as their analysis shows, created, in particular, the possibility of double counting. The main objective of the second section is to assess the impact of the inter-republican reallocation of census forms of prisoners on the indicators of the structure of the populations of Russia, Ukraine and Kazakhstan.

MARK TOLTS (mtolts@huji.ac.il), HEBREW UNIVERSITY OF JERUSALEM, ISRAEL.

THE ORIGINAL ARTICLE IN RUSSIAN WAS PUBLISHED IN DEMOGRAPHIC REVIEW IN 2019, 7(1), 100-117.
DOI: 10.17323/DEMREVIEW.V7I1.10822

1.

The 1939 census, as a result of direct falsification of its results, estimated the total population of the country at 170.6 million people¹. This was achieved in several successive steps. In the official results of the census, in addition to the actually enumerated population, data from special forms for checking the correctness of the count in the census (“control forms”) were included, probably often unjustifiably. The resulting total figure of the population of the USSR was once again increased by one percent, allegedly due to a possible undercounting of the census population. Distortions had their own peculiarities for certain territories, including with the aim of concealing the number of servicemen and prisoners, enumerated using a special procedure. The original materials of the 1939 census and the methods of its falsification became known only after the declassification of the Soviet archives (Bogoyavlensky 2013).

The controversy over the results of this census has a long history. Probably the first specialist to question the results of the 1939 census was the former head of one of the regional statistical offices, who ended up in the West after the end of World War II, under a pseudonym (“P. Galin”). This was done in a work specially devoted to Soviet censuses, which appeared as one of the first publications of the Munich Institute for the Study of the History and Culture of the USSR, founded in July 1950 by a group of émigré scholars from the Soviet Union. Judging by the text of this work, its author was directly related to the 1937 and 1939 population censuses in his region (Galín 1951). The most interesting places in his work are full of personal memories of a well-informed witness about the peculiarities of the functioning of Soviet demographic statistics in the 1930s. This is the only and undoubtedly valuable addition to the well-known memoirs of Mikhail V. Kurman (1993), one of the repressed leaders of Soviet statistics of that period. Galin in his publication was the first to point out, in particular, that manipulations of the data of the control forms introduced in the census were aimed at inflating the population size in the 1939 census.

Galín soon received a forceful objection from Basilius Martschenko (Vasily P. Marchenko) in a work prepared in the same institute, but published in the USA, which relied on the official data of the 1939 census. Here is what was written by its author, a former senior researcher of the Ukrainian Academy of Sciences who, while himself not taking part in conducting this census, as an economist-planner was a consumer of its official results, including those not published in the open press: “Any falsification of the basic absolute results of the census is, in general, an operation so complicated and risky that even the Soviet statistical apparatus, in other cases ready to falsify all kinds of statistical data for the needs of Soviet propaganda, had to abstain from it” (Martschenko 1953: 2). However, the seemingly impossible, as we now know, was done - the results of the 1939 census were distorted in accordance with a special top-secret algorithm developed by the organizers of this census. Note that Martschenko’s point of view on the results of the 1939 census remained dominant even among specialists until the opening of the Soviet archives.

However, even after all the archival materials of the 1939 census became available to researchers, attitudes towards them differed greatly. Let’s note two extreme positions. According to one of them, a longtime researcher of the problem continues to believe that “the 1939 census

¹ Initially, a slightly lower figure was published - 170.5 million.

was carried out with great care and is probably the most accurate” (Maksudov 2014: 308). In contrast, an internationally renowned scholar of the history of the Soviet period only reinforced his former opinion that the materials of this census were “totally worthless” (Conquest 2000: 145).

In any statistical work, especially one as complex as a population census, instructions on the collection of primary material are of paramount importance. In the case of a census, these are the instructions for filling in the census form. It is with them that any analysis of the problems of the adequacy of the results of the 1939 census should begin, yet this, unfortunately, is usually not done. After all, it is known that it is at the stage of collecting material that a very serious distortion of statistical information can occur; an example of this is the crop yield statistics of the Stalinist period (Wheatcroft, Davies 1994).

Therefore, it makes sense to compare the content of the instructions for the censuses of 1939 and 1959 (TsUNKhU SSSR 1938; TsSU SSSR 1958). The first post-war census was prepared in a calm atmosphere. Its instructional materials were reviewed in detail well in advance, in 1957 at the All-Union conference of statisticians, in which a very wide circle of specialists took part (Vsesoyuznoye soveshchaniye ... 1959). In contrast, the compilation of instructional materials for the 1939 census was strictly controlled by the top leadership of the USSR and was not discussed openly, which could not but affect their quality. This was the most difficult time for Soviet demography (Vishnevsky 1996).

A comparison shows that the two instructions for filling in the census form were far from identical. The instructions for the 1959 census are much more extensive and precise. In particular, they included a new category, important for the accuracy of the census results, of temporary residents, which was completely absent in 1939. Another drawback of the instructions for the 1939 census, noted by one of its active participants, who was to lead the two subsequent Soviet censuses, was that that in it "the question of cases in which it is necessary to draw up a control form was not sufficiently clear and detailed" (Podyachikh 1957: 151-152). During the 1959 census, this part of the instructions was expanded and concretized. The already noted imperfection of the instructions does not allow us to consider the number of control forms received in 1939 as adequate.

