



Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Konrad-Zuse-Strasse 1 · D-18057 Rostock · GERMANY
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;
<http://www.demogr.mpg.de>

MPIDR WORKING PAPER WP 2014-008
AUGUST 2014

**A golden age before serfdom?
The human capital of Central-Eastern
and Eastern Europe in the
17th-19th centuries**

Jörg Baten
Mikołaj Szoltysek (szoltysek@demogr.mpg.de)

This working paper has been approved for release by: Frans Willekens (willekens@demogr.mpg.de),
Head of the Research Group on International Migration.

© Copyright is held by the authors.

Working papers of the Max Planck Institute for Demographic Research receive only limited review.
Views or opinions expressed in working papers are attributable to the authors and do not necessarily
reflect those of the Institute.

A GOLDEN AGE BEFORE SERFDOM? THE HUMAN CAPITAL OF CENTRAL-EASTERN AND EASTERN EUROPE IN THE 17TH-19TH CENTURIES

By Jörg Baten and Mikołaj Szoltysek

SUMMARY

Can the 16th and early 17th centuries in Poland-Lithuania and some other east-central European countries be characterized as a “Golden Age” in human capital? We trace the development of a specific human capital indicator during this period: numeracy. We draw upon new evidence for Poland and Russia from the early 17th century onwards; and for Belarus, Ukraine, and Lithuania from the 18th century onwards; controlling for potential selectivity issues. Poland had quite high levels of numeracy during the early 17th century, but these levels subsequently fell below those of even southern Europe. As in other countries in the area, numeracy levels in Poland were lower than those of western Europe during the 17th, 18th, and early 19th centuries. This finding might support the hypothesis that the second serfdom process, which gained momentum during the 17th century, was one of the core reasons why human capital accumulation was delayed in eastern Europe. The major wars in the region also had devastating effects on numeracy levels.

KEYWORDS

Central-Eastern Europe; historical Demography; Eastern Europe; Human Capital; Numeracy; Age-Heaping; census microdata

CONTACT DETAILS

Jörg Baten, Department Economics, Univ. Tübingen, CEPR and CESifo. Email: joerg.baten@uni-tuebingen.de (corresponding author)

Mikołaj Szoltysek, Laboratory of Historical Demography, Max Planck Institute for Demographic Research, Rostock; szoltysek@demogr.mpg.de

Acknowledgements: We wish to thank Inna Vasylchuk for her very important input in the first phase of the project. We also wish to thank Nicholas Meinzer, Nadine Raimann, Mikołaj Malinowski, Max-Stephan Schulze, Tamas Vonyo, Jacob Weisdorf, Nikolaus Wolf, Jan Luiten van Zanden, participants in the Warsaw FRESH meeting, the Utrecht human capital workshop, and members of the Tübingen economic history research group for their comments. Financial assistance was provided by the HIPOD EU program and ClioInfra (N.W.O.). Important input came from participants in the ClioInfra meeting in Utrecht 2011.

Introduction

A number of scholars have argued that eastern central Europe was once a region with a high standard of living, even in comparison to western Europe. Van Zanden (1999) found that wages expressed as the purchasing power of grain were higher in the Polish cities of Warsaw and Krakow during the 16th and early 17th centuries (and in Lviv, which is in today's west Ukraine, during the 16th century) than in many western European locations.¹ Using a broader cost-of-living basket, Allen (2001) showed that during the 16th century, English and Polish cities had similar real wage levels (based on maxima in 1599), and that the English cities did not pull ahead of the Polish cities in terms of wage levels until the 17th century. A recent study by Malinowski (2013) found that, relative to their counterparts in England, urban workers in Poland even had a substantial real wage advantage (based on barebones baskets), but that when the real wages of rural workers were also taken into account, Poland was roughly on par with England during the late 16th century, and had lower wage levels before and after.²

Anthropometric indices are useful complements to these results because they also include the relatively large groups of farmers and other self-employed individuals; although the sample sizes are relatively small for the early modern period. Koepke and Baten (2005, 2008) found that northern and eastern European health and nutrition levels

¹ Of course, the cities for which evidence is available may not be representative for all of eastern Europe. In fact, in our conclusion we will argue that western and central Poland, to which they refer, still had a remarkably high level of numeracy in the early 17th century, unlike in other regions.

² The situation of small farmers is slightly less clear, as estimating their income requires us to make assumptions about their marketing of products (Malinowski 2013, p. 17) .

were more favorable than those of western and southern Europe. According to estimates in Koepke and Baten (2005, Table 3), during the 16th century the average height of eastern European men was 171.4 cm, compared to 170.4 cm among British men, 169.3 cm among southern German men, and 170.0 cm among the men of the "North Rhine" (Holland/western Germany).³

However, during the 19th century, real wages as well as human capital (which is typically correlated with income), were clearly lower in eastern Europe. These indicators were, of course, affected by economic changes, wars, and social transformations over the centuries. What happened? And what were the determinants of these economic changes? In an attempt to answer these and related questions, we will study the development of numeracy in this region. Numeracy is clearly a core component of human capital, especially in agricultural societies, as individuals who were making decisions about the timing of activities had to take a number of issues into account, such as the weather, the status of plants and animals, and other similar variables. We will therefore use techniques for measuring age-heaping to compare numeracy in several eastern and central-eastern European regions with evidence from western and southern Europe. Although Poland belongs to central-eastern Europe, we will use the term eastern Europe in the following for reasons of brevity.

