Center for Global Economic History

Universiteit Utrecht

CGEH Working Paper Series

A dataset on human capital in the former Soviet Union area

Sources, methods, and first results

Dmitry Didenko, Vnesheconombank

Peter Foldvari, Debrecen University

Bas van Leeuwen, Utrecht University

August 2012

Working paper no. 35

http://www.cgeh.nl/working-paper-series/

A dataset on human capital in the former Soviet Union area

Sources, methods, and first results*

Dmitry Didenko[1]

Péter Földvári[2]

Bas van Leeuwen[3]

Abstract

To date, the rise and fall of the (former) USSR has triggered a lot of research. Many have focused on the accumulation of physical capital, growth, and consumption. Recently, also the accumulation of human capital has increasingly been incorporated in this picture. However, few datasets exist that cover this crucial variable for this vast area. Therefore, our main objective is to introduce a new dataset that contains human capital related time series for the USSR (and the Newly Independent States (NIS) after its dissolution), constructed mostly on an annual basis. These data were drawn from various primary sources, available datasets and secondary literature where our focus was on constructing a dataset as clear, transparent and consistent as possible. It is our hope that, by supplying these data in electronic format, it will significantly advance quantitative economic history research on Russia and all over the former Soviet Union area (FSU) and will inspire further research in various new fields relating to intellectual production. The data presented in this paper follow after the discussion of the information value of the primary sources utilised, and the various problems that arose when linking and splicing the data from various sources. After constructing series of human capital indicators we perform a time-series and spatial analysis in order to identify the long-term trends of education penetration and of the human capital development in the FSU area with a strong emphasis on inequality issues between the NIS. Applying these results in a simple growth accounting framework provides us with some preliminary insights on the role of human capital in economic development in the FSU area.

Keywords: human capital, education, book production, economic development, socialism, USSR

JEL Codes: P23, P24, E24, N14, N15

Corresponding author: Dmitry Didenko, Didenko_D_V@veb.ru

Acknowledgements: The findings, interpretations, and conclusions are the authors' own views which may not be shared by the institutions of their affiliation. The authors acknowledge the financial support from the Netherlands Organisation for Scientific Research (NWO) under the CLIO-INFRA project.

[1] Senior Analyst, State corporation 'Bank for Development and Foreign Economic Affairs (Vnesheconombank)' (Russia).
[2] Associate Professor, University of Debrecen (Hungary), post-doc researcher, Utrecht University (the Netherlands).
[3] Postdoc Researcher, Utrecht University (the Netherlands).

1. Introduction

It is undisputed that human capital plays an important role in economic growth and human development. It is seen as indicative of long run growth, reduction in corruption, participation in decision making, etc (e.g. Lucas 1988; Romer 1990; Perotti 1996; Alesina and Perotti 1996). However, especially for the former socialist countries, very little information on this variable is available. Recently, some papers on long run development of human capital and growth have appeared dealing with China and Eastern Europe (e.g. Foldvari and Van Leeuwen 2009; 2011; Van Leeuwen and Foldvari 2011; Van Leeuwen, Van Leeuwen-Li, and Foldvari 2011), but research on how it affects economic development in these countries is still in its infancy.

This is especially true for the former Soviet Union area (FSU)[4] where the standard datasets do hardly ever include human capital. For example, the dataset 'Soviet Economic Statistical Series' constructed by the Slavic Research Center at Hokkaido University, is primarily focused on external trade while Easterly and Fisher (1995) do not include human capital as a monetary measure. Even the big international datasets from Cohen and Soto (2007) and Morrisson and Murtin (2009) do not include estimates for the USSR (although Morrisson and Murtin in their paper do make some guesstimates).

In Section 2 we develop a new and consistent dataset on human capital and related measures for the USSR and the Newly Independent States after its dissolution. We constructed the data series of various human capital indicators (both in natural- and monetary units), basically on an annual basis stretching back in most cases to 1920s, and in some instances even to the 19th century Russian Empire. To this dataset we added population (which is a crucial variable in many human capital estimates) in age-cohort breakdown, as well as comparable macroeconomic indicators like GDP, fixed (physical) capital stock, size of the general government expenditures, and the total wage bill. These data were drawn from various primary and secondary sources (including available datasets and literature) where our focus lay in constructing a dataset as clear, transparent, and consist as possible. Section 3 discusses the construction of the human capital indicators as well as their spread throughout the FSU area, while Section 4 deals with economic development and spatial growth of human capital in the FSU comparing it with China. We end with a brief conclusion.

2. Primary and secondary sources, description, and data discussion

2.1 General description of the sources

The starting point in constructing the dataset consisted of the official statistics, available datasets and the research literature based on them (Table 1). The official statistical data are easiest to reach. Indeed, as pointed out in Davis and Wheatcroft (1994) as well as in other literature starting at least from Gerschenkron (1947), the Soviet official series contain the information that at least was not intentionally falsified in a straightforward way as the government statistical offices preferred either to not to publish the unpleasant data or to adjust the methodology to let the resulting figures look better.

The basic official publication used for this study is the statistical yearbook "The national economy of the USSR". In addition, the USSR statistical office also published topical volumes like "Labour", "Construction of culture", "Culture, education and science", "Females and children",

[4] 'The former Soviet Union' (the FSU or ex-USSR) is the mostly common term used hereinafter for all time periods and for all territorial coverage of both the Russian Empire, Soviet states after its fall, the USSR and the Newly Independent States after its collapse. The terms 'USSR' or 'Soviet Union' are used for the period of 1922-1991 only when this state existed within its actual borders. The term 'Newly Independent States' refers to multiple of existing states on the territory of the former USSR, both to the period after its dissolution and to the period when they were the Soviet republics, basically within their current borders. Russia refers to the territory basically within the borders of the contemporary Russian Federation, in various periods.

Table 1: Basic human capital related indicators for the FSU area available in the dataset

Category	Indicator	Period	Basic Sources and Literature	Notes
Human capital (self-sufficient proxies)	Literacy	ca. 1250-2010	SRSO*, HSE IDEM (2011), Mironov (1985, 1991, 1994, 2003)	Except the NIS other than Russia for 1990-2010.
	Age heaping	1897-2010		Calculated based on distribution of 1-year cohorts of population at age 23-62.
	Average years of education	1897-2002	Russian Empire Statistical Office (Troinitskii, N.A., ed., 1905), SRSO, HSE IDEM (2011), Poliakov, ed. (1992, 1999, 2007)	Calculated based on inputs. Except the NIS other than Russia for 1990-2010.
	Educational enrolment	ca. 1800-2010	SRSO, Johnson (1950)	
	Government expenditure on education	1923-2010	Soviet and Russian Ministries of Finance (NarKomFin, MinFin, Kaznacheistvo Rossii), SRSO, SU–HSE (2005, 2007, 2010, 2010a), UIS UNESCO (2011), De Witt (1961), Noah (1966), Plotnikov (1954), Subbotina (1965)	
	Book production	1913-2010	SRSO	Except the NIS other than Russia for 1990-2010.
Human capital (proxies with differentials)	Wages	1985-2010	SRSO, CIS Statistical Committee (StatKom SNG)	
		1935-1984	SRSO, Chapman (1963), Zaleski (1980)	
		1923-1934	SRSO	For the entire USSR and for urban sector basically.
		1913-1922	Soviet Statistical Office, Krumin, ed. (1923, 1924)	
Population	Total persons	1885-2010	Andreev et al. (1993, 1998), Gel'fand (1992), Maddison (2010), Volkov (1930)	Except the NIS other than Russia for 1990-2010.
	Male/Female	1897-2010	HSE IDEM (2011), Poliakov, ed. (1992, 1999, 2007)	
Size of the economy	GNP/GDP	1885-2010	Becker (1969), Bergson (1961), Gregory (1982), Easterly and Fischer (2001), Harrison (1998), Maddison (2010), Markevich and Harrison (2011), Ponomarenko (2002), Steinberg (1990)	
	NMP	1928-1990	Khanin (1991), Steinberg (1990)	For the entire USSR.

Category	Indicator	Period	Basic Sources and Literature	Notes
Fixed (physical) capital	Stock	1928-2010	Easterly and Fischer (2001), Moorsteen and Powell (1966)	Gross stock, until ca. 1990 includes residential property. Except the NIS other than Russia for 1990-2010.
	Annual investment	1928-2010	World Bank (2011), Bergson (1961), Moorsteen and Powell (1966), Steinberg (1990)	
Prices	GNP/GDP deflator	1886-2010	World Bank (2011), Becker (1969), Bergson (1961), Steinberg (1990)	
	Consumer price index	1886-2010	SRSO, World Bank (2011), Chapman (1963), Gregory (1982)	

* Soviet and Russian Statistical Offices – respectively of the USSR and Russia[5].

since end 1950s normally once per decade. Besides these publications, the government financial office (Ministry of Finance since 1946) published the national budget execution reports on a 5-yearly basis since 1962 (providing annual historical data for the latest 5-year period and back to 1940 with 10- and 5-year intervals). Such publications had not been regular before. In the late 1980s they launched such reporting on an annual basis. Prior to mid-1930s the budget reporting was ordered by ministry (as it was in the Imperial period) which is not comparable with the later publications that preferred the functional (by topic) structure. The financial office also published topical volumes on educational-, cultural services-, and research expenditures twice (in 1939 and 1958).

The population data were obtained from the published census data. There were 9 comparable censuses in the FSU: 1897, 1920, 1926, 1937, 1939, 1959, 1970, 1979 and 1989. Almost all of their aggregate data were officially published some years after the respective censuses except 1937 and 1939. However the questions varied from census to census and so did the depth of coverage in age and regional breakdown. The population censuses covered the whole country territory within its actual borders except the one in 1920 that included the civilian and military population of the European part of Russia and the national regions controlled by autonomous communist governments but did not cover even the entire Russian territory (i.e. it excluded most of the Caucasus and Central Asia). Finally, we used some official volumes (e.g. "Labour in the USSR" of 1975 and 1983 editions) which were not available to the scholars at the time of their publication but have been disclosed after the Soviet Union collapsed.

2.2. Population size and literacy

Most of the censuses aggregated data are available in the electronic publication effected by the Institute of Demography at the National Research University Higher School of Economics (HSE IDEM). Besides the published data, they also include additional information from the archived records. Therefore HSE IDEM (2011) provides more detailed information for some years than the

[5] Tsentral'noe statisticheskoe upravlenie – TsSU (1918-1922 in Russia, 1923-1930 both in the USSR and Russia), Tsentral'noe upravlenie narodnokhoziaistvennogo uchota – TsUNKhU (1931-1948 in the USSR), Upravlenie narodnokhoziaistvennogo uchota RSFSR – UNKhU (1931-1948 in Russia), Tsentral'noe statisticheskoe upravlenie – TsSU (1948-1987 both in the USSR and Russia), Gosudarstvennyi komitet po statistike – GKS (1987-1991 in the USSR, 1988-2004 in Russia), Federal State Statistics Service – Rosstat (2004-present in Russia).

published census volumes do. This primarily relates to the data on age distribution of the male and female population and its literacy that are generally available with 1-year breakdown in HSE IDEM. That allows us to calculate the first indicator of literacy and numeracy: age heaping, which requires single year observations. Innumeracy (age heaping) is measured as the excess of people reporting their ages ending on multiples of -5 and -0 (i.e. 25, 30, 35 etc). This measure is then converted into the ABCC index, proposed by A'Hearn et al. (2009), which captures the percentage of persons correctly reporting their ages. Since, numeracy measurement is based on indirect responses; it captures functional skills instead of formal ones. Moreover, it was less politically sensitive topic and therefore probably less upward-biased than literacy.

Our second measure of education (i.e. literacy) is better recorded in the printed volumes. These volumes use various age-group breakdowns but no less than 5-year cohorts as a rule and 1-year for typical schooling ages. In order to check their reliability, we compared the data from HSE IDEM (2011) after their aggregation with those from the published census volumes for selected years (1897, 1926, 1937 and 1939 (the Russian Empire / USSR as a whole)). The discrepancies that were found (mainly for 1937) do not seem to be significant. The data for the 1959, 1970, 1979 and 1989 censuses were assessed as even more reliable and, therefore, we decided not to make a check for their consistency.

For both age heaping and literacy, all these census official publications contained rather detailed information on the whole country except those of 1920, 1937 and 1939. The population generally included both available and permanent residents. However detailed figures were published until 1970 for the available population only. As regards the 1979 census we have both the data on 5-year age-cohorts of the constant population from the official publication (GKS, 1989) and more detailed data (1-year age-groups) of the available population from HSE IDEM (2011). Therefore we chose to use the census data for the total FSU as regards the available population. The difference is not so large when taking the USSR as a whole but is evidently more (though not substantially) different for its constituent republics. However during 2002 census the Russian statistical office switched to counting population by permanent residents only thus making the historical age-cohort data not fully comparable with the last ones. The same principle was applied during the 2010 Russian census.

For 1920 the data on total population were not comprehensive as the civil war was going on and some territories were not controlled by the central government in Moscow. In addition, some data were also lost when being stored and processed. Therefore, the final data on 1920 census detailed by age cohorts are available only for 44 regions of the European part of Russia (43.3% of the total population within the borders of the USSR in 1925-1939 as estimated by Andreev et al., 1993). Therefore in order to arrive at the 1920 literacy level we used the assessments from either later official publications or from the research literature. We did the same for the earlier period as regards the European part of Russia, having taken the data from Mironov (1985, 1996, 2003) with both time-series interpolation and retropolation of the data that were either missing or inconsistent with the other estimates.