The most important changes in the instructions for the 1959 census were corrections of those provisions that might have led to double counting in the 1939 census. Above all, this was the clear indication in paragraph 5 of the instructions for the 1959 census that “the present population [*nalichnoye naseniye*] includes ... everyone who spent the night from 14 to 15 January in this building, regardless of whether they live here or not (except for those specified in paragraph 5i)". The above-mentioned paragraph 5i specifies that “all those who were not at home, but on the territory of the same city, settlement or village council (for example, visiting relatives and friends)”, should not fill in the census form for the place where they spent the night (TsSU SSSR 1958: 33-34). However, the exception stipulated in 1959 ("except for those specified in paragraph 5i") was absent in the corresponding place of the instruction for the 1939 census. On the contrary, the instruction indicated that "the present population includes all those who spent the night from January 16 to January 17 in this building and all those living in it who that night were on the territory of the same city, settlement or village council” (TsUNKhU SSSR 1938: 250-251). This created a real possibility of double counting for the relevant population group. In 1959,

paragraph 5z of the instructions was also clarified, which indicated that the present population included “those who had gone to the bazaar (fair) and were not staying where they could be enumerated (in kolkhoz guest houses, in hotels, with relatives, acquaintances” (TsSU SSSR 1958: 34). In 1939, those “not staying there where they could be enumerated” were not mentioned in the same paragraph of the instructions (TsUNKhU SSSR 1938: 251), which again could have led to double counting.

All these omissions and problematic provisions of the instructions for the 1939 census do not allow us to consider it “the most accurate”. A worse written instruction cannot give a better result. This is an axiom of statistical practice. But there were other factors that negatively influenced the adequacy of its primary materials – above all, the pursuit of higher numbers when collecting them. “The efforts of the organizers of the [1939] census more likely led to overcounting than undercounting of the population,” correctly wrote Evgeny M. Andreev, Leonid E. Darsky and Tatyana L. Kharkova (1993: 33). However, these authors did not take into account the possibility of double counting when preparing their most famous estimates of the population of the USSR based on the materials of this census. Then again, I myself do not know how to numerically express the influence of this factor.

Table 1. Corrections of the total population of the USSR based on the results of the censuses of 1937 and 1939 proposed by some authors

Authors	1937 census		1939 census	
	Estimated population, millions	Upward correction, %	Estimated population, millions	Downward correction, %
Andreev, Darsky, Kharkova	162.7	0.4	168.9*	1.0
Volkov			167.6	1.7
Maksudov	162.8	0.5	168.6	1.2
Official census result	162.0		170.6	

Note: * - Main variant; according to the lower variant - 168.3 million people.

Sources: (Andreev, Darsky, Kharkova 1993: 29; Volkov 1997: 18; Maksudov 2019: 244, 265).

Now, having an idea of the problematic nature of the guidelines for counting the population in the 1939 census, let us consider post-Soviet estimates of the total population of the USSR based on it (Table 1). Their differences remain significant, and the range of proposed corrections is much larger than in the case of the 1937 census. The downward corrections for the 1939 census range from 1.7 to almost 3 million (1.0-1.7%), while upward corrections to the 1937 census are concentrated in a very narrow interval between 0.7 and 0.8 million (0.4-0.5%). While the methods of calculating estimates for 1939 used by the just named three authors and Sergei Maksudov are well known and described in detail in their works (Andreev, Darsky, Kharkova 1993; Maksudov 2014), the estimate of the results of this census given by Andrei G. Volkov requires special consideration, especially since it remains undeservedly forgotten to this day.

Volkov (1997: 18) expressed his opinion clearly in the following words: “The census of 1939, despite the strictest control and even direct calls to inflate the population size, gave only 167.6 million. Knowing that they would be in trouble, the new heads of TsSU and Gosplan artificially exaggerated the results of the census by almost 3 million people in order to “reach” the

population size announced by [Joseph Stalin at the XVIII Party Congress]." Volkov was certainly firmly committed to this view, since he had expressed it earlier (Volkov, Gozulov, Grigoryants 1994: 312). A similar numerical estimate is given by such well-known researchers of the 1939 census as Dmitry D. Bogoyavlensky (2013) and Valentina B. Zhiromskaya (2001).

Today, the stages of getting the approval of the country's top authorities for the total population size based on the results of the 1939 census are well known (Davies et al. 2018). Volkov may not have known about all of them, but the artificial inflation of its results was clear even then. When considering the significance of his assessment of the results of the 1939 census, it is important to take into account that Volkov was undoubtedly the best informed expert when he expressed his opinion, and his knowledge went far beyond the boundaries of formal sources². Volkov's position in the system of Soviet state statistics was uniquely significant, despite the fact that he did not hold any high administrative position there, but was only the head of the Demography Department of the Research Institute of Statistics (Vishnevsky 2014).