Our sources include various kinds of census and census-like microdata from the territories of historical Poland-Lithuania, the Russian Empire, as well as regional data for 1880 in Prussia and Austria-Hungary. In each case, the listings contained descriptions of the characteristics of each of the individuals who were grouped into a given household,

³ During the 17th century, however, the heights of eastern Europeans started to decline sharply.

including the individual's age, sex, marital status, and his or her relationship to the household head (see more in Data section).

These sources allowed us to estimate numeracy in several regions of what is today Poland, Belarus, Ukraine, Lithuania, and Russia. The application of age-heaping-based numeracy estimates to this newly available dataset is performed here for the first time for such a large region and for the time frame (but see Mironov 1991 and Kaiser and Peyton 1993 on Russian samples). We carefully discuss the potential selectivity biases of these sources. While social selectivity is not a major problem due to the non-exclusive “census-type” character of these sources, regional selectivity is a large challenge that requires special treatment in this study. Given the regional character of our sources, we compare the places we can cover with representative nationwide census evidence of the early 19th century.⁴

The remainder of the study is organized as follows. We first review some findings and hypotheses of the previous literature regarding the reasons why eastern Europe lagged behind western Europe in educational and welfare levels. In the second section, we present our new dataset and explain its capabilities and limitations. In Section 3, we briefly explain our age-heaping methodology, and address some concerns scholars may have about it. We also offer some responses to potential criticism of age-heaping proponents. In Section 4, we discuss potential selectivities of this new evidence, report the results at the regional level, and present a method that can be used to adjust for

⁴ We decided to aggregate the numeracy estimates using today's national borders rather than historical empires or other regional units. This will allow for the comparison of the estimates with other historical evidence in the future (such as GDP estimates, anthropometric welfare, and other indicators).

regional selectivities. Finally, in Section 5, we present national estimates since the 17th century for the five eastern European countries in our sample, and compare these estimates with evidence from western and southern Europe. Finally, we present a tentative discussion of the implications of our findings for our understanding of economic growth in early modern Europe.

1. Review of the literature

In his study of literacy rates around 1800, Reis (2005) reported a rather low value for Hungary (six percent), the only eastern (or central-eastern) country he investigated. This value was substantially lower than the values found in other European countries. Mironov (1991) studied Russian literacy over a very long period of time. He cited the estimate by the Russian historian Sapunov that a mere one to 1.5 percent of the Russian population in the mid-13th century may have been literate before the Mongol invasion (based on the assumption that monks, clergymen, and the upper strata of secular society were literate). Mironov reported that, by the end of the 17th century, the number of books, records, and similar literacy-related items had increased. His estimates of literacy rates based on the signatures of witnesses in legal sources yielded a very rough literacy rate of between two and 2.5 percent for the late 17th century. Finally, after organizing the 1897 census by birth cohorts, he arrived at the following rough estimates of literacy rates for Russia: four percent in 1800, 13 percent in 1850, and 30 percent in 1900.

By contrast, the literacy rates in western Europe were estimated at between 15 and 65 percent in the early modern period until around 1800 (A'Hearn et al. 2009, p. 802). Mironov also looked at local samples of Baltic peasants and other sources to assess the

degree of age-heaping, but did not organize the data by birth cohorts of adults. Kaiser and Peyton (1993), who studied the urban communities of Tula and Viatka around 1700, unearthed very important evidence, but also did not conduct a cohort analysis. A'Hearn et al. (2009) argued that eastern Europe lagged behind the west in numeracy.

What might have caused the relatively low educational levels and the relatively modest welfare levels observed in eastern Europe of the 19th century? A number of prominent explanations for these adverse developments in eastern Europe have been given in the previous literature.

a) Hajnal famously argued that differences in the age at marriage and other aspects of household formation behavior differed between eastern and western Europe. He identified a border following a line between St. Petersburg and Trieste, which might have left most of the Baltic and western Poland in the “western” part; and Ukraine, Russia, Belarus, and eastern Poland in the “eastern” part. It is possible that early marriage might have resulted in less educational investment per child. As early as in 1970, Hajnal’s observations had been rejected as being too simplistic by J. Sklar in her dissertation, and more recently by several other scholars who offered even harsher criticisms of Hajnal’s work (Sklar 1970, Plakans and Wetherell 2005, Szołtysek 2007, 2008, 2012a, 2014). Yet despite these challenges, the Hajnal hypothesis remains a persistent stereotype in economic history, as well as in the demographic literature. The nature of this discussion may be influenced by the traditional dividing line between economists, who tend to accept simplifications and the statistical concept of the average; and family historians, who tend to adopt more nuanced and contextualized perspectives that stress micro- and meso-level variation (see Szołtysek 2012b).