The 1937 census was found inappropriately conducted and was cancelled by the government order soon after preliminary calculation of its outcomes as they appeared to be below the government's expectations. The census 1939 main outcomes were originally published in the official media but the information was clearly insufficient. More or less detailed information on this census was published only in 1990-2000s in academic volumes based on the sources extracted from the archives.

For the total population series with adjusted census data in respective years and in the years between the censuses we used the data from Volkov (1930) and Andreev et al. (1993) that were at year start and from Maddison (2010). The latter were taken as average of the figures for two neighbour years as the original figures were estimations for mid-year.

As for the availability of literacy for the Republics (NIS), most of the data appeared in 1926. For the NIS total population series we used the data from Maddison (2010) as averages of the figures for two neighbour years (similar to the USSR). The resulting figures were close to those

from HSE IDEM (2011) which compiled their series from the national and supra-national statistical offices (CIS Statistical Committee, Eurostat).The official data on literacy, of course, had their intrinsic shortcomings. In all the FSU censuses literacy was defined as the ability to read at least one language. Hence, writing skills were not taken into account at all. However in the 1926 census instruction it was stated that an ability to write ones name was not enough to be considered as literate person. It is quite possible, however, that by 1939 such ability was often sufficient for a person to be counted as literate. In our opinion, conventional measurement of literacy based on direct questions left much room for reading proficiency criteria to be eased, especially in adverse environment of mass terror.

While the 1897 and 1926 data contain the best age distribution of literate population (1-year age groups) the later data become much less detailed. The 1939 census was the last one with official publication of literacy data for the total male and female population with breakdown between age cohorts of 9-49 years and 50+ years. For 1959 the age distribution of the overall literate population is available in HSE IDEM (2011). In none of the later census publications we have comprehensive data. It was disclosed for the age cohorts of 9-49 years and for the total population of 15 years and older in GKS RF (1992). We can arrive at the level of literacy for the total population of age 9 and over if we assume that the percentage of literate in the age of 9-14 was the same as in the age of 9-50 (that was already close to 100%). Our reconstruction of the literacy level for the age cohort of age 50+ was based on exponential-function interpolation of it for all the ages (male and female population separately). The results show that illiteracy had not been eradicated completely even by the fall of the Soviet era: in 1989 almost 8% of the USSR female population of 50 and older could neither read nor write at all.

2.3. Educational attainment

Our third educational variable (besides age heaping and literacy) concerns educational attainment. We expressed educational attainment for the male, female and total population separately in 6 ISCED levels to which the national systems of the Russian Empire (less), the Soviet Union and the NIS after its dissolution (more) generally fit.

A question on educational attainment was asked in 1897 census but was not included in the 1920 and 1926 questionnaires. In other censuses education-level grouping varied from census to census. It was the most detailed in 1959 and the least detailed in 1937 and 1939. The data on educational attainment for the NIS started to appear in 1939 but without any breakdown by age groups (except Russia). The official publications started to disclose such kind of data since 1959 only. The 1937 census had the most detailed and comprehensive age structure, all later census were more aggregated. In the case of the 1970 census we had to choose people aged 10-15 as the first age group, which did not correspond with the first category in other censuses and with the other categories of the same census ending on either 4 or 9. In the 1989 census no data were available for the age category of 10-14 years while better education-level coverage was provided.

Following the previous cross-country datasets on educational attainment (Barro and Lee, 2010; Cohen and Soto, 2007) and the age structure of the FSU published data on censuses we chose as our balanced solution to select 5-year intervals for our age groups starting with 10 years and completing with 70+ years.

The duration of each level of education was attached to each broad based educational group (complete lower secondary, incomplete and complete upper-secondary, incomplete and complete vocational non-tertiary and tertiary for this example). For incomplete levels of education as reported in the census, we assigned the average value of the nearby completed ones. However the distribution of people between the smaller categories was not equal. E.g. for 1939 all the levels of education from complete ISCED 2 to incomplete ISCED 5 are merged together with average duration of 9.92 years of schooling. This is not too far from reality at first glance but after looking at the structure of previous enrolment it becomes evident that this average duration is upward-biased: most of the people in the merged category could belong to those who had completed ISCED 2 and had no more (with only 7 years of schooling).

Hence, without knowledge of the previous life-time enrolment structure and completion rates for various age groups it was not possible to define the weights for the smaller categories. Indeed, it is likely that the duration of primary and lower secondary education changed over time though not significantly (±1 year). In most cases we assigned to each education level those durations that were normatively prescribed as of the census date. This lead to a slight overestimation in 1970 and 1979 when significant part of the population obtained their lower secondary education at the time when its duration was 7 years (instead of 8 years later) while the proportion of people who obtained only primary education under older rules (duration was reduced from 4 to 3 years) was evidently less. In earlier years actual term of schooling level tended to be shorter than normatively prescribed one. To take this into account we used the evidence from Allen (2003) and Mironov (1991, 1994).

ISCED 4 and ISCED 5 data for 1939 and especially subsequent census years are definitely upward-biased as the share of part-time (after-job evening and correspondence) study was growing. Persons with correspondence education in ISCED 5 level were included starting from 1939. As it follows from TsSU (1971, 1977), in 1960s part-timers reached almost a half of all the ISCED 5 education enrolment and up to 20% of the ISCED 4 enrolment. Though the period of correspondence study was 0.5-1 years longer it evidently failed to compensate the lack of learning time for part-time students relative to full-time ones. And since we do not include ISCED 6 graduates their little quantities relative to those of ISCED 5 do not compensate the upward bias anyway.

2.4. Educational enrolment

The Soviet-era official publications started their enrolment series since 1914/15 school year (that could already be somewhat negatively affected by the WWI) as a base to the Soviet-era comparison with one of the last years of the 'old regime'. After the 6-year pause the data on all education levels become available beginning with 1920/21, normally on an annual basis.

We included the data for the selected years of the 19th and early 20th centuries from Johnson (1950) to better highlight the place of the Imperial and the Soviet periods in education growth in the FSU. These data cannot be compared directly without taking into account decreases in population resulting from the loss of territory and wars. The actual enrolment data used was combined with the attainment data from the censuses to estimate educational attainment in the years between censuses (see section 3). Therefore we express educational enrolment in ISCED levels similar to educational attainment. As noted above, ISCED classification is generally comparable to the USSR/NIS national ones. However, we had to pay attention to some special cases.

The first case was pre-tertiary education institutions that operated in 1922-1940 as 'rabfaki' ('faculties for workers' in direct translation) that generally provided evening classes. These institutions served as an educational lift for working and socially active but low-educated people allowing them to get eligibility for entering tertiary education institutions without taking full-time secondary school course. Having taken into consideration all these features we assigned the ISCED 3 to these institutions.

Another special case were the various institutions of lower vocational education. The composition of such institutions and their level of general education significantly varied over time. Not all of them provided the trainees with general education of ISCED 3 level before 1970s like the contemporary institutions of 'PTU' and 'TU' did. The latter gradually replaced other institutions of lower level. In 1920-1940s many of lower vocational training institutions were basically the courses of on-the-job training with added school hours for elementary theoretical knowledge. We assume that the average level of general education for their graduates was ISCED 1 in 1920-1940s, ISCED 2 in 1950-1960s and ISCED 3 in 1970-1980s.

We should admit the very good quality of the series as regards post-secondary education (ISCED 4 and 5 levels). Post-graduate education (ISCED 6) was restored in 1928-29 and has good coverage despite a few breaks. The major problem in operating with the Soviet-era enrolment series was their lack of comprehension as regards primary and secondary schools (ISCED 1-3 levels). The

official statistics provided the continuous series for ordinary types of various-level education institutions but often omitted the series for special ones like general education of adults, schools with specialised classes for children and schools for handicapped children. Happily though, there are reliable series on total enrolment that can be compared with those for various levels to restore and allocate the residual quantity. We use both incomplete series and those on total enrolment to predict the complete ones on education levels for the USSR as a whole, predominantly for the pre-WWII period (1920/21-1940/41 school years), but also for some subsequent ones (1946/47-1955/56, 1962/63-1968/69, 1988/89-1990/91). It should be noted that the data availability was worsening significantly when major school reforms were launched and their implementations were in progress.

The gender composition of students is presented much worse in the official publications primarily due to later start of the coverage (1927/28 except ISCED 4 since 1921/22) and larger intervals between the data points (10-15 years maximum as regards primary and secondary schools, 3-5 years for the higher levels). However, as approximate gender parity was achieved there by 1940 we found it possible to make interpolations for primary and secondary schools until 1970/71, assuming actual data points as fluctuations around the 1940-1971 sideways trend and extrapolating the latter for the rest of the period under coverage (1972-1989). Our approximations for post-secondary non-tertiary and tertiary education are thought to have better fit to reality due to availability of more intermediate data points. Our interpolations for the period of 1955-1984 for post-graduate education (ISCED 6), that never experienced a gender parity, seem to be less in precision, due to larger intervals between data points, but also close to reality in that sense that the share of women, after its significant decline in post-WWII years, was gradually increasing since 1960.[6]

2.5. Financial data on human capital expenditures

One way to valuate human capital is to estimate expenditure on education (creating a cost-based measure of human capital). However, to do so we require estimates on government expenditure eon education. Unfortunately, the data on education expenditures for the period prior 1917 are at most indicative and not sufficient to analyse them with sophisticated methods. Most probably they are not inclusive and have the potential to be revised upwards. We provided them with the purpose of comparing their relative level to that of the subsequent period only to have a general idea of the process. As for the Soviet period the financial data are more or less reliable only after the 1923/24 fiscal year (from October to September), being the first complete one after the new relatively firm ruble was introduced and the USSR was established in December 1922.

The USSR National government consolidated budget (‘svodny biudzhet’ or ‘gosudarstvenny biudzhet SSSR’) included all levels of the state finances: the USSR central governement (‘soiuzny biudzhet’), the union republican governments (‘respublikanskii biudzhet’) as well as regional and local governments (‘mestnyi biudzhet’). The government budgets of autonomous republics that were in subordination to the union republics as well as of other similar administrative units were considered as regional government budgets.

Educational institutions of ISCED 1-3 levels were generally financed from local budgets with some co-financing from regional ones. The institutions of ISCED 4-6 levels were predominantly financed from the Union republican and the USSR central government budgets depending on their size and significance perceptions. However, most of the funds for ISCED 5-6 level institutions were supplied by the USSR central government budget.

For the USSR for the official financial reporting the most commonly used term was ‘prosveshchenie’ (‘enlightenment’ in direct translation). However, the meaning of this term underwent changes over time. Besides education proper it included other items like cultural services

[6] We tried as much as possible to take into account those changes in duration of various schooling levels that took effect over time. However, the period prior 1930s could be subject to some revisions in this aspect.

and, in certain periods, it also included expenditures on research. In certain (1989 and probably 1990) years 'enlightenment' also included some items not covered by educational, cultural services and research expenditures.

The expenditures for education proper consisted of two major groups: general education ('obshchee obrazovanie, vospitanie') and vocational education ('podgotovka kadrov'). The former generally included kindergartens (ISCED 0), schools of various types for general education for both children and adults (ISCED 1-3) as well as homes for orphan children, additional after-classes services, certain types of courses for children moral upbringing; while the latter encompassed vocational non-tertiary and tertiary education, and adult training. There was no division of the general education financing between the levels (most often they were in the same school and the same teachers could give classes to both ISCED 2 and 3 pupils). Such a classification was adopted in 1930s but in the subsequent official publications some data were recalculated backwards to the end of the 1920s.

We used primarily the data from public reporting of the USSR Financial Office.[7] In addition, we made use of the data from the USSR Statistical Office when they proved to be more compatible with the other values or if the data from the former institution were missing. The statistical office reported the expenditures with inclusion of nongovernmental institutional sources more frequently. For the early years (1920-1940s) we often preferred to take the revised national budget expenditures data from Plotnikov (1954) and Subbotina (1965) because of their later reclassification of the expenditure categories. We compared the aggregate data from Plotnikov (1954) for 1928/29-1954 with the alternative data from the Russian émigré scholar Kovankovski (1956). The differences prior to 1945 could be explained by the subsequent reclassification of the budget expenditures that Plotnikov, as an insider scholar, should have taken into account more precisely. Evidently, he was not interested in any upward revision of the earlier data that could make the subsequent growth to become more modest. Small differences for 1941-1945 arose because of the rounding in Kovankovsky (1956).

Like in the case of enrolment, we assigned some special-case education institutions to the recipients of the respective level of financing. These were various institutions of lower vocational education ('ISCED 1-3 vocational' as a special sub-category) and certain institutions of pre-tertiary academic education for adults ('rabfaki') in 1920-1930s (inside the subcategory 'Other ISCED 1-6 vocational').

The official expenditure figures included both current (for wages, scholarships and stipends, books etc.) and capital (for construction and renovation, equipment purchase and repairs). The latter accounted for about 8-10% of overall expenditures on educational, cultural services and research.The official publications provided not only the government expenditures from the budget but also from various institutional sources (that were basically under the government control). They also captured the part of private expenditures that was union republican budget revenues as tuition fees in upper secondary school grades, vocational non-tertiary and tertiary education. The fees size assessment is based on MinFin (1957) with our assumption that Republican budgets received 90% of the fees and the USSR central government received the remaining 10%. These fees were introduced in 1940 and abolished in 1956. Very approximate estimations of the other private expenditures were taken from Noah (1966) for selected years in 1950s and from Rogovin (1982) for 1976-1980.