The assessment given by Volkov means that he not only did not agree with the one-percent correction for underestimation, but he also did not accept the data from the processing of control forms, which were partially taken into account in their assessment by Andreev, Darsky and Kharkova, who worked in his department. For this it was necessary to look at the problem differently and have solid evidence. But did Volkov know the results of the processing of control forms? Absolutely. Maksudov (2014: 332), their great enthusiast, reports that he received a copy of the results of their processing from Darsky "25 years ago". Consequently, Volkov, under whose leadership Darsky and his co-authors then worked, could not but know about them. There are two possible explanations for Volkov's position. Either he believed that the refusal to take into account the results of the processing of control forms counterbalanced the double counting, or he believed, based on some information known to him, that these results were completely inaccurate and should not be taken into account. It is worth recalling that it has been mathematically proven that the country lacked the large mobility of the population which would correspond to the official results of the processing of control forms for the 1939 census; moreover, to the researchers who performed the corresponding calculations, their very number seems to be doubtful (Andreev, Darsky, Kharkova 1998: 36).

It is now natural to apply Volkov's figure of 167.6 million people based on the 1939 census to assess the reliability of the results of the previous 1937 census. To do this, we will also use the results of two alternative calculations by Andreev, Darsky and Kharkova on the value of natural increase in 1937 and 1938 - 5.4 and 6.0 million, respectively (Andreev, Darsky, Kharkova 1993: 48). An approximate calculation based on them gives figures that differ from the result of this census, equal to 162.0 million people - 161.6 and 162.2 million. The resulting large figure is not much higher than the census result, while the estimates of other authors significantly exceed it, reaching 162.8 million (Table 1). The lower estimate is even less than the official census figure. As my previous analysis of the 1937 census instructions showed, some of their provisions also led to double counting (Tolts 1991). It can be assumed that this factor seriously counterbalanced the

² As a confirmation, I can tell that it was from Volkov that I was fortunate enough to first find out the population of the USSR from the 1937 census, although the Goskomstat leadership persistently denied then the very fact that something had survived from this census.

undercounting in this census, of which it was always suspected. However, if the real population size according to the 1939 census is less than the estimate based on it given by Volkov, then the estimates for 1937 will be even lower.

Table 2. Correction of the official results of the censuses of 1937, 1939 and 1959 for early childhood ages proposed by Andreev, Darsky and Kharkova for the population of the USSR, %

Age, years	1937 census	1939 census	1959 census
0	+2.1	-0.8	+4.0
1	+0.5	-0.8	+2.7

Sources: (Andreev, Darsky, Kharkova 1993: 62; Kharkova 1995: 8).

There is another classic way of evaluating the accuracy of censuses - by analyzing the correction values for younger children. This is possible according to the results of calculations by Andreev, Darsky and Kharkova for the censuses of 1937, 1939 and 1959 (Table 2). The comparison shows that the magnitude of the corrections is noticeably smaller for the 1937 census than for the 1959 census, the accuracy of which has never been seriously questioned. These three authors, in the course of their study of the dynamics of the population of the USSR, corrected upward the overall results of the 1959 census by only 0.1% (Andreev, Darsky, Kharkova 1993: 63). The unusual negative correction of the 1939 census data for both the youngest ages (-0.8%) cannot be explained solely by an unjustified total 1% upward adjustment. Even after its removal, the underestimation, especially in the first year of life, remains unusually low, which can be interpreted as confirmation of the hypothesis about the role of double counting in this census. However, it is possible, looking just at these figures, to continue to assert that the 1939 census was "the most accurate."

The technical side of the mechanized processing of data from the 1939 census has been described in sufficient detail (Zhak 1958). Today, supposedly, all of its surviving materials are open to researchers, yet they too do not contain a specific algorithm for inflating the population size and concealing classified contingents, primarily the army, in the results. In general, there is data on it, but it is not known how the structural characteristics of the army contingents were included in the materials of individual regions.

2.

The main purpose of every census is to capture the composition of the population. However, the 1939 census marked the beginning of the practice of territorial reallocation of a part of the recorded population in the census results, a practice which existed until the end of the Soviet period (Tolts 2001). After the declassification of the 1939 census materials, it became known that during the processing of its results, census forms for 758.7 thousand people were sent to Ukraine and Kazakhstan (Simchenko 1990: 18-19, 24-25). This was done in order to conceal the decrease in the population of these two union republics as a result of the catastrophic events of the first half of the 1930s. A quarter of a century ago, when analyzing the results of the 1939 census, I hypothesized that the census forms sent there belonged to a part of prisoners in forced labor camps located in the northern and eastern regions of Russia (Tolts 1995). The organizers' purpose

in manipulating the census materials was not only to make it possible to inflate the population of Ukraine and Kazakhstan, but at the same time to conceal the very high concentration of prisoners in their places of detention (Simchenko 1990: 2770). The hypothesis of the inter-republican reallocation of prisoners' census forms has been accepted by specialists (Bogoyavlensky 2013; Rudnytskyi et al. 2015).

The number of prisoners from Russia added to the population of Ukraine was only 8.4 thousand more than the number added to the population of Kazakhstan (Table 3). However, the large difference in the number of people living in the two republics led to a noticeable difference in the impact of this manipulation on their population. The prisoners included in the census results totaled 1.2% of the official population of Ukraine, while in Kazakhstan it came to 6.1%. Although all of the prisoners' census forms attached to the population of these two republics were removed from the population of Russia, this had a lesser impact on Russia, due to its much larger size. The number of prisoners excluded from the Russian census results totaled only 0.7% of the entire official population.

Table 3. The number of prisoners in forced labor camps whose census forms were reallocated from Russia to Ukraine and Kazakhstan during processing of the 1939 census materials

Reallocated prisoners	Removed from the population of Russia	Added to the population of:	
		Ukraine	Kazakhstan
Total	758 743	383 563	375 180
<i>of these:</i>			
men	700 238	383 563	316 675
women	58 505	–	58 505
As % of officially recorded in the census:			
Entire population	0.7	1.2	6.1
<i>of which among:</i>			
men	1.4	2.6	9.9
women	0.1	–	2.0
Rural population	1.0*	1.9	8.4
<i>of which among:</i>			
men	2.1	4.1	13.7
women	0.2	–	2.7

Note: * - When attributing to this population group all prisoners of forced labor camps whose census forms during processing of the 1939 census materials were reallocated outside of Russia.

Source: (Simchenko 1990: 18-19, 24-25).

All prisoners' census forms sent to Ukraine and Kazakhstan were added to the rural population. Therefore, their share was even greater than the official size of this part of the population of the two republics: 1.9% in Ukraine and 8.4% in Kazakhstan. Census forms for all 58.5 thousand female prisoners removed from the population of Russia were added to the rural population of Kazakhstan. If we conditionally attribute all prisoners of forced labor camps whose census forms were reallocated outside Russia during the processing of census materials to the official number of its rural population, even then their share in it will be only 1.0%. This figure gives an idea of the maximum possible impact of the inter-republican reallocation of prisoner census forms on the size of this part of the population there. After all, if part of these census forms belonged to the urban population - something we cannot know - then they should not be fully

attributed to the rural population of Russia, although it was precisely to this particular part of the population of Ukraine and Kazakhstan that all were added.

Table 4. Characteristics of prisoners in forced labor camps according to Gulag statistics as of January 1, 1939

Indicator	%	Indicator	%
Age		Ethnic group	
under 16	0.1	Russians	62.9
16-17	1.1	Ukrainians	14.2
18-21	9.7	Belarusians	3.4
22-25	16.7	Tatars	1.9
26-30	20.2	Uzbeks	1.8
31-40	28.5	Jews	1.5
41-50	15.8	Germans	1.4
51-60	6.3	Kazakhs	1.3
61 and older	1.3	Poles	1.3
unknown	0.3	Azerbaijanis*	1.1
Sex		Education	
men	91.7	higher	1.9
women	8.3	illiterate	8.5

Note: * - In the source they are designated as "Türks".

Source: (Yakovlev 2000: 416-417).

Unfortunately, the declassified census materials do not contain information about the composition of prisoners in forced labor camps whose census forms were reallocated from Russia to Ukraine and Kazakhstan during processing, since they were processed not separately, but in the general data set of the census. At the same time, the statistics of the Gulag have been published, which give the main characteristics of 1,289,491 prisoners of forced labor camps as of January 1, 1939 (Yakovlev 2000: 416-417). Prisoners whose census forms fell into the inter-republican redistribution during the processing of the census results accounted for 59% of the corresponding Gulag statistics. Comparison of the only data that are available from both sources - on sex composition - shows their greater similarity. Among the prisoners whose census forms underwent inter-republican reallocation, there were 92.3% of men and 7.7% of women, and among all the inhabitants of forced labor camps, according to Gulag statistics, - 91.7% of men and 8.3% of women (Tables 3 and 4).

Of course, the structure of prisoners in forced labor camps had its own pronounced features that distinguished it from the entire population, as has already been shown by the data cited on sex composition. Among the prisoners of the Gulag camps, there were almost no people of pre-working age. Their ethnic composition also had its own special features. There was a higher level of education (Table 4). These indicators of the structure of prisoners in the Gulag camps, known to us from its statistics, have been used by me to roughly assess the possible impact of the inter-republican reallocation of prisoner census forms on the official results of the census in Russia, Kazakhstan and Ukraine. When making the recalculation, the structural indicators corresponding to the data of the Gulag statistics were superimposed on a known number of census forms removed from the population of Russia. To obtain indicators for Ukraine and Kazakhstan, the results of the computation were divided in proportion to the share of prisoners whose forms went to each of these republics. The exception was the indicators of sex composition, which were known and were taken from declassified census materials (Table 3). On the basis of all these data, correction of the

official figures gave absolute numbers according to which, taking into account the change in the total size of the corresponding population, recalculated structural indicators were obtained for the entire and rural population of the three republics (Tables 5-7).