The educational financial data were much better represented for the USSR as a whole than for its constituent republics. Therefore we used the former to estimate the latter when it was necessary. Another approach was to estimate the share of a republic in total expenditures and then converting it into absolute numbers. Logarithmic transformation was sometimes used to estimate the data in periods of high inflation (end 1920s-1930s, 1990s). We made allowance for the border changes in 1929 when Tajikistan split off from Uzbekistan and in 1936 when Kazakhstan and Kyrgyzstan split off from Russia becoming republics of the USSR.

[7] Narodnyi Komissariat Finansov SSSR – NarKomFin in 1930s, Ministerstvo Finansov SSSR – MinFin in 1950-1980s.

For the Soviet era we used our assumption for the allocation of the Union budget residual (consolidated USSR budget minus the sum of all the republican budgets, effectively the USSR central government budget) between the Republics. The size of consolidated budget of a particular republic was chosen as a single criterion to define its weight among the other republics in expenditures of the USSR central government.

The information on execution of the consolidated budget of the Russian Federation has been provided by the Treasury since 2003 (Federal'noe Kaznacheistvo – Kaznacheistvo Rossii, 2011). For earlier period (1995-2002) it is reported in various topical volumes of the State (currently 'National') Research University Higher School of Economics (SU–HSE). The latter institution also provided assessment of institutional and private education expenditures in Russia since 1995.

2.6 Book production

Besides literacy, age heaping, enrolment and government finance, another indicator of education is book production per annum. The two indicators of book production (number of titles, number of copies) capture codified knowledge production (the former more, the latter less) and consumption (the latter more, the former less) in terms of natural output. They may be considered as a reliable proxy for human capital in the long-run before the ICT revolution (i.e. for the entire Soviet period until 1990s). However they fail to capture the quantity of information and we have no data on text volume in the books published for an extended period. The evidence provided in Mironov (2003) suggests that the share of brochures was significantly higher in the FSU than in other countries. Official publications and propaganda texts are also included into the Soviet-era book statistics while in other countries they are normally omitted.

Another feature of the book production indicators is that they are sensitive to unfavourable changes in macroeconomic environment that accompany wars and economic crises. These indicators have a more rapid and more significant response to such shocks than enrolment and education expenditures.

Nevertheless, books may be considered a useful proxy. Hence, we included them in our datasets within current country borders. This means that we made allowance for the border changes in 1929 (Tajikistan split off from Uzbekistan) and in 1936 (Kazakhstan and Kyrgyzstan split off from Russia), the same way as in educational finance. Similarly, our approach in interpolation or retropolation of the data consisted in estimating the share of a republic in total expenditures and then converting it into absolute numbers.

2.7. Labour market (employment and wages)

So far, we discussed literacy and age heaping, attainment, enrolment, and book production. These data may be considered first order human capital indicators in the sense that, besides corrections for statistical problems, they do not require further calculations. However, from these underlying data, we may be able to calculate average years of education (based on population, attainment and enrolment), and cost- and income-based human capital measures (based on expenditure on education and wages respectively). Since we already discussed expenditure on education, here we turn to wages.

The Soviet labour market was strictly regulated throughout the whole period beginning with the early (1920s) until the late (1966-1991) Soviet era. The most severe restrictions were in effect from 1940 to 1956 when not only farmers but also all other employees could not change their employment without permission from the management. For collective farm employees such restrictions were lifted in 1965. However, excluding the period of mass compulsory labour during and some time before and after World War II, a typical Soviet worker (both blue- and white-collar) had a relative freedom of choice as to what education to obtain and what occupation to choose. Moreover, the available evidence suggests that many of the formal restrictions effectively were not obstacles to a high degree of social mobility.

At the same time, in the centrally-planned Soviet economy wage proportions were defined and set by the government. However, they were set to address the shortage or abundance of

particular skills and therefore affect their supply and demand. The government planners had to set the qualification tariffs, industries' and enterprises' wage bill limits in such a way as to provide greater or lesser incentives for present and prospective employees working in a particular field. In certain periods a great deal of power was delegated by the central planners to the enterprise management to define the remuneration of individual employees or groups of employees within defined limits. Therefore it is possible to argue that wage distribution was a part of the Soviet economy (almost totally regulated by the government) that experienced the outcome of market forces, i.e. the supply of and demand for labour. Evidence of their feedback reactions (based on cross-correlation analysis) is provided in Didenko (2006).

The most significant structural shortcoming of the available official statistics on wages (including salaries) is their lack of an intra-industry dimension, so it is not possible to study wage differentials on the level of employees' occupations or educational attainment. The major exclusion were the wages of the blue- and white-collar workers in industry, construction and agriculture. The Soviet ruling elite considered the industrial sector as the key one in the national economy. That is why the relation between the wages of blue- and white-collar industrial workers may be considered as the core of the overall income distribution and, hence, as a reliable proxy for the trends of human capital private returns. Hence, our assumption is that the visible and non-visible (i.e. not reflected in official data) income relation was the same for the blue- and white-collar industrial workers in any particular year.

There were four major periods in methods of grouping the wages data in Soviet and Russian official publications: 1913-1918, 1923-1938, 1940-2004, since 2005. We tried to splice the data for these periods and make those adjustments, which we found necessary to make them as much comparable as possible. However, the data relating to different periods should be compared taking into account some imprecision that may arise of this. We believe that these discrepancies are not as big as to cast doubts on the indicator trend directions.

The figures for the years 1913 - 1917 are taken from the early Soviet-era publications. These data were based on the Industrial census in 1918 that covered a sample of 3043 enterprises that operated all over the period of 1913-1918 and were located on the territory controlled by the Communist (Bolshevik) government in Moscow. In that period not only official figures could be published but some independent estimations as well. We believe that the independent calculations in Krumin, ed. (1924) and quasi-official ones in TsBST (1924) were more reliable. Therefore we corrected the 1918 census data for the respective coefficient for the year 1913, for which all the sources have the wage data for blue- and white-collar employees.

The official series for the Soviet era start from 1922. It was the year when monetary wages were replacing the predominance of in-kind remuneration of industrial employees: monetary share in wages jumped from 25-30% to 77% of their total wage during 1922 (TsBST, 1924). The reason was in that the monetary reform was developed and launched with the introduction of a new ruble in December 1922. This relatively firm currency circulated in parallel with previous one (subject to hyperinflation) until February 1924. From the political side the labour market gained support as the USSR was officially established in December 1922 after the central government in Moscow gained control over the territory within its borders (except some parts of the Central Asia).

The wage data prior to 1940 were based on sample surveys of enterprises, the large-scale ones predominantly. Enterprises of either 30 employees or 16 employees with any engine-power equipment were considered as large-scale in 1920s and 1930s. In 1928 they included about 72% of the blue-collar labour force. Before the Bolshevik Revolution large-scale enterprises differed substantially with the artisan industry ('kustarnaia i remeslennaia promyshlennost') in terms of wages. It follows from TsSU (1924) that in 1913 the average annual wage of the artisan blue-collar workers was 73.64 rubles versus 291.5 rubles in large-scale enterprises. This spread tightened during the early Soviet times. As of TsSU (1929), average wage in small-scale enterprises was 33% lower than in the large-scale ones in the period 1925/26-1928/29. And in the same period the large-scale industry wages were on average 4.1% higher than in the entire industry. However we lack

more extensive data on the early Soviet small-scale industry. Nor do the sources provide data on blue- and white-collar wages within the small-scale industry.

We did not calculate wages for the period of demonetarised and hyper-inflationary economy before a relatively firm currency was introduced, since most of remuneration consisted of in-kind payments then. However, when estimating white/blue-collar wage differential for 1918-1922 we relied on retro- and extrapolations that were derived from the time-series of the monetary wages in the periods that preceded and followed.

We also avoided using the monetary data on wages for 1941-1944 due to both its poor statistics and the lack of economic sense. Although the monetary system appeared to be more stable this time than during the Civil War, the major part of consumer goods was also sold for non-monetary means of payment.

Our average wage figures include various types of monetary and in-kind remuneration of employees. However they do not include the cost of subsidies for various social services consumed by them. As such subsidies could not be normally substituted by employees' choice (i.e. used as any means of exchange) they cannot be considered as a marketable remuneration. Though we bear in mind that an accessibility to such kind of services and subsidies affected employees' preference for a particular job place and position. It follows from TsSU (1983) the share of such subsidies from the public welfare funds ('obshchestvennye fondy potrebleniia') in overall employees' income increased from 18% in 1940 to 28% in early 1980s.

The reported wages are not refined of compulsory and quasi-compulsory deductions (direct taxes, allegedly voluntary cash contributions and the government quasi-bond subscriptions). This issue was explored in Chapman (1963) only for selected years and the discrepancy between the official reported figures and the actual employees' incomes, which they could use by their own choice, was not significant in 1928 and 1937 (3-4%), but was extremely high during the WWII (40%), and was moderating thereafter (from 16% in 1948 to 12% in 1954). As the period (1928-1954) explored in Chapman (1963) was harder than that of the 1960s-1980s, when the government tended not to be so persistent in restricting personal consumption, we believe this discrepancy diminished further over time. If we assume that these deductions were equally distributed between the white- and blue-collar industrial workers then their wage differential should not be affected. However, all the evidence says that this was not the case for the distribution between the farm and non-farm employees. The former had lower wages and were stripped off sometimes below their subsistence level so that higher potential for additional extraction remained in the non-farm sector. Therefore an upward-bias (about half the size of the official/actual income differential pointed out above) arises in our average wage in the national economy (including agricultural non-state enterprises) for 1930s-1950s.

The official statistics provide much more frequent data on employment (the number of workers) than on their wages. But employment data have their own weak points that resulted in upward bias of the respective data on average wages in the national economy. This bias would increase when going back in time and becomes especially significant in republics with large rural sectors such as Kazakhstan.

In terms of employment the structure of data from official sources is definitely biased to industrial manufacturing sector. Annual data on agricultural sector include only enterprises of the state property ('sovkhozy', MTS, RTS) that in 1920-1960s constituted just a minor part of the rural labour force. Service sector is poorly represented in early times when the substantial part of it was private one (before 1930s). However, these data are consistent with the respective average wage data while more comprehensive employment data from Harrison (1998) are not. Average wage data are biased upwards (especially from 1930s to 1960s) as wages in agriculture were significantly lower than in urban-based sectors of the economy. As the share of rural employment decreased over time (like the share of non-state enterprises in agricultural employment did) the elder the data the more they are downward-biased in employment and upward-biased in average wage.

Only from 1940 scarce official data appear on agricultural enterprises with collective property ('kolkhozy') that constituted the major part of it until 1970s. So that direct calculation of

unbiased average wage becomes possible for selected years. But for the FSU republics except Russia we have unbiased average wage data only from mid-1980s. To address this problem we used a retropolation correcting for the change in urban/rural population ratio. This corrected average wage series allow us to calculate our unbiased income-based human capital measure in Section 4.

Based on Steinberg (1990) we found out that the total wage bill coverage exceeded that of the employment at least since 1965. Therefore the average wage in the national economy (including agricultural non-state enterprises) for 1960s-1980s, calculated on the basis of the Soviet official data, has an upward-bias diminishing from 12% in 1965 to 3% in 1985. This bias seems to be less in the pre-WWII years as extremely low paid farm employment had substantially larger share in overall employment than it did in 1960s-1980s.

Some data on blue- and white-collar workers were omitted in the above sources. We interpolated them based on the total employment and average wage in the state-owned sector. In some cases (mainly for 1920s) we used time-series retropolation. The data for the last year of the USSR was predicted using the StatKom SNG (1992) data. They were compiled after the USSR dissolution (December 1991) and were for the CIS countries only, i.e. the USSR republics excluding Georgia, Latvia, Lithuania, Estonia. In 1990 the CIS countries accounted for more than 95% of the USSR population. The StatKom SNG (1992) figures were more comprehensive as they included all the sectors of the economy while the other official figures did not.

2.8. National accounts (GDP, fixed capital) and their price indices

Obviously, any analysis of human capital is severely limited if we cannot calculate its relationship with per capita income. Therefore, we also include series of per capita national income in the USSR. However, initially, the structure of the national income of the former USSR was quite different from that in most Western economies. The epistemological fundamental for national income calculations under Soviet-type socialism was the belief that no new value added may be created outside sectors of material production. Those industries which produced intangibles (i.e. knowledge producing) were classified as of intermediate consumption and non-productive. Therefore the Soviet official Net Material Product (**NMP**) figures omitted most of services until mid-1980s. For the period of 1985-1990 the USSR **GNP** data (in established prices) were calculated by the late Soviet statistical office (Goskomstat) with the IMF and the World Bank assistance and were published in IMF, WB, OECD, EBRD (1991). With application of the same methodology the data for the USSR GNP for 1965-1984 were calculated in Steinberg (1990).