Table 5. Characteristics of the composition of the population of Russia according to the official data of the 1939 census and according to a recalculation eliminating the effect of the removal of prisoners' census forms sent to Kazakhstan and Ukraine during processing of the census materials, %

Indicator	According to official data		According to results of recalculation		Discrepancy between official and estimated indicators	
	Entire population	Rural population	Entire population	Rural population	Entire population	Rural population*
Age						
under 16	38.8	42.6	38.5	42.1	-0.3	-0.5
16-17	3.5	3.3	3.5	3.2	0.0	-0.1
18-21	5.9	5.2	5.9	5.3	0.0	+0.1
22-25	7.2	6.3	7.3	6.4	+0.1	+0.1
26-30	9.5	8.3	9.6	8.5	+0.1	+0.2
31-40	14.2	13.0	14.3	13.2	+0.1	+0.2
41-50	8.6	8.2	8.6	8.3	0.0	+0.1
51-60	6.3	6.4	6.3	6.4	0.0	0.0
61 and older	6.0	6.7	6.0	6.6	0.0	-0.1
Sex						
men	47.2	47.0	47.5	47.5	+0.3	+0.5
women	52.8	53.0	52.5	52.5	-0.3	-0.5
Ethnic group						
Russians	82.5	80.1	82.4	79.9	-0.1	-0.2
Ukrainians	3.1	3.0	3.1	3.2	0.0	+0.2
Belarusians	0.4	0.3	0.4	0.4	0.0	+0.1
Tatars	3.6	4.2	3.6	4.2	0.0	0.0
Jews	0.9	0.1	0.9	0.2	0.0	+0.1
Germans	0.8	1.0	0.8	1.0	0.0	0.0
Kazakhs	0.3	0.4	0.3	0.4	0.0	0.0
Poles	0.1	0.1	0.1	0.1	0.0	0.0
Education						
higher**	1.3	0.3	1.3	0.4	0.0	+0.1
illiterate***	21.9	28.4	21.8	28.1	-0.1	-0.3

Notes: * - Highest estimate (see text); ** - per 100 people aged 22 and over; *** - at the age of 15 years and older.

Sources: (Demoscope Weekly 2020; Zhiromskaya 1999: 105; Simchenko 1990: 18-19, 24-25; Yakovlev 2000: 416-417).

The results of the recalculations show that for the entire population of Russia the removal of census forms should not have affected the value of most of the indicators - 13 out of 21 figures remain completely unchanged (Table 5). In another five cases, the difference obtained for the entire population are within the rounding accuracy, which means that they should not be considered as significant discrepancies. Only for the largest age group under 16 is its share in the total population adjusted by 0.3 percentage points. The proportion of women in the entire population decreases by the same amount and, accordingly, the proportion of men increases. For the rural population of Russia, the impact of the removal of census forms could, of course, be higher, but even for them, the calculation gives smaller maximum possible discrepancies, with one exception, than for the entire population of Ukraine.

The discrepancies of the indicators for Ukraine are not just larger in numbers compared to Russia. The results of recalculations in some cases reverse our idea of the ratio of the indicators themselves in the two republics. Thus, the inter-republican reallocation of census forms led to the fact that in the official results of the census, the prevalence of women was more pronounced in Russia (52.8%) than in Ukraine (52.3%). The recalculation results show the opposite picture: in Ukraine there was a higher proportion of women (53.0%) compared with Russia (52.5%).

Table 6. Characteristics of the composition of the population of Ukraine according to the official data of the 1939 census and a recalculation eliminating the influence of the inclusion of prisoners' census forms from Russia during processing of the census materials, %

Indicator	According to official data		According to results of recalculation		Discrepancy between official and estimated indicators	
	Entire population	Rural population	Entire population	Rural population	Entire population	Rural population
Age						
under 16	35.2	38.0	35.6	38.8	+0.4	+0.8
16-17	3.9	3.9	4.0	4.0	+0.1	+0.1
18-21	6.6	6.1	6.6	6.0	0.0	-0.1
22-25	7.9	7.2	7.8	7.0	-0.1	-0.2
26-30	10.3	9.4	10.2	9.2	-0.1	-0.2
31-40	15.4	14.8	15.2	14.5	-0.2	-0.3
41-50	9.5	9.3	9.4	9.1	-0.1	-0.2
51-60	6.1	6.1	6.1	6.1	0.0	0.0
61 and older	5.1	5.2	5.1	5.3	0.0	+0.1
Sex						
men	47.7	47.7	47.0	46.6	-0.7	-1.1
women	52.3	52.3	53.0	53.4	+0.7	+1.1
Ethnic group						
Russians	13.5	7.6	12.9	6.5	-0.6	-1.1
Ukrainians	76.5	85.7	77.3	87.1	+0.8	+1.4
Belarusians	0.5	0.3	0.5	0.3	0.0	0.0
Tatars	0.2	0.1	0.2	0.1	0.0	0.0
Jews	4.9	1.1	5.0	1.1	+0.1	0.0
Germans	1.3	1.6	1.3	1.6	0.0	0.0
Poles	1.2	1.3	1.2	1.3	0.0	0.0
Education						
higher*	1.3	0.5	1.3	0.4	0.0	-0.1
illiterate**	17.6	21.7	17.8	22.2	+0.2	+0.5

Notes: * - Per 100 people aged 22 and over; ** - at the age of 15 years and older.

Sources: (*Demoscope Weekly* 2020; *RGAE. F. 1562. Op. 336. D. 604. L. 19, 24*; *Simchenko 1990: 24-25*; *Yakovlev 2000: 416-417*); The archival materials used in the calculations for this and the following table were kindly provided by Dmitry D. Bogoyavlensky, for which the author is deeply grateful to him.

For Ukraine, the recalculation gives the maximum difference for the share of the titular ethnic group (Table 6). According to the official census data, Ukrainians accounted for 76.5% of the total and 85.7% of the rural population, while according to the recalculation, their share increases to 77.3% of the whole and 87.1% of the rural population. At the same time, the share of Russians decreases: from 13.5 to 12.9% in the entire population and, even more noticeably, from 7.6 to 6.5% in the rural population. According to the recalculation, the proportion of Jews in the entire population of Ukraine increases to 5.0% or 0.1 percentage point, i.e., within the rounding accuracy.

Since the relative number of Russia's prisoners whose data are included in the population of Kazakhstan was much higher than for Ukraine (Table 3), the influence of this factor was significantly greater in Kazakhstan. Moreover, the results of recalculations in some cases reverse our understanding of the order of some of the most important indicators in this republic (Table 7). Thus, the official and recalculated indicators paint a diametrically opposite picture of the pre-war ethnic structure of the population of Kazakhstan. According to the official results of the 1939 census, in the entire population of Kazakhstan, Russians (40.0%) numerically prevailed over Kazakhs (37.8%). The recalculation shows the opposite was true: Kazakhs (40.2%) definitely outnumbered Russians (38.5%)³.

Table 7. Characteristics of the composition of the population of Kazakhstan according to the official data of the 1939 census and according to a recalculation eliminating the influence of the inclusion of prisoners' census forms from Russia during processing of the census materials, %

Indicator	According to official data		According to results of recalculation		Discrepancy between official and estimated indicators	
	Entire population	Rural population	Entire population	Rural population	Entire population	Rural population
Age						
under 16	36.0	37.0	38.3	40.5	+2.3	+3.5
16-17	3.2	3.1	3.4	3.3	+0.2	+0.2
18-21	7.1	6.7	7.0	6.4	-0.1	-0.3
22-25	8.9	8.2	8.4	7.4	-0.5	-0.8
26-30	10.0	9.7	9.4	8.7	-0.6	-1.0
31-40	15.6	15.6	14.8	14.4	-0.8	-1.2
41-50	9.1	9.2	8.6	8.6	-0.5	-0.6
51-60	5.7	5.9	5.6	5.8	-0.1	-0.1
61 and older	4.4	4.6	4.5	4.9	+0.1	+0.3
Sex						
men	52.1	52.1	50.0	49.1	-2.1	-3.0
women	47.9	47.9	50.0	50.9	+2.1	+3.0
Ethnic group						
Russians	40.0	33.1	38.5	30.4	-1.5	-2.7
Ukrainians	10.7	11.7	10.5	11.5	-0.2	-0.2
Belarusians	0.5	0.5	0.3	0.2	-0.2	-0.3
Tatars	1.7	1.1	1.7	1.0	0.0	-0.1
Uzbeks	2.0	1.9	2.0	1.9	0.0	0.0
Jews	0.3	0.2	0.2	0.1	-0.1	-0.1
Germans	1.5	1.8	1.5	1.8	0.0	0.0
Kazakhs	37.8	44.0	40.2	47.9	+2.4	+3.9
Poles	0.9	1.1	0.9	1.1	0.0	0.0
Azerbaijanis	0.2	0.2	0.2	0.2	0.0	0.0
Education						
higher*	0.7	0.5	0.6	0.3	-0.1	-0.2
illiterate**	27.8	30.8	29.7	34.1	+1.9	+3.3

Notes: * - Per 100 people aged 22 and over; ** - at the age of 15 years and older.

Sources: (Demoscope Weekly 2020; RGAE. F. 1562. Op. 336. D. 604. L. 91, 95; Simchenko 1990: 18-19; Yakovlev 2000: 416-417).

³ Earlier, on the basis of an alternative source of information on the ethnic composition of prisoners in forced labor camps, I obtained very close estimates: 40.4% for Kazakhs and 38.4% for Russians (Tolts 1995).

The addition of census forms of prisoners from Russia, in which men were sharply predominant, to the population of Kazakhstan, led to the fact that in the official results of the census of this republic women were in the minority in the entire population and in the rural population equally (47.9%). The recalculated results give a different picture: in the entire population, the proportions of men and women were equal (50.0%), and in the rural population there were more women (50.9%) than men (49.1%). In Kazakhstan, the level of education was also significantly overestimated. This is especially noticeable for the rural population. According to the official data of the census, the share of illiterates in it at the age of 15 years and older was 30.8%, while the recalculation increases it to 34.1%. The recalculated results show that half of the persons with higher education officially shown in the results of the census in the rural areas of Kazakhstan did not live there, but were imprisoned in Russia.