For the USSR period we constructed the series of NMP (based primarily on official figures for material production sector), GNP (based on the research literature for the overall economy) in established current prices, its deflators and gross fixed capital series (also in current prices) based on the available data. The figures applied to the territory within actual USSR borders. For the WWII period the USSR territory as of 1940 temporarily occupied by the Nazi troops was also included.

We checked various estimation of the USSR GNP in current prices taken from the previous literature by the monetary indicators that were originally expressed in current prices: total wage bill and the national budget total expenditures. The same procedure was applied to our estimations. We chose to link together those series that had generally the same concepts and close values in neighbour time points. We also used the series of both NMP and GDP in current prices for their cross-check.

For the period of 1885-1913 we used the data from Gregory (1982) on Net National Product (calculated by final use) to define GDP nominal growth rate relative to 1913 for which we had both the NNP from Gregory (1982) and the GDP from Markevich and Harrison (2011). Their data were also used for the period of 1913-1928 after conversion from constant 1913 to current prices employing our preferred deflators for that period. It follows from Markevich and Harrison (2011) that their Net (or Real) National Income calculated by sector of origin (precisely by net value-added that was refined from intermediary consumption but was the source for gross investment) was assumed to be equal to Gross Domestic Product as in the framework of the UN System of National

Accounts. We assumed that it was approximately equal to the GNP bearing in mind it was closer to reality after 1917 when capital flows to and from abroad were strongly limited.

For later periods we took basically the current price data points and series from Bergson (1961) for 1928-1955, Becker (1969) for 1958-1964 and Steinberg (1990) for 1965-1990 of his own calculation (not replicating the late Soviet official methodology).

The series of the USSR GNP from Harrison (1998), Easterly and Fischer (2001) were expressed in constant prices but seem to be quite artificial in monetary terms. E.g. they spliced the data from the literature expressed in various denominations (pre-1961 and 1961) without making conversion. However their data exposed short-term trends of the GNP dynamics and we used them for interpolation of the appropriate data in current prices.

Our **gross investment series** were basically taken from the same sources as GNP except Bergson (1961) because we preferred the series from Moorsteen and Powell (1966) for 1928-1957 as more complete and with application of more strict methodology. For 1958-1961 we used averages of the estimates from Moorsteen and Powell (1966) and Becker (1969). Gross fixed investment values did not include those in livestock, inventories but did include those in residential housing and capital repairs in construction and installation services. In contrast to many countries much of the residential housing was on corporate balance sheets, especially in 1960-1990s.

Our **gross fixed capital stock** estimation (in current prices) is based on gross fixed capital to GNP (at factor cost) ratio derived from Easterly and Fischer (2001) assuming that this relationship is correct for particular year regardless of its monetary expression. Easterly and Fischer (2001) used the series based on Western estimates that were generally the same that we used for our GNP and gross investment values for the period prior to 1956 at least. Our calculation of implied retirement rate showed that either *a)* the fixed capital series from Easterly and Fischer (2001) made very modest allowance for its retirement or *b)* the gross investment values from our sources were underestimated or *c)* our GNP deflator was less than that of fixed capital stock. The latter seems quite improbable as the Soviet-era prices for consumer products generally outperformed those for investment goods. The principal difference in the FSU economic growth rates assessments arises from application of different measurements of inflation, both the indicators and their size. Therefore finding an appropriate price index to evaluate the FSU human capital in monetary units is rather a complicated but very important issue. Our preferable inflation indicator was **GNP deflator** as it is the most comprehensive price index that covers an entire economy. It includes not only consumer goods and services but also government consumption and capital assets. However, we used consumer price indices as a cross-check where it was possible.

For 1885-1913 our GNP deflator was taken as an average of the 2 indices derived from Gregory (1982) who applied both Podtiagin wholesale price index of Russian regional markets and the combined retail price index of Russian capital cities St.-Petersburg and Moscow. The former was a price basket of 66 commodities, the latter was based on 38 commodities excluding housing rents.

For 1913-1928 the General retail weighted-average price index ('Biudzhetnyi indeks TsBST, obshchetovarnyi srednevzveshennyi') was chosen as our preferred one. It was constructed after the Bolshevik Revolution by the Central Bureau for Labour Statistics (TsBST) that was the joint body of the Soviet official Labour union organisation (VTsSPS), the official Central Statistical Board (TsSU) and the Government labour office (NKT). This index had the longest and the most detailed record among the other price indices published by the official statistical office. The other retail price indices had values close to it.

For the period of 1928-1955 we constructed our Chain Deflator Index (hereinafter referred to as CDI) as neither of the available price indices for this period provided any satisfactory tool for us to capture the structural changes in the Soviet economy in an optimal way.

First of all, application of any price indices to a centrally-planned economy requires some aspects to be taken into account. One of the basic features of such an economy is the government's

control over prices that is achieved either by setting them directly or by limiting their fluctuation. Finding a single reliable price index in an economy with centralised price setting and distorted price structure is a tricky issue. The prices were changed under pressure of the market situation but reactions usually lagged behind. Therefore inflation often appeared to be hidden: goods (assets) availability was deteriorating or/and their quality was worsening.

Neither consolidated price index can capture such an effect. The prices of uncontrolled market (either legal or illegal one) can give some guidance in consumer goods but not as much. These prices experienced sharp fluctuations reflecting those in supply as the majority of the goods were distributed through the controlled-price system. The consumption structure was rather complex and the proportions of goods purchased by artificial and equilibrium prices changed over time. Another obstacle is the limited availability of data. There are some data on prices of consumer goods in legal markets but there is almost no evidence on prices of consumer goods in illegal markets and on those for capital goods (that involved some sort of bribes). Summing up, any price index of a centrally-planned economy is misleading if considered on a short-term basis. However in the long-run deflator indices can capture major changes in price levels as any economic system attempts to an equilibrium state with more or less efficiency and speed to avoid self-disintegration.

It was pointed out in the literature (Ofer, 1987) that the pre-World War II period (1928-40) is the one in which the deflator index number problem has the most considerable implications for economic growth rates in the USSR. Anyway, prices after WWII and before the USSR dissolution exhibited much less increase than before. Inflation was quite moderate until mid-1960s and only by 1990s it approached to an annual 10% threshold.

Gerschenkron (1947) was the first who identified one problem in using the Soviet official statistics that estimated 'real' growth rates taking earlier years as the base (1913 or 1926/27). The so-called 'Gerschenkron effect' is the upward bias in output indices weighted by base-year prices during a process of industrialisation. This bias is caused by underestimation of inflation which is calculated employing Paasche index. The underlying negative correlation between the quantity and price of certain goods leads to an overrepresentation of goods that were scarce and costly in the base year compared to the situation later. The longer period we take to create an index, the more aggravated the problem can become. In the inter-war USSR the 'Gerschenkron effect' was further complicated by distorted price structure. On the other hand, if we take a later year as base for a price index and calculate inflation employing Laspeyres index this will result in a reverse effect: the more we go back from the later base year, the more we tend to overestimate inflation.

We chose our way to address this problem in making a synthetic deflator where weights (e.g. of 1928 and 1937) were to change when they approached or diverged from the respective weights base. It appeared to be some kind of a substitute for a chain index that would be certainly better than a sudden change of commodities/services basket. In implementing such guidelines we used the GNP estimates in Bergson (1961) for benchmark years (1928, 1937, 1940 and 1944) and on annual basis for the period 1948-1955 in current and constant price levels (1928, 1937 and 1950) to derive the price indices with different year base.

To test our CDI for its relevancy we applied it to the 1928 average wage as of the Soviet official sources. The assumptions were the following: 1) inflated 1928 wage should be close to actual one in a particular year; 2) the difference between them would indicate to changes in people's material well-being from 1928; 3) as the latter's dynamics has more or less reliable empirical evidence the difference between actual and theoretical wages could provide a good guidance in testing various deflator estimates. We also compared our CDI testing outcomes with those for basic price indices derived from Bergson (1961) estimations of the USSR GNP in various prices. The outcomes (Table 2a) generally confirmed our assumptions and better fit of our CDI comparing to previous price indices for the period prior 1950. To additionally check the relevancy of our CDI we constructed our index of average wage to GDP (GNP) per capita (AW/GDPpc) with 1928 as the benchmark, effectively index of wage bill to GDP ratio. Its dynamics (Table 2b) also generally fit the trends reported in empirical literature (e.g., Chapman, 1963; Mironov, 2004).

Table 2. Monetary economic indicators with application of our Chained Deflator Index

a) Average wage

Year	Actual average monthly wage*	What average monthly wage should have 1928 PPP				Actual vs theoretical average monthly wage			
		at 1928 weights deflator	*at 1937 weights deflator*	*at 1950 weights deflator*	***at our chain deflator***	*at 1928 weights deflator*	*at 1937 weights deflator*	*at 1950 weights deflator*	***at our chain deflator***
1928	5,86	5,86	5,86	5,86	**5,86**				
1929	6,67	6,68	7,10	7,10	**6,73**	0%	-6%	-6%	**-1%**
1930	7,80	7,63	8,60	8,61	**7,78**	2%	-9%	-9%	**0%**
1931	9,39	8,70	10,42	10,45	**9,07**	8%	-10%	-10%	**4%**
1932	11,89	9,93	12,63	12,67	**10,63**	20%	-6%	-6%	**12%**
1933	13,05	11,33	15,31	15,36	**12,54**	15%	-15%	-15%	**4%**
1934	15,48	12,92	18,55	18,63	**14,91**	20%	-17%	-17%	**4%**
1935	18,91	14,74	22,48	22,59	**17,83**	28%	-16%	-16%	**6%**
1936	20,70	16,82	27,23	27,39	**21,46**	23%	-24%	-24%	**-4%**
1937	25,32	19,19	33,00	33,22	**26,01**	32%	-23%	-24%	**-3%**
1938	28,89		36,13	36,78	**28,49**		-20%	-21%	**1%**
1939	30,31		39,54	40,72	**31,24**		-23%	-26%	**-3%**
1940	33,10		43,29	45,08	**34,29**		-24%	-27%	**-3%**
1941			44,59	46,88	**35,43**				
1942			45,93	48,76	**36,63**				
1943			47,31	50,71	**37,89**				
1944			48,73	52,73	**39,24**				
1945	43,90		54,81	58,70	**43,85**		-20%	-25%	**0%**
1946	47,50		61,65	65,33	**48,97**		-23%	-27%	**-3%**
1947	56,93		69,34	72,72	**54,64**		-18%	-22%	**4%**
1948	60,23		78,00	80,95	**60,91**		-23%	-26%	**-1%**
1949	62,04		82,44	85,25	**64,17**		-25%	-27%	**-3%**
1950	64,20		73,24	75,42	**56,77**		-12%	-15%	**13%**
1951	65,60		70,74	73,02	**54,96**		-7%	-10%	**19%**
1952	66,90		68,48	70,93	**53,39**		-2%	-6%	**25%**
1953	67,90		65,85	67,88	**51,10**		3%	0%	**33%**
1954	70,60		65,55	66,72	**50,22**		8%	6%	**41%**
1955	71,50		64,07	65,90	**49,61**		12%	8%	**44%**

* Basically in urban sector of the national economy (excluding agricultural non-state enterprises) in rubles of 1961 denomination, current prices; based on official data.

b) Average wage to GDP per capita (wage bill to GDP ratio)

Year	Real GDP per capita change to 1928 level**	Average wage to GDP per capita (AW/GDPpc) change to 1928 level***	Average wage to GNP per capita (AW/GDPpc) change to 1928 level****	Average wage to GNP per capita (AW/GDPpc) change to 1928 level*****	White/blue-collar wage differential in industry change to 1928 level
	Maddison GDP	*our CDI and Maddison GDP*	*Bergson GNP*	*our GNP*	
1928					
1929	1%	**-2%**		-2%	-2%
1930	6%	**-5%**		-8%	11%
1931	7%	**-3%**		-11%	9%
1932	5%	**7%**		-11%	26%
1933	9%	**-5%**		-22%	39%
1934	19%	**-13%**		-30%	29%
1935	36%	**-22%**		-33%	3%
1936	45%	**-34%**		-41%	-4%
1937	57%	**-38%**	-47%	-43%	-15%
1938	57%	**-35%**		-35%	-7%
1939	63%	**-41%**		-38%	-9%
1940	56%	**-38%**		-43%	-14%
1941					-4%
1942					6%
1943					6%
1944					6%
1945					6%
1946	40%	**-31%**		-47%	-6%
1947	55%	**-33%**		-44%	-22%
1948	75%	**-44%**	-51%	-47%	-30%
1949	91%	**-50%**	-56%	-53%	-35%
1950	107%	**-45%**	-54%	-51%	-40%
1951	105%	**-42%**	-55%	-52%	-44%
1952	114%	**-42%**	-56%	-52%	-47%
1953	120%	**-40%**	-55%	-52%	-48%
1954	127%	**-38%**	-55%	-52%	-48%
1955	142%	**-40%**	-57%	-54%	-45%

** Calculation based on Maddison (2010).

*** Calculation based on official data on wages, our CDI and Maddison (2010) data on real GDP per capita.

**** Calculation based on official data on wages, Bergson (1961) data on GNP and Andreev et at (1993) data on total population.

***** Calculation based on official data on wages, our data on GNP in current prices and Andreev et at (1993) data on total population.

As much of our annual calculations are based on interpolation of price indices we did not give too much weight to calculation outcomes for 1929-1936, 1938-1939 and 1946-1947 periods. Some caution is recommended in interpreting the 1945 results as in this case we did not have direct figures for the GNP (1945).