The recalculation makes it possible to see some important general consequences for the proper understanding of the age structure of the population of the republics where the removed census forms were sent. In Ukraine and Kazakhstan, the share of younger ages was undercounted and, accordingly, the share of prime working ages, which prevailed among prisoners, was overcounted. On the contrary, in Russia this manipulation of the census materials led, as already noted, to a certain inflation of the share of younger ages.

* * *

The materials of the processing of the 1939 census are not indisputable, but the extreme points of view - their total denial or the assessment of this census as "the most accurate" - cannot be considered justified. Our analysis shows that the instructions for filling in the census form in 1939 were imperfect. This could not but affect the numerical results of the census, leading to double counting of part of the population. However, there are simply no other materials describing in such detail the population of the USSR on the eve of World War II. The recalculations of the structure of the population of the three union republics, which eliminate the influence of the inter-republican reallocation of prisoners' census forms, give a concrete idea of the possible influence of this manipulation of the materials of the 1939 census. For the entire population of Russia, by most indicators these recalculations either do not change the picture previously given by official census data or, more rarely, only slightly refine it. In contrast, for Ukraine, and especially for Kazakhstan, the recalculations give noticeable changes, which in some cases significantly clarify our understanding of the composition of their pre-war population.

REFERENCES

- Andreev E.M., Darsky L.Ye., Kharkova T.L. (1993). *Naseleniye Sovetskogo Soyuza. 1922-1991* [The population of the Soviet Union. 1922-1991]. Moscow: Nauka. (In Russ.).
- Andreev E.M., Darsky L.Ye., Kharkova T.L. (1998). *Demograficheskaya istoriya Rossii: 1927-1959* [The demographic history of Russia: 1927-1959]. Moscow: Informatika. (In Russ.).
- Bogoyavlensky D.D. (2013). O pripiskakh v perepisi 1939 g. [On intentional inflation of the 1939 census results]. *Demoscope Weekly*, 571-572. (In Russ.). URL: <http://demoscope.ru/weekly/2013/0571/axiv01.php>.

- Conquest R. (2000). *Reflections on a ravaged century*. New York and London: W.W. Norton & Company.
- Davies R.W., Harrison M., Khlevniuk O., Wheatcroft S.G. (2018). The Soviet population and the censuses of 1937 and 1939. In *The Industrialisation of Soviet Russia*. Volume 7: The Soviet Economy and the Approach of War, 1937–1939 (pp. 129-155). London: Palgrave Macmillan.
- Demoscope Weekly (2020). Prilozheniye. Vsesoyuznaya perepis' naseleniya 1939 goda [Supplement. The All-Union population census of 1939]. (In Russ.). URL: <http://www.demoscope.ru/weekly/ssp/census.php?cy=2> (data downloaded on 01.01.2020).
- Galin P. (1951). *Kak proizvodilis' perepisi naseleniya v SSSR* [How censuses were made in the USSR]. Munich: Institut po izucheniyu istorii i kul'tury SSSR [Institute for the Study of the History and Culture of the USSR]. (In Russ.).
- Kharkova T.L. (1995). *The reliability of the USSR statistics for the pre-war period. Paper presented at the International Conference on Soviet Population in the 1920s and 1930s* (Toronto, 27-29 January 1995).
- Kurman M.V. (1993). Vospominaniya [Memories]. *Cahiers du monde russe et soviétique*, 34(4), 591–629.
- Maksudov S. (2014). Perepisi naseleniya 1937 i 1939 gg. [The population censuses of 1937 and 1939]. In V.B. Zhiromskaya (Comp.), *Mir glazami istorika: pamyati akademika YUriya Aleksandrovicha Polyakova* [The world through the eyes of a historian: in memory of academician Yuri Alexandrovich Polyakov] (pp. 307-332). Moscow: Institut rossiyskoy istorii RAN [Institute of Russian History, Russian Academy of Sciences]. (In Russ.).
- Maksudov S. (2019). *Pobeda nad derevney: demograficheskiye poteri kollektivizatsii* [Victory over the village: the demographic losses of the collectivization]. Moscow; Chelyabinsk: Sotsium. (In Russ.).
- Martschenko B. [Marchenko V.P.] (1953). *Soviet population trends, 1926-1939*. New York: Research Program on the USSR.
- Podyachikh P.G. (1957). *Vsesoyuznaya perepis' naseleniya 1939 goda. Metodologiya i organizatsiya provedeniya perepisi i razrabotki itogov* [The All-Union Census of 1939. Methodology and organization of the census and processing of its results]. 2nd ed. Moscow: Gosstatizdat. (In Russ.).
- Rudnytskyi O., Levchuk N., Wolowyna O., Shevchuk P., Kovbasiuk A. (2015). Demography of a man-made human catastrophe: the case of massive famine in Ukraine 1932–1933. *Canadian Studies in Population*, 42(1-2), 53-80.
- Simchenko Yu.B. (1990). *Perepis' 1939 goda: Dokumental'nyye istochniki Tsentral'nogo gosudarstvennogo arkhiva narodnogo khozyaystva (TsGANKH) SSSR* [Documentary sources of the Central State Archive of the National Economy (TsGANH) of the USSR]. In 15 parts. Moscow: Institut etnologii i antropologii imeni N.N. Miklukho-Maklaya AN SSSR. (In Russ.).
- Tolts M. (1991). Perepis', prigovorennaya k zabveniyu [The census condemned to obscurity]. In A.G. Vishnevsky (Ed.), *Sem'ya i semejnyaya politika* [Family and family policy] (pp. 161-178). Moscow: Institut sotsial'no-ekonomicheskikh problem narodonaseleniya AN SSSR. (In Russ.). URL: <http://www.demoscope.ru/weekly/2012/0523/arxiv02.php>
- Tolts M. (1995). *The Soviet censuses of 1937 and 1939: some problems of data evaluation*. Paper presented at the International Conference on Soviet Population in the 1920s and 1930s (Toronto, 27-29 January 1995) [Revised as of 30 May 2009]. URL: <https://www.academia.edu/1522451>.