Comparing 1937-1940 and 1928 levels we found that the rapid industrialisation did not have substantial effect on real incomes. However, these incomes were mainly of urban employees. And our outcomes rather support the idea that the USSR industrialisation was enforced at the expense of urban consumption at least: real incomes fluctuated around the same level while real GDP per capita as of Maddison (2010) had grown notably (though not so significantly as it could be

supposed from official figures). Other deflators indicate that real incomes either grew or fell by about a quarter. Neither of these outcomes find support in available evidence.

We found that in two years after the monetary reform (1947-1948) and stabilisation of the government finance people's real incomes started to grow. That is caused by both rising wages and falling retail prices. Our CDI indicates 1950 as the turning point when real wages exceeded their benchmark 1928 level. Alternative deflators indicate 1953 instead. However the post-war dynamics of all the deflators have generally the same patterns while the difference in levels with our CDI is almost completely explained by the different assessment of the pre-war period.

The dynamics of our AW/GDPpc index exposed that the share of wages in GDP contracted significantly (almost twice) in 1928-1937 and went on diminishing modestly until 1950. The period of 1928-1937 is just the period of the accelerated forced industrialisation. Afterwards the share of wages either fluctuated at the reached low level or started to recover slightly since 1950. We compared AW/GDPpc ratio with our series of white/blue-collar wage differential in industry. The results were basically consistent with Williamson (1998) suggestion that a share of wages in GDP trend generally acts as converse proxy for income inequality trend: if wages lag behind GDP p.c., inequality tends to increase; conversely, if wages grow faster than GDP p.c., inequality tends to decline. We found that in the period of mobilisation-driven industrialisation both the wage bill to GDP ratio declined and the wage differential widened. However moderation of inequality preceded the expansion of wages in GDP. That was probably due to massive losses of human and physical capital during the World War II and subsequent defence burden during the Cold War.

We also calculated our CDIs using the methods described above for subsequent years where it was possible: for 1958-1964 using the GNP data from Becker (1969), 1973-1982 using the GNP data from Steinberg (1990). For the period 1956-1958 we assumed there was no inflation in absence of any other contradictory data as a transitional state between the periods of overall deflation and modest inflation starting from 1959. We derived the GNP deflator from Becker (1969) for 1959-1964 and from Steinberg (1990) for 1965-1972 using their GNP data (on the basis of the same methodology) expressed both in current and constant prices (for nearby year).

The previous research literature paid much more attention to the USSR as a whole than to its constituent Union republics as regards the national account series. Therefore they have quite less time span for the NIS. For the NIS we borrowed the data on GDP and on gross fixed capital formation, in current prices, primarily from the World Bank (2011) country data series and additionally from the World Bank (1992). Facing the two alternative estimations in World Bank (1992) we preferred to take their weighted average, with weights changing over time: we gave most possible weight to the optimistic estimation for 1980 and the most possible weight to the pessimistic estimation for 1990. The linked series generally cover the period since 1980 but for some countries (Georgia, Latvia, Ukraine) stretch as far back as 1965 for the GDP at least.

For Russia we also used the series from Ponomarenko (2002) for 1961-1979, average of the estimates from the latter and the World Bank (1992) for 1980-1989 and the latest official data that are generally in accord with the World Bank (2011). The GDP and its deflator official data prior 2002 are not quite compatible with those for later years as the Rosstat changed their methods of financial services assessment. However, we believe that any inconsistencies that arise are not significant.

Some discrepancies remained in the national accounts data between the sum of Union republics and the whole Union. However, no reconciliation of data is possible here as the GNP of the whole Union should always be less than the sum of its constituents. The latter is subject to clearing the inter-republican flows of goods, services and assets. Besides, following our data sources for the USSR as a whole, we constructed the *GNP* series while only the *GDP* series were available for the NIS. Such a difference in concepts, similar but not completely identical, could add to discrepancies.

We should admit that inflation level in the FSU prior 1991 is definitely underestimated by the World Bank (1992, 2011) as it follows from our comparison of their GDP deflator figures with

those from Steinberg (1990) and derived from Ponomarenko (2002) for Russia. Obviously a single country with centralised government price setting could not afford such an extreme difference in price index movement in its different parts in the long-run. Therefore, for derivative calculations we preferred to apply the single price index for the USSR to all of its republics for the period before its dissolution (1991).

After the construction of our GNP (GDP) and its deflator series we compared the resulting real growth rates with those derived from Maddison (2010) for the USSR (1928-1990). The discrepancy between ours and Maddison's real GDP growth rates might arise from different deflator base. Probably Maddison deflated the USSR GDP with retail price indexes that evidently exceeded the GDP deflators by about a factor of two in 1930-40s. Maddison also ignore the data that demonstrated deflation in 1950-1955. We find consistent with other evidence that it was the case when deep deflation of retail prices (after two decades of their outperforming rise) was combined with only modest increases in prices for investment (capital) goods.

3. Methods of human capital evaluation in their application to the FSU case

Above data were used to calculate human capital indicators. Such natural indicators like book production numbers and volumes, literacy, numeracy, and average years of education are surely not human capital proper but rather its proxies. However, in our case they may well be used to verify the monetary indicators or to go back in time where input monetary data are too scarce. Yet, for more recent periods, especially when literacy and numeracy reached 100% and thus did not reflect any more changes in educational attainment of a society, it became necessary and possible to calculate a monetary indicator of human capital. This can either be done by using a cost- or income based measure.

The most basic natural indicator is book production. This is often thought to be indicative of the level of literacy (Baten and Van Zanden 2008), or the accumulation of existing knowledge (Eisenstein 1979). The most notable is the decline in book titles in the USSR in 1960-1980s (Table 3), despite growth in number of book copies. Our explanation

Table 3: Book titles per million persons in Europe and the FSU

	FSU	Total Europe (without FSU)
1920s	219.2	
1930s	239.0	
1940s	161.3	321.9
1950s	269.2	343.2
1960s	335.2	430.4
1970s	327.8	570.6
1980s	292.2	702.9
1990s	190.8	751.7

Source: Plopeanu et al.2012; own calculations

considers this observation as indirect evidence of a deterioration in human capital quality in the USSR. In 1990s economic collapse contributed much to the further decline, both in number of titles and number of copies. However, the number of book titles not merely recovered in Russia, but is at historical high at present (902.0 per million inhabitants in 2009). This suggests that diversity of knowledge flows, even leaving electronic media aside, may have gotten a boost under open market system. The evidence that book printing (number of copies) in Russia has not recovered may be

explained in the way that electronic publishing (Internet most, CD/DVDs too) is replacing printed press.
Another evidence of this is that the number of translated Western titles went up. Indeed, as can be seen in below Table 4,

Table 4: Number of book titles translated in Europe per million persons

	Translations within Europe
1980	22.10
1985	26.37
1990	27.38
1995	45.20
2000	45.24
2005	53.92

Source: Plopeanu et al. (2012)

be seen in below Table, the number of translations went up quite considerably, largely because more Western European books were translated into Slavic languages (see also Abramitzky and Sin 2010).

Of course, if book production were purely an indicator of literacy, its effect on economic growth should decline when adult literacy approaches 100%. This is also true for the percentage persons reporting their correct age (ABCC index or age heaping). The results (Table 5) show that literacy rose

Table 5: Literacy and age heaping in the FSU

	Literacy	Age heaping (ABCC index)
1897		79.4
1920	44.1	
1926	54.7	85.2
1939	87.4	96.8
1959	98.4	97.7
1970	99.7	99.7
1979	99.8	99.8
1989	99.8	100.0

after numeracy. However, in both cases, after 1950 there was almost full literacy and numeracy which hardly changed in the later part of the twentieth century.

Clearly, even though literacy and numeracy reached its zenith in the 1950s, human capital formation did not. After all, if almost everyone is literate, or can count, still people may acquire more formal schooling. This is often captured by the average years of education. We used the method as proposed by Foldvari and Van Leeuwen (2009). They basically use census data by level of education as described in the previous section. The in-between years were calculated using the

Barro and Lee (2001) perpetual inventory method. However, this results in a bias: when calculating backwards, one will overestimate average years of education and when forward estimating one will underestimate it with an equal amount. Therefore, they propose to calculate each number back and forward and take an average. This estimate seems reasonably

Table 6: Reliability ratio of average years of education in the former USSR area (based on panel least squares)

	Dependent variable: Barro & Lee (2010) average years of education		**Dependent variable: average years of education (our series)**	
	coefficient	*t-value*	*coefficient*	*t-value*
constant	2.323	1.083	5.578	6.487
average years of education (this text)	0.640	2.239		
average years of education (Barro & Lee 2010)			0.270	2.239
	time and region dummies		*time and region dummies*	
No. Obs	41		41	
R^2	0.981		0.993	

plausible. If we compare it with the only available series for all Soviet republics from Barro and Lee (2010), we get an error to signal variance ratio of 56% versus 270% in the Barro & Lee series (as it follows from our test results in Table 6).[8]

Yet, calculating average years of education still does not capture all important aspects of human capital properly. After all, measuring human capital in terms of average years of education is similar to calculating physical capital in terms of number of machines: their heterogeneity makes it impossible to aggregating them by simple addition. Therefore, it is important to valuate human capital. This can be done using the cost- and the income-based measure.

In calculating the cost-based measure, we follow Judson (2002), updated by Van Leeuwen and Foldvari (2008). As suggested by Judson (2002), the cost-based human capital indicator is similar to the measurement of physical capital stock. Her method allows to calculate the per capita (or per worker) stock of human capital at the replacement value of a single year of education. By multiplying it by average years of education, we arrive at the total accumulated stock of human capital per capita at its replacement cost, as proposed by Van Leeuwen and Földvári (2008):

$$h_t = S_t \sum_j d_{jt} a_{jt}$$

where h_t denotes the average human capital stock per worker in year t, S_t is the average years of formal education in year t, d_{jt} is the public expenditure on education per level j in year t (per student enrolled), a_{jt} denotes the share of the labour force (population at the age of 15+ in the FSU case) in year t with a certain level of education.

This method does not include foregone wages and non-government spending on education largely because these data are unavailable for many countries and adding them would make these

[8] If series x and y are both unbiased estimators of a latent variable s, that is: $y = s + \varepsilon$ and $x = s + \eta$ then the limit of the OLS estimator of β_1 in the equation: $y = \beta_0 + \beta_1 x + u$ will be $\frac{Cov(x,y)}{Var(x)} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_\eta^2}$. From this we can estiamte which of the two proxies x and y has lower error to signal ratio.

series incomparable with other countries. However, it is based on the key component of schooling costs, which normally defines their dynamics. And the above-mentioned shortcoming can be remedied in principle by adding private expenditure and foregone earnings.

Using the income-based measure, we follow Foldvari and Van Leeuwen (2011) and calculate the expected future wage flow arising from education. Human capital is then treated in parallel with investments: the price of an asset, like a bond or a stock, will tend to be the present value of all expected future flows of income from it. Since, the present value of the future expected labour income of a worker, assuming continuous time and his/her retirement age at 65, can be expressed as:

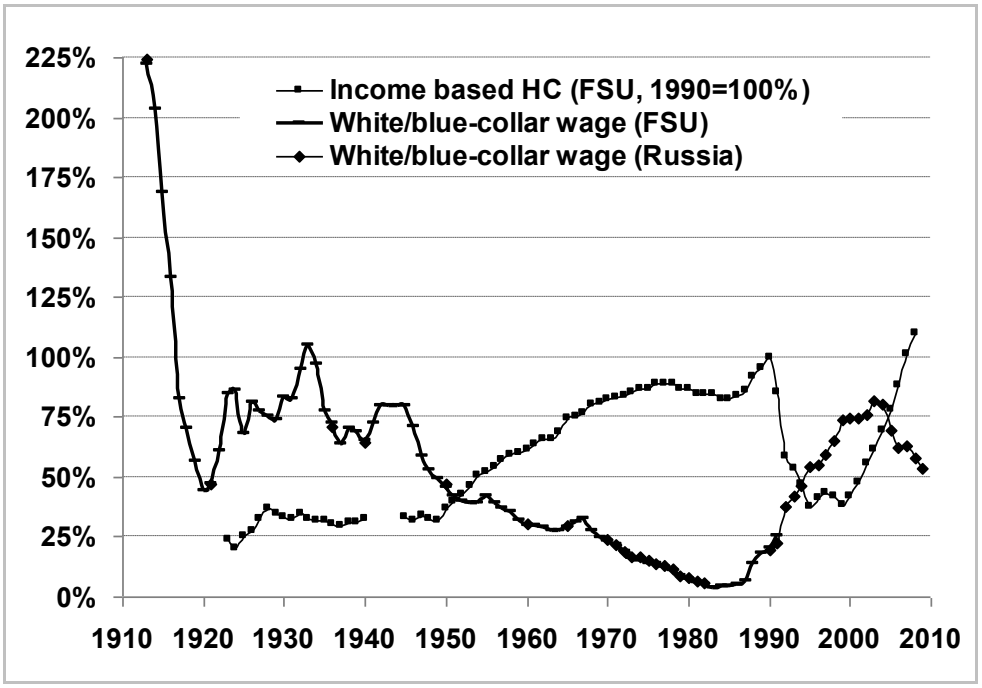
$$\bar{h} = \int_{t=0}^{65-\bar{x}} \bar{w} e^{(g-q)t} dt = \frac{\bar{w}}{g-q}\left(e^{(g-q)(65-\bar{x})} - 1\right)$$

where $\bar{h}$ is per worker stock of human capital in monetary units, $\bar{w}$ is actual average wage, $\bar{x}$ is the average age in the population, g is constant rate of expected real wage growth and q is the discount factor. We assume that q-p=0.02, as people expect their utility resulting from higher wages will increase with time, in line with Dagum and Slottje (2001) at micro-level.

This measure is not affected by intra-country wage differentials. However, if we assume that earnings of unschooled workers were the same among the FSU republics in a particular year then their difference in average wages would display the rewards for schooling. Including future earnings of unschooled allows us to capture not only private but also social returns to education if their wages increase due to investments into education made by other individuals or the state.

Notably, in the FSU the fluctuations of the human capital income-based measure tend to move reversely with those of white/blue-collar wage differential in industry (see Figure 1). This

Figure 1: Income-based HC and white/blue-collar wage differential in the FSU (1910s-2000s)



highlights the pattern where positive social returns to education are gained in much at the expense of private ones. Although similar trends may be observed in more developed countries with market

economy, the wage compression in the FSU appeared to be rather sharp and astonishing in 1910s (even before the Bolshevik Revolution) and particularly reinforced in mid-1940s – early 1980s.

The result of various human capital measurement results is given in below Table 7. Basically all series move about in the same direction, while the income-based measure is strongly influenced by abrupt real wage dynamics, especially in 1990s. However, it has been recovered at least by 2008. One more special thing to note is that, when corrected for the urban/rural population change, income based human capital in

Table 7: Human capital in the FSU

	Average years of education*	Cost based**	Income based***	Income based****
		1990 GK dollars	*1990 GK dollars*	*1990 GK dollars*
1924	1.6	128	36,390	
1940	3.5	1,510	59,014	60,189
1959	5.1	3,140	109,859	116,956
1970	6.4	5,180	150,004	156,555
1979	8.1	8,580	157,371	166,264
1989	9.8	11,673	174,014	190,593
2000			74,805	80,085
2008			198,433	212,439

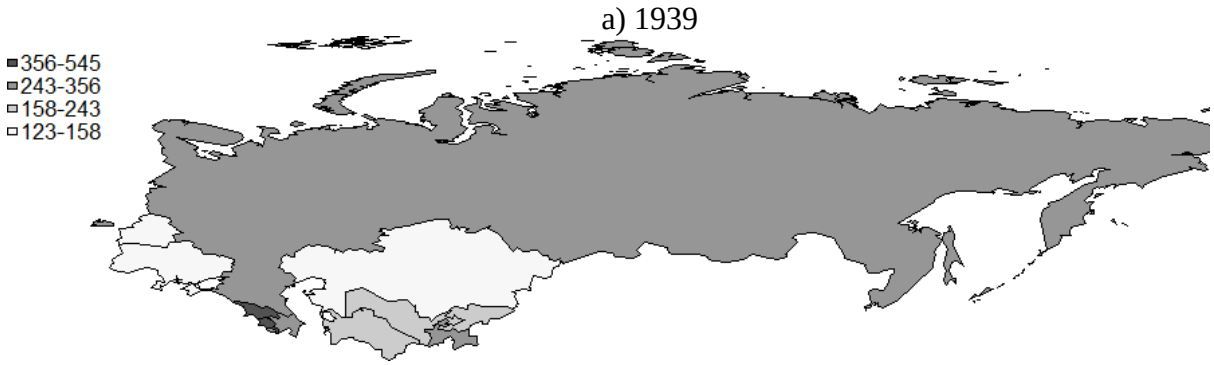
* Population at the age of 10+.
** Per capita (population at the age of 15+) stock calculated based on education expenditures data for the entire FSU.
*** Per worker stock calculated based on average wage data for public (basically urban) sector of the FSU economy (1924), average wage data for the entire FSU (1940-1989) and average wage data for the NIS (2000-2008).
**** Per worker stock calculated based on average wage data for the FSU republics (NIS), corrected for their change in urban/rural population ratio and weighted by their labour force.

the FSU is slightly lower in the 1920s and slightly higher in the 1970-2000s. However, we expect that this change will be bigger in countries with a larger agricultural sector.

Clearly, various human capital indicators kept growing in all of the FSU republics but unequally, so the republics could change their positions (see Maps 1-5, borders of 1989). E.g., Russia appeared to be the loser in all 3 rankings while Central Asian and, to some extent, Transcaucasian republics advanced in average years of education. This was probably due to previous relatively high expenditures, above the USSR average level. However, they produced only limited effect on these republics ranking in income-based measure. In this respect they did not change much over time.

Map 1: Book titles per million persons

a) 1939



b) 1989

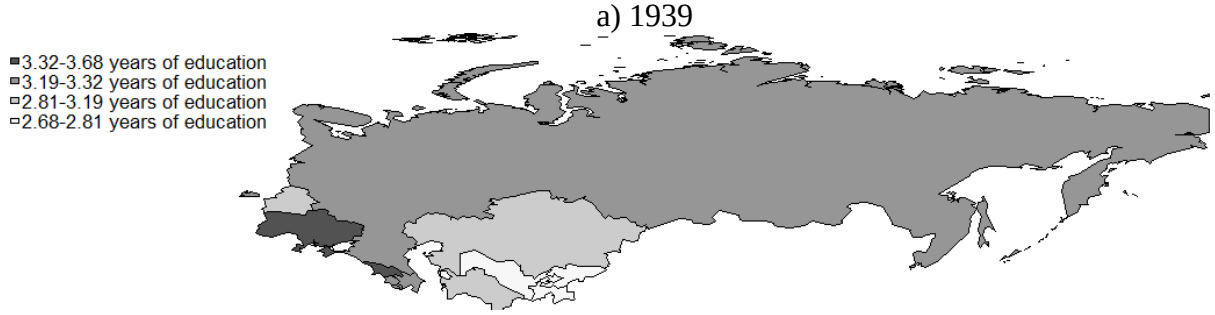
Map 2: Numeracy as measured by ABCC index
(indicates the % persons able to report their own age correctly, i.e. to count)

a) 1926

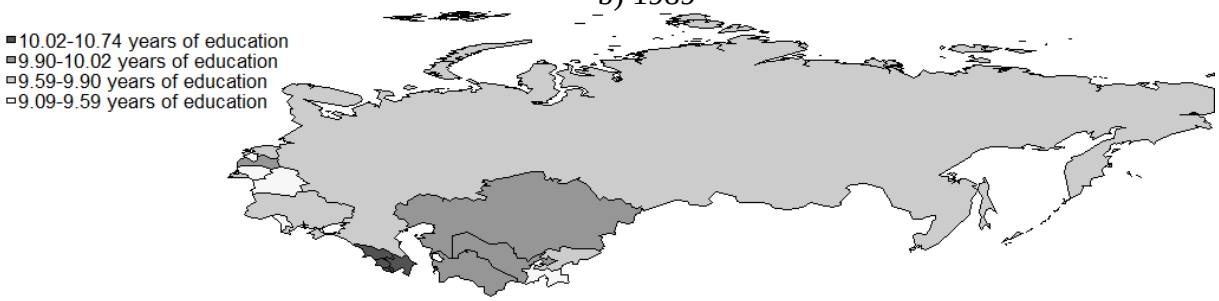
b) 1979

Map 3: Average years of education

a) 1939



b) 1989



Map 4: Cost based HC per worker

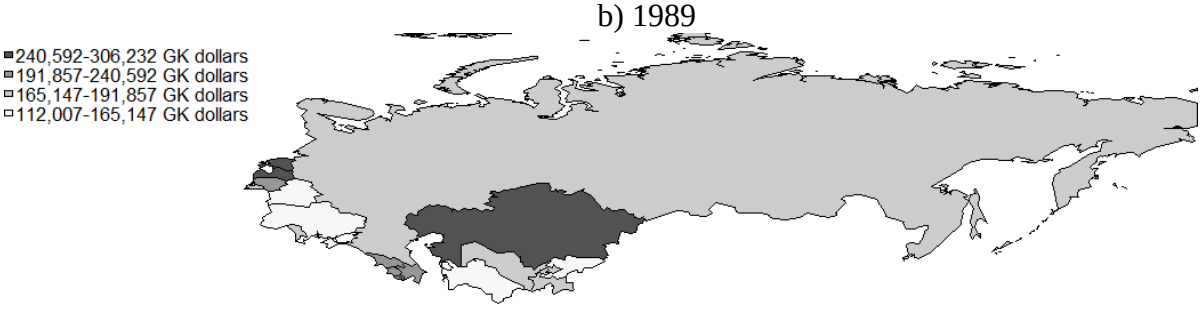
a) 1939

b) 1989

Map 5: Income based HC per worker

a) 1939

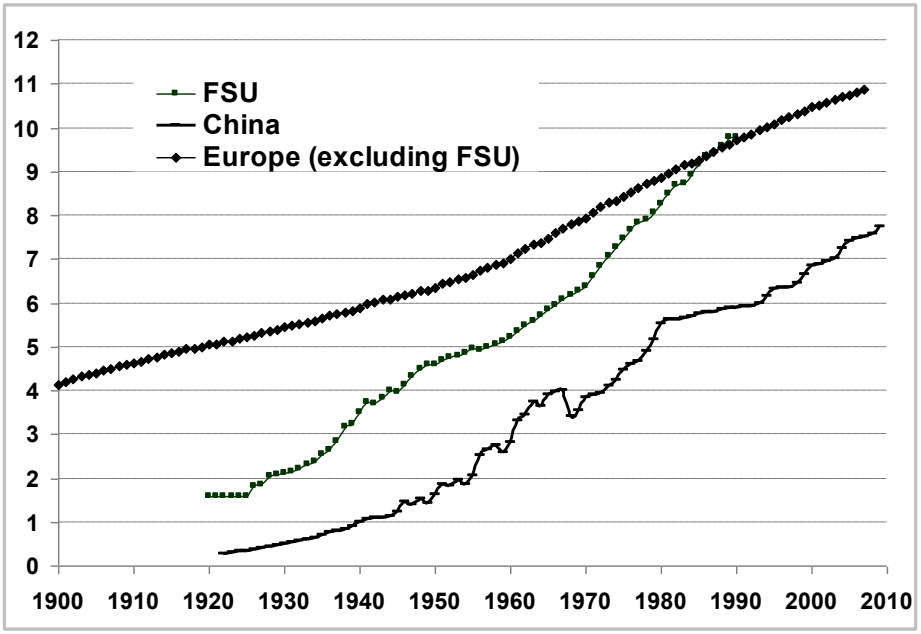
b) 1989



4. The spread of human capital in the FSU in a comparative perspective

The development of human capital in the USSR has been quite remarkable in an international perspective. Comparing with China, both countries started with a low cost based on human capital measure. However, where China started from almost the absolute 0-level, the USSR

Figure 2: Average years of education in Europe, China and the USSR



Source: Foldvari, Van Leeuwen and Van leeuwen-Li (2009); Van Leeuwen et al. (2011); own calculations

already had quite a human capital base in the 1920s. In that respect they more represented Europe (see Figure 2). In addition, it witnessed a fast growth by catching up to Europe in average years of education (but probably not in cost- or income-based human capital). Indeed, looking at

Figure 3: Cost based human capital per capita in China and the USSR (1990 GK dollars)

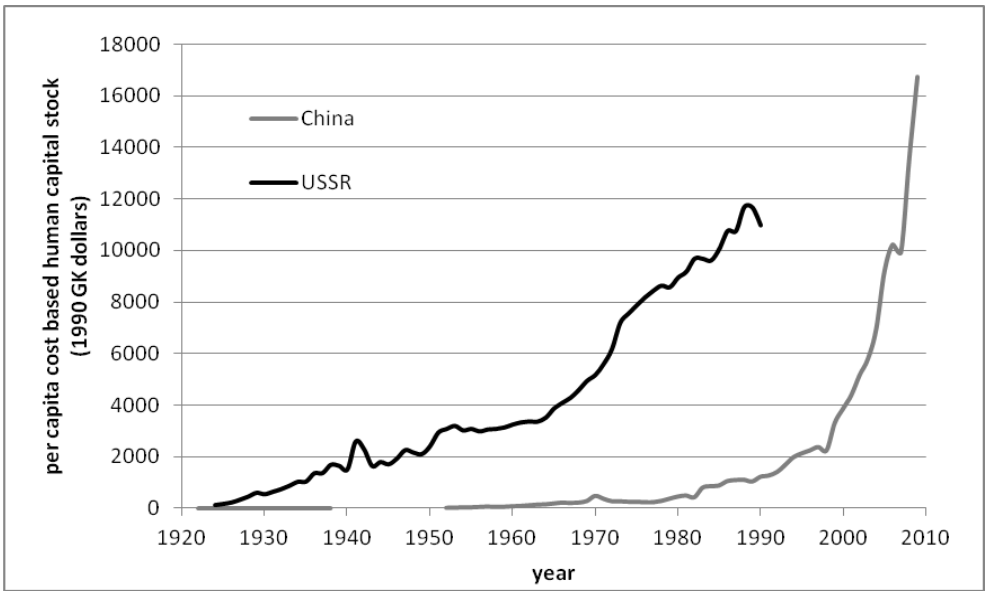


Figure 3, we note that the human capital in China in recent years grows much faster than it did in the USSR in the most part of the twentieth century.