- Tolts M. (2001). *The failure of demographic statistics: a Soviet response to population troubles*. Paper presented at the IUSSP XXIVth General Population Conference (Salvador-Bahia, Brazil, 18-24 August 2001) [Revised as of 12 July 2012]. URL: <https://www.academia.edu/1473864>.
- TsSU SSSR (1958). Instruksiya Tsentral'nogo statisticheskogo upravleniya pri Sovete Ministrov SSSR o poryadke provedeniya Vsesoyuznoy perepisi naseleniya 1959 goda i zapolneniya perepisnogo lista [Instruction of the Central Statistical Office under the Council of Ministers of the USSR on the procedure for the 1959 All-Union population census and filling in the census form]. In *Vsesoyuznaya perepis' naseleniya 1959 goda* [The All-Union population census of 1959] (pp. 33-50). Moscow: Gospolitizdat. (In Russ.).
- TsUNKhU SSSR (1938). Instruksiya Tsentral'nogo Upravleniya Narodnokhozyaystvennogo Ucheta pri Gosplane SSSR po zapolneniyu perepisnogo lista Vsesoyuznoy perepisi naseleniya 1939 goda [Instruction of the Central Administration for National Economic Accounting under the State Planning Commission of the USSR on filling in the census form of the 1939 All-Union population census]. In Boyarsky A.Ya., *Kurs demograficheskoy statistiki* [Lessons in statistical demography] (pp. 250-259). Moscow: Gosplanizdat. (In Russ.).
- Vishnevsky A.G. (1996). Trudnoye vozrozhdeniye demografii [The difficult rebirth of demography]. *Sotsiologicheskii zhurnal*, 1/2, 93-116. (In Russ.).
- Vishnevsky A.G. (2014). Vstupitel'naya stat'ya [Introductory article]. In Volkov A.G., *Izbrannyye demograficheskiye trudy* [Selected demographic works] (pp. 4-12). Moscow: Izdatel'skiy dom Vysshey shkoly ekonomiki. (In Russ.).
- Volkov A.G. (1997). Kak stalo krivym zerkalo obshchestva (K 60-letiyu perepisi 1937 g.) [How the mirror of society became crooked (On the 60th anniversary of the 1937 census)]. *Voprosy statistiki*, 3, 14-21. (In Russ.).
- Volkov A.G., Gozulov A.I., Grigor'yants M.G. (1994). Perepis' naseleniya: Perepisi naseleniya v Rossii i byvshem SSSR [Population census: Censuses of population in Russia and the former USSR]. In A.Ya. Kvasha et al. (Eds.), *Narodonaseleniye. Entsiklopedicheskiy slovar'* [Population. Encyclopedic dictionary] (pp. 309-313). Moscow: Bol'shaya Rossiyskaya entsiklopediya. (In Russ.).
- Vsesoyuznoye soveshchaniye statistikov 4-8 iyunya 1957 g. [All-Union conference of statisticians, 4-8 June 1957] (1959). Moscow: Gosstatizdat. (In Russ.).
- Wheatcroft S.G., Davies R.W. (1994). The crooked mirror of Soviet economic statistics. In R.W. Davies, M. Harrison, S.G. Wheatcroft (Eds.), *The Economic Transformation of the Soviet Union, 1913-1945* (pp. 24-37). Cambridge: Cambridge University Press.
- Yakovlev A.N. (Ed.) (2000). *GULAG: Glavnoye upravleniye lagerey, 1918-1960* [GULAG: Chief Directorate of Camps, 1918-1960]. Moscow: MFD. (In Russ.).
- Zhak D.K. (1958). *Mekhanizirovannaya razrabotka materialov perepisey naseleniya SSSR* [Mechanized processing of census materials in the USSR]. Gosstatizdat. (In Russ.).
- Zhiromskaya V.B. (2001). *Demograficheskaya istoriya Rossii v 1930-e gg. Vzglyad v neizvestnoye* [The demographic history of Russia in the 1930s. A look into the unknown]. Moscow: ROSSPEN. (In Russ.).
- Zhiromskaya V.B. (Comp.) (1999). *Vsesoyuznaya perepis' naseleniya 1939 goda: Osnovnyye itogi. Rossiya* [The All-Union population census of 1939: Main results. Russia]. St. Petersburg: BLITS. (In Russ.).