Possibly, the faster development of the USSR in the early twentieth century is one of the reasons it outperformed China during that period. Indeed, human capital played a very important role. This period of fast growth of human- and physical capital is also the period with the highest negative TFP growth (Table 8). As pointed out by Van Leeuwen, Van Leeuwen-Li and Foldvari (2011), in

Table 8: GDP growth and TFP

	Factor share of human capital (HC)	Factor share of physical capital (FC)	Growth of GDP	Growth of HC	Growth of FC	TFP growth
			FSU			
1920-1940	40%	60%	6%	18%	8%	**-6%**
1950-1966	40%	60%	4%	4%	10%	**-4%**
1966-1977	40%	60%	3%	7%	5%	**-3%**
1978-1993	40%	60%	1%	2%	3%	**-2%**
1994-2006	40%	60%	3%			
			China			
1920-1940						
1950-1966	53%	47%	2%	16%	7%	**-10%**
1966-1977	44%	56%	2%	1%	5%	**-1%**
1978-1993	54%	46%	6%	12%	9%	**-5%**
1994-2006	54%	46%	8%	15%	11%	**-5%**

China this was largely caused by a reduction in technical efficiency of the factors of production: the continuous increase in human- and physical capital reduced the returns, while general technical change kept increasing. In the later part of the century, when the growth of the factors of production slowed down, technical inefficiency did not decline so much anymore, and TFP grew increasingly positive since it became largely driven by general technical development. However, this apparently did not work for the USSR since economic growth continued to be low. Only after the fall of socialism and deep decline in 1990s economic growth recovered in end 1990s - 2000s. The basic argument is that technical inefficiency reduced, which allowed for more TFP growth. Given our previous discussions, this may be caused either by integration of human capital (lagging provinces and countries grow harder in terms of human capital), because market economies have more efficient allocation of capital, or because general technical development became faster, possibly because more knowledge came available via the West.

For China, this growth was largely caused by a reduction in technical inefficiency paired with increased general technical growth. The provinces, however, remained as divided in terms of human capital as they had been in the 1920s.

Despite the evidence is limited at this moment, this process was similar for the USSR. If we compare different FSU republics (see Table 9), we see that factors of production kept growing in all of them, notwithstanding their level of economic development. Of course this did vary by human capital indicator. Whereas human capital inequality across republics in terms of age heaping went down (unsurprisingly since numeracy went up), inequality in books per capita went up considerably. However, the cost- and

Table 9: Human capital indicators in the FSU

			Age heaping	Books (no. titles per mln people)	Average years of education	Cost based HC	Income based HC (average wage)**	Income based HC (average wage, corrected for urban/rural change)***
1939*	USSR		97%	227.8	3.2	1,649	85,054	59,014
	of which	Armenia	94%	530.2	3.3	1,634	69,616	47,186
		Azerbaijan	91%	351.4	3.3	1,856	70,974	50,081
		Belarus	97%	147.2	3.2	1,298	47,647	32,996
		Georgia	89%	400.1	3.7	1,930	63,975	44,017
		Kazakhstan	97%	102.3	3.0	2,517	92,659	64,746
		Kyrgyzstan	95%	240.1	2.7	1,730	67,053	44,029
		Russia	98%	297.1	3.2	1,931	94,132	66,642
		Tajikistan	87%	190.1	2.7	1,563	80,485	49,188
		Turkmenistan	92%	231.8	2.9	2,483	84,058	58,298
		Ukraine	99%	152.7	3.4	1,083	59,523	42,855
		Uzbekistan	90%	160.8	2.8	1,111	58,592	39,770
	Gini		**1.0**	**14.3**	**2.2**	**10.5**	**10.8**	**8.9**
1989**	USSR		100%	268.5	9.8	11,673	174,014	190,593
	of which	Armenia	100%	301.6	10.6	19,319	190,096	
		Azerbaijan	99%	171.0	10.7	14,473	169,670	
		Belarus	100%	292.6	9.5	15,313	171,773	
		Estonia	100%	1317.7	9.9	25,581	244,758	
		Georgia	98%	365.1	10.6	20,909	167,605	
		Kazakhstan	100%	118.9	9.9	17,157	218,182	
		Kyrgyzstan	100%	236.2	9.9	14,804	152,980	
		Latvia	99%	722.5	10.0	21,980	228,378	
		Lithuania	100%	729.6	9.4	21,565	181,588	
		Moldova	99%	339.3	9.1	15,994	151,491	
		Russia	100%	313.0	9.8	12,189	206,924	
		Tajikistan	100%	169.0	9.6	12,381	152,100	
		Turkmenistan	100%	185.0	9.9	12,840	118,211	
		Ukraine	100%	164.1	9.7	12,336	170,612	
		Uzbekistan	100%	116.6	10.0	11,320	127,990	
	Gini		**0.0**	**20.8**	**1.1**	**14.1**	**7.3**	

* 1940 for income-based measures.

** Per worker stock; for the FSU republics calculated based on average wage data for public (basically urban) sector of their economies; for the USSR calculated based weighted average for its republics (by their labour force).

*** Per worker stock calculated based on average wage data for the FSU republics (NIS), corrected for their change in urban/rural population ratio and weighted by their labour force.

income-based measures, inequality remains about equal, suggesting there is no catch up and investment in human capital remains constant, irrespective of its level. This is similar as was noticed in China, where rich provinces witnessed equal (or even faster) growth of human capital,

irrespective of its level of income. This increased technical inefficiency (i.e. a lower return to capital).

5. Conclusion

In this paper, we develop a new dataset on human capital and related indicators for the former USSR area, most of them between ca. 1920 and 2000. This fills a gap in the literature since so far very few estimates of this vast area have been made available.

We use official statistics, combined with more recently available information and secondary literature. Combing all information we arrive at consistent estimates of literacy, book production (no. titles and total book output), average years of education, and cost- and income-based human capital measures. In addition, we add information on physical capital, GDP, and labour force.

We find that the USSR and its republics increased its human capital fast in the most part of the twentieth century. However, very little integration took place among the Republics. Also some of the indicators provide evidence on deterioration in human capital spending level and its quality during late Soviet era. This most likely caused increased technical inefficiency, causing a reduction in TFP growth. The same situation occurred in China. However, whereas China managed to keep technical inefficiency relatively moderate and, in addition, managed to increase general technology, in the former USSR area we do so far not find much evidence for a similar development. Just a few promising signs of the recovery appeared in 2000s.

References

A'Hearn, B., Baten, J., and D. Crayen (2009), "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital,"*Journal of Economic History*, 68(3), 783-808.

Abramitzky, R, and I. Sin (2010), *Book translations as idea flows: The Effects of the collapse of Communism on the diffusion of knowledge?* SIEPR Discussion Paper 09-032.

Alesina, A. and R. Perotti (1996), "Income Distribution, Political Instability, and Investment," *European Economic Review*, 40(6), 1203-1228.

Allen, R. (2003), *Farm to Factory: A Reinterpretation of the Soviet Industrial Revolution*, Princeton: Princeton University Press.

Andreev, E., E. Darskii, and T. Khar'kova, (1993), *Naselenie Sovetskogo Soiuza (1922-1991)* (Population of the Soviet Union (1922-1991),. Moscow: Nauka (in Russian).

Andreev, E., E. Darskii, and T. Khar'kova (1998), *Demograficheskaia istoria Rossii: 1927-1959* (Demographic history of Russia: 1927-1959), Moscow: Informatika (in Russian).

Barro, R. and J-W. Lee (2010), "A New Data Set of Educational Attainment in the World, 1950-2010," NBER Working Paper No. 15902.

Barro, R., and J.-W. Lee (2001), "International Data on Educational Attainment: Updates and Implications," *Oxford Economic papers*, 53(3), 541-563.

Baten, J., and J.L. van Zanden (2008), "Book production and the onset of modern economic growth," *Journal of Economic Growt*h, 13(3), 217-235.

Becker, A. (1969), *Soviet national income, 1958-1964: National accounts of the USSR in the seven year plan period.* Berkeley: Univ. of California Press.

Bergson, A. (1961), *Real National Income of Soviet Russia since 1928,* Cambridge (Mass.): Harvard University Press.

Chapman, J. (1963), *Real Wages in Soviet Russia since 1928,* Cambridge, Mass.: Harvard University Press.

Cohen, D., and M. Soto (2007), "Growth and human capital: good data, good results," *Journal of Economic Growth*, 12(1), 51-76.

Dagum, C., and D. Slottje (2000), "A New Method to Estimate the Level and Distribution of the Household Human Capital with Application," *Journal of Structural Change and Economic*

Dynamics, 11(1–2), 67–94.
Davis, R., and S. Wheatcroft (1994), "The crooked mirror of Soviet economic statistics," In: Davis, R.W. et al. (eds.). *The Economic Transformation of the Soviet Union, 1913-1945.* N.Y. – Melbourne: Cambridge University Press, 24-37.
De Witt, N. (1961), *Education and Professional Employment in the U.S.S.R.*, Washington: National Science Foundation.
Didenko, D. (2006), "Dinamika oplaty rabotnikov intellektual'nogo truda kak kharakteristika roli 'industrii znaniia' v SSSR" (The knowledge workers' wage dynamics as a proxy for the 'knowledge industry' performance in the USSR), In: Vinogradov, V.A. et al. (eds.). *Rossiia v kontekste mirovogo ekonomicheskogo razvitiia vo vtoroi polovine XX veka* (Russia in a context of the global economic development during the second half of the XXth century), Moscow: Moscow University Press, 276–288 (in Russian).
Easterly, W, and S. Fischer (2001), "The Soviet Economic Decline: Historical and Republican data," *World Bank Policy Research Working Paper* Number 1284, April 2001. Also published in the *World Bank Economic Review* 9(3) 1995, 341-371.
Eisenstein, E. (1979), *The Printing Press as an Agent of Change*, Cambridge: Cambridge University Press.
Federal State Statistics Service – Rosstat (2011). http://www.gks.ru, accessed November 2011.
Federal'noe Kaznacheistvo – Kaznacheistvo Rossii (2011), *The information on execution of the consolidated budget of the Russian Federation*. http://www.roskazna.ru/reports/cb.html, accessed November 2011.
Földvári, P., and B. van Leeuwen (2009), "Average years of education in Hungary: annual estimates 1920-2006," *Eastern European Economics*, 47(2), 5-20.
Gel'fand, V. (1992), *Naselenie SSSR za 50 let (1941 - 1990): Statisticheskii spravochnik* (Population of the USSR for 50 years (1941-1990): Statistical Abstract), Perm: Izdatel'stvo Permskogo Universiteta (in Russian).
Gerschenkron, A. (1947), "The Soviet Indices of Industrial Production," *Review of Economics and Statistics*, 29(4), 217-226.
Gosudarstvennyi komitet RF po stastistike (1992), *Sbornik analiticheskikh dokladov po materialam perepisi naseleniya 1989 goda* (Collection of analytical reports based on 1989 population census data), Moscow: Respublikanskii informatsionno-izdatel'skii tsentr (in Russian).
Gosudarstvennyi komitet SSSR po statistike (1988), *Trud v SSSR* (Labour in the USSR), Moscow: Finansy i Statistika (in Russian).
Gosudarstvennyi komitet SSSR po stastistike (1989), *Itogi Vsesoiuznoi perepisi naselenia 1979 goda. Pol, vozrast i sostoyanie v brake naseleniya SSSR* (Results of the 1979 All-Union population census. Gender, age and marital status of the population of the USSR), Tom II. Chast' 1. Statisticheskii sbornik. Moscow (in Russian).
Gosudarstvennyi komitet SSSR po statistike (1989a), *Narodnoe obrazovanie, nauka i kultura v SSSR. Statisticheskii sbornik* (Public education, science and culture in the USSR. Statistical abstract), Moscow: Finansy i Statistika (in Russian).
Gregory, P. (1982), *Russian National Income, 1885-1913,* London: Cambridge University Press.
Harrison, M. (1998), "Trends in Soviet Labour Productivity, 1928–1985: War, Postwar Recovery, and Slowdown," *European Review of Economic History*, 2(2), 171-200.
IMF, WB, OECD, EBRD (1991), *A Study of the Soviet Economy,* Paris.
Institute of Demography at the State Research University Higher School of Economics (2011), Demoscope Weekly, http://demoscope.ru, accessed November 2011 (in Russian).
Johnson, W. (1950), *Russia's educational heritage,* Pittsburg (Pennsylvania): Carnegie Press.
Judson, R. (2002), "Measuring Human Capital Like Physical Capital: What Does It Tell Us?" *Bulletin of Economic Research*, 54, 209-31.
Khanin, G. (1991), *Dinamika ekonomicheskogo razvitiia SSSR* (The dynamics of economic development of the USSR), Novosibirsk: Nauka, Sibirskoe otdelenie (in Russian).
Kovankovski, P. (1956), *Biudzhet SSSR: Istoriko-kriticheskii obzor* (The budget of the USSR:

Historical and critical review), München: Institute for the Study of the USSR (in Russian).
Krueger, A., and M. Lindahl (2001), "Education for growth: why and for whom?" *Journal of Economic Literature*, 39(4), 1101–1136.
Krumin, G. (ed.) (1923), *Narodnoe khoziaistvo Rossii za 1921/22 god. Statistiko-Ekonomicheskii Ezhegodnik* (National economy of Russia in 1921/22. Yearbook of economic statistics), Moscow: Ekonomicheskaia Zhizn' (in Russian).
Krumin, G. (ed.) (1924), *Narodnoe khoziaistvo S.S.S.R. za 1922-1923 god. Statistiko-Ekonomicheskii Ezhegodnik* (National economy of the USSR in 1922/23. Yearbook of economic statistics), Moscow - Leningrad: Ekonomicheskaia Zhizn' (in Russian).
Lucas, R. (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), 3-42.
Maddison, A. (2010), *Statistics on World Population, GDP and Per Capita GDP, 1-2006 AD, last update: February 2010*. http://www.ggdc.net/MADDISON/oriindex.htm, accessed November 2011.
Markevich, A., and M. Harrison, M. (2011), "Great War, Civil War, and Recovery: Russia's National Income, 1913 to 1928," *Journal of Economic History*, 71(3), 672-703.
Ministerstvo Finansov SSSR (1957), *Gosudarstvennye biudzhety soiuznykh respublik v piatoi piatiletke: Statisticheskii sbornik* (State budgets of the union republics in the fifth five-year-plan period: Statistical abstract), Moscow: Gosfinizdat (in Russian).
Ministerstvo Finansov SSSR (1962, 1966, 1972, 1976, 1982), *Gosudarstvennyi biudzhet SSSR i biudzhety soiuznykh respublik* (State budget of the USSR and budgets of the union republics), Moscow (in Russian).
Ministerstvo Finansov SSSR (1987-1990), *Gosudarstvennyi biudzhet SSSR* (State budget of the USSR), Moscow Finansy i Statistika (in Russian).
Ministerstvo Finansov SSSR (1958), *Raskhody na sotsial'no-kul'turnye meropriiatia po edinomu gosudarstvennomu biudzhetu SSSR: Statisticheskii sbornik* (Expenditures on social and cultural purposes of the consolidated state budget of the USSR. Statistical abstract), Moscow: Gosfinizdat (in Russian).
Mironov, B. (1985), "Gramotnost' v Rossii 1797-1917 godov" (Literacy in Russia in 1797-1917). *Istoria SSSR*, 4, 137-153 (in Russian).
Mironov, B. (1991), "The Development of Literacy in Russia and the USSR from the Tenth to the Twentieth Centuries," *History of Education Quarterly*, 31 (2), 229-252.
Mironov, B. (1994), "Ekonomicheskii rost i obrazovanie v Rossii i SSSR в XIX – XX vekakh" (Economic growth and education in Russia and the USSR in XIX – XX centuries), *Otechestvennaia istoria*, 4-5, 111-125 (in Russian).
Mironov, B. (2003), *Sotsial'naia istoria Rossii perioda imperii (XVIII - nachalo XX v.): Genezis lichnosti, demokraticheskoi sem'i, grazhdanskogo obshchestva i pravovogo gosudarstva*, Vol. 1, 2. 3rd ed. Saint-Petersburg: Dm. Bulanin (in Russian). Also published in English (with less detailed data): Mironov, B. (2000), *The Social History of Imperial Russia, 1700-1917*, Vol. 1, 2. Boulder: Westview Press.
Mironov, B. (2004), "Zhiznennyi uroven' v Sovetskoi Rossii pri Staline po antropometricheskim dannym" (The standard of living in the Soviet Russia under Stalin by anthropometric data). In: Borodkin, L.I and Petrov, Yu.A. (eds.), *Ekonomicheskaia istoriia. Ezhegodnik, 2004* (Economic History Yearbook, 2004), Moscow: ROSSPEN, 565-588 (in Russian).
Moorsteen, R., and R. Powell (1966), *The Soviet Capital Stock, 1928-1962*, Homewood (Illinois): Richard D. Irwin, Inc.
Morrisson, C., and F. Murtin (2009), "The Century of Education, Journal of Human Capital. Volume 3, Issue 1, Page 1–42, Mar 2009.
Narodnyi Komissariat Finansov SSSR (1939), *Gosudarstvennyi biudzhet Soiuza SSR za vtoruiu piatiletku (1933 -1937 gg.)* (State budget of the Union of the SSR in the second five-year-plan period), Moscow: Gosfinizdat (in Russian).

Narodnyi Komissariat Finansov SSSR (1939a), *Raskhody na sotsial'no-kul'turnye meropriiatia po edinomu gosudarstvennomu biudzhetu SSSR (soiuznomu, respublikanskomu i mestnym) za 1 i 2 piatiletki* (Expenditures on social and cultural purposes of the consolidated state budget of the USSR (union, republican and local) in the 1st and 2nd five-year-plan periods), Moscow: Gosfinizdat (in Russian).

Noah, H. (1966), *Financing Soviet Schools,* N.Y.: Russian Institute of Columbia University, Teachers College Press.

Ofer, G. (1987), "Soviet economic Growth: 1928-85," *Journal of Economic Literature*, 25(4), 1767-1833.

Perotti, R. (1996), "Growth, Income Distribution, and Democracy: What the Data Say," *Journal of Eonomic Growth*, 1(2), 149-187.

Plopeanu, A., P. Foldvari, B. van Leeuwen, and J.L. van Zanden (2012), "Where do Ideas come from? Book production and patents in a global and temporal perspective," CGEH working paper no. 33.

Plotnikov, K. (1954), *Ocherki istorii biudzheta sovetskogo gosudarstva* (Essays on history of the budget of the Soviet state), Moscow: Gosfinizdat (in Russian).

Poliakov, Iu. (ed.) (1992), *Vsesoiuznaia perepis' naselenia 1939 g. Osnovnye Itogi* (All-Union 1939 population census. Principal results), Moscow: Nauka (in Russian).

Poliakov, Iu. (ed.) (1999), *Vsesoiuznaia perepis' naselenia 1939 g. Rossia* (All-Union 1939 population census. Russia), Moscow: Nauka (in Russian).

Poliakov, Iu. (ed.) (2007), *Vsesoiuznaia perepis' naselenia 1937 g. Obshchie Itogi. Sbornik dokumentov i materialov* (All-Union 1937 population census. Aggregated results. Collection of documents and data), Moscow: ROSSPEN (in Russian).

Ponomarenko, A. (2002), *Retrospektivnye natsional'nye scheta Rossii: 1961-1990* (Retrospective national accounts of Russia: 1961-1990), Moscow: Finansy i statistika (in Russian).

Rogovin, V. (1982), "Raspredelitel'nye otnoshenia kak faktor intensifikatsii proizviodstva" (Distributive relations as a factor of production process intensification), *Sotsiologicheskie issledovania*, 1, 7-18 (in Russian).

Romer, P. (1990), "Endogenous Technological Change," *Journal of Political Economy*, 98(5), 71-102.

Slavic Research Center, Hokkaido University, *Soviet Economic Statistical Series*, http://src-home.slav.hokudai.ac.jp/database/SESS.html, accessed March 2012.

State University – Higher School of Economics (2005), *Biudzhetnoe finansirovanie obrazovaniia. Informatsionnyi biulleten'* (Budgetary financing of education. Information bulletin), Moscow: GU – VShE.

State University – Higher School of Economics (2007, 2010). *Obrazovanie v Rossiiskoi Federatsii: Statisticheskii Ezhegodnik* (Education in Russian Federation: Statistical Yearbook), Moscow: GU – VShE (in Russian).

State University – Higher School of Economics. (2010a). Ekonomika obrazovaniia: itogi monitoringa, 2009. Informatsionnyi biulleten' (Economics of education: monitoring results, 2009). Moscow: GU – VShE.

Statisticheskii Komitet Sodruzhestva Nezavisimykh Gosudarstv – StatKom SNG (1992), *Strany-chleny SNG: Statisticheskii ezhegodnik* (CIS member countries: Statistical Yearbook), Moscow: Finansovyi Inzhiniring (in Russian).

Steinberg, D. (1990), *The Soviet Economy 1970 - 1990: A statistical analysis*, San Francisco: Intern. Trade Press.

Subbotina, K. (1965), *Narodnoe obrazovanie i biudzhet* (Public education and the budget), Moscow: Finansy (in Russian).

Troinitskii, N. (ed.) (1905), *Obshchii svod po Imperii rezul'tatov razrabotki dannyh pervoi vseobshchei perepisi naseleniia, proizvedennoi 28 ianvaria 1897 goda* (General aggregation on Empire level of the data processing results of the compulsory population census effected on 28 January 1897), Vol. I. St.-Petersburg (in Russian).

Tsentral'noe buro statistiki truda VTsSPS, TsSu, NKT – TsSBT (1924), *Trud v SSSR. Statistiko-Ekonomicheskii Obzor. (oktiabr' 1922 - mart 1924 g.)* (Labour in the USSR. Statistical and economic review), Moscow: Izdanie VTsSPS (in Russian).
Tsentral'noe statisticheskoe upravlenie SSSR (1924), *Narodnoe khoziaistvo Soiuza SSR v thifrakh: Kratkii spravochnik* (National economy of the Union of the SSR in figures: Short Abstract), Moscow (in Russian).
Tsentral'noe statisticheskoe upravlenie SSSR (1929), *Statisticheskii spravochnik SSSR za 1928* (Statistical Abstract of the USSR in 1928), Moscow: Statisticheskoe Izdatel'stvo TsSU SSSR (in Russian).
Tsentral'noe statisticheskoe upravlenie SSSR (1956), *Kul'turnoe stroitel'stvo SSSR: Statisticheskii sbornik* (Construction of culture in the USSR: Statistical Abstract), Moscow: Gosstatizdat (in Russian).
Tsentral'noe statisticheskoe upravlenie SSSR (1963, 1985). *Zhenshchiny i deti v SSSR: Statisticheskii sbornik* (Females and children in the USSR: Statistical Abstract), Moscow: Finansy i Statistika (in Russian).
Tsentral'noe statisticheskoe upravlenie SSSR (1968, 1975, 1983), *Trud v SSSR* (Labour in the USSR), Moscow: Statistika (in Russian).
Tsentral'noe statisticheskoe upravlenie SSSR (1971, 1977), *Narodnoe obrazovanie, nauka i kultura v SSSR: Statisticheskii sbornik* (Public education, science and culture in the USSR. Statistical abstract), Moscow: Statistika (in Russian).
Tsentral'noe statisticheskoe upravlenie SSSR – Gosudarstvennyi komitet SSSR po statistike (various years), *Narodnoe khoziaistvo SSSR* (National economy of the USSR), Moscow (in Russian).
Tsentral'noe statisticheskoe upravlenie RSFSR (various years), *Narodnoe khoziaistvo RSFSR* (National economy of the RSFSR), Moscow (in Russian).
Tsentral'noe upravlenie narodnokhoziaistvennogo uchota SSSR (1932-1936), *Trud v SSSR* (Labour in the USSR), Moscow (in Russian).
Tsentral'noe upravlenie narodnokhoziaistvennogo uchota SSSR (1937), *Zhenshchina v SSSR* (Female in the USSR: Statistical Abstract), Moscow: Soiuzorguchot (in Russian).
Tsentral'noe upravlenie narodnokhoziaistvennogo uchota SSSR (1940), *Kul'turnoe stroitel'stvo SSSR: Statisticheskii sbornik* (Construction of culture in the USSR: Statistical Abstract), Moscow - Leningrad: Gosplanizdat (in Russian).
United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics (2011). http://stats.uis.unesco.org, accessed November 2011.
Van Leeuwen, B., and P. Földvári (2011), "Capital accumulation and growth in Central Europe, 1920-2006," CGEH working paper no. 23.
Van Leeuwen, B., and P. Foldvari (2011), "Capital accumulation and growth in Hungary, 1924-2006," *Acta Oeconomica* , 61(2), 143-164.
Van Leeuwen, B., and Földvári, P. (2008), "How Much Human Capital does Eastern Europe have? Measurement Methods and Results," *Post-Communist Economies*, 20(2), 189–201.
Van Leeuwen, B., and P. Földvári (2008), "Human Capital and Economic Growth in Asia 1890-2000: a time-series analysis," *Asian Economic Journal*, 22(3), 225-240.
Van Leeuwen, B., J. van Leeuwen-Li, and P. Földvári (2011), "Regional human capital in Republican and New China: Its spread, quality and effects on economic growth," Paper presented at the workshop "Quantifying knowledge", Utrecht, 9-10 December 2011.
Volkov, E. (1930), *Dinamika Narodonaseleniia SSSR za 80 let* (Demographical dynamics of the USSR for 80 years), Moscow: Gosizdat (in Russian).
Williamson, J. (1998), "Globalization and the Labor Market: Using History to Inform Policy," In: Aghion, P. and Williamson, J. (eds.), *Growth, Inequality and Globalization: Theory, History, and Policy,* N.Y. – Cambridge: Cambridge University Press, 103-193.
World Bank (1992), "Measuring the Incomes of Economies of the Former Soviet Union," Policy Research Working Paper 1057, December 1992.

World Bank (2011), “Data by Country,” http://api.worldbank.org, accessed November 2011.
Zaleski, E. (1980), *Stalinist Planning for Economic Growth, 1933 - 1952*, Chapel Hill: The University of North Carolina Press.