But it is also possible that the east-west divide was caused by other factors mentioned below. We should note upfront that the Hajnal hypothesis is considered here with very strong reservations.

b) A lack of “girl power” (de Moor and van Zanden 2010, similarly Foreman-Peck 2011) may have also played a role. De Moor and van Zanden have argued that in the west, and especially in the North Sea region, women had more customary rights in the labor market and in other aspects of family economies (such as inheritance; see, however, Guzowski 2010 and Dennison 2011, and Dennison and Ogilvie 2013 for criticism). Educational gender inequalities might have led to the people of eastern Europe having less education on average, as the women were mainly responsible for basic education in the household.

c) The second serfdom hypothesis is also commonly cited in the economic history literature (Kula, 1976; Millward 1982; Cerman 2008; Ogilvie and Edwards 2000; also Sosnowska 2004). Scholars have noted that historical Poland and Russia in particular were affected by noble landlordism and village subjection (Hagen, 1998; also Mironov, 1996).⁵ The dramatic expansion of the powers landlords had over the rural population in these areas was closely related to a rapid rise in agricultural commodity values in the west caused by the 16th-century “price revolution.” The eastern European landowners responded to this trend by expanding their previously modest familial manor farms into

⁵ These were: juridical subjection, migration regulations, legal attachment to a particular social status, subjection to communal payments and duties (including the harshest forms of compulsory labor), limited right to private property, limited choice of occupation, and unprotected personal dignity; see Mironov, 1996, p. 323.

large-scale domanial economies designed to produce surpluses for sale on the urban markets of western Europe. This type of seigneurialism prompted landlords to demand from their peasant subjects not only rents in cash and kind, but above all labor services, which were essential to the very functioning of the demesne farms (Szołtysek 2008a). Serfs therefore had relatively few incentives or opportunities to invest in the kind of basic education which would have enabled them to understand the numeracy concept applied in this study.

There were of course very different forms of serfdom (Cerman 2012). The most extreme form was the manorial system based on peasants' personal and hereditary subjection, as well as on their labor obligations (*corvée*) to the manors. While this system was introduced in the territories of Poland-Lithuania during the 16th to early 17th centuries, the strongest manorial systems developed in western Poland and in some parts of Ukraine (esp. Volhynia). On the other hand, there were parts of eastern Europe where this type of serfdom was never fully introduced (Polessia in Belarus; Subcarpathian Ukraine)⁶. For example, in large areas of the historical Grand Duchy of Lithuania, a softened version of the system—based on cash quitrents rather than on *corvée*, or a mixture of the two—emerged in the second half of the 17th century, and prevailed until the end of the Polish republic in 1795 (Szołtysek 2008, earlier Kozlovskij 1961).

As a minor digression from the literature review, we will look briefly at the regional distribution of serfs, as opposed to free or manumitted persons, and of the people who lived on government-owned estates.

⁶ Please note that the relatively “weak” serfdom system that covered large parts of the Belarusian territories of the late Polish Republic turned into more stringent forms of serf exploitation and control after these territories came under Russian governance after the final partition of Poland in 1795.

An overwhelming majority of the population of all of the territories under investigation lived in personal and hereditary subjection up until the 19th-century reforms, with their property rights limited to an indeterminate leasehold. For this period, however, the share of hereditary or emphyteutic freeholders (i.e., peasants who held more advantageous property rights) appears to decline substantially and continuously when moving eastward across the Polish territories towards Russia (Rutkowski, 1986; Moon 1999). Moreover, if we look at the distribution of the share of serfs in the Russian Empire during the mid-19th century, a clear regional pattern emerges (Figure 3).⁷ Especially in a central corridor between Belarus (Minsk) and Nishniy Novgorod, the share of serfs was particularly large. By contrast, there were relatively few serfs in the thinly populated regions of both the northeast and the southeast. During the late serfdom period, the southeast in particular had a slightly less oppressive system of *Obrok* (defined as feudal obligations that were paid in money or kind), whereas the *corvée* system of compulsory labor was more typical in other regions. The share of serfs actually corresponds quite well with the regional distribution of numeracy and literacy. In Figures 1 and 2, the ABCC index of numeracy is compared with the numeracy rate. The former is defined as the share of persons who probably know their age with an annual resolution.⁸

⁷ Note, however, that no data are provided for Polish territories.

⁸ The ABCC Index used in these figures reports the share of individuals in a given society who probably know their true age (named after A'Hearn, Baten and Crayen as well as Greg Clark, who developed that measure). The formula is

$$ABCC = \left(1 - \frac{(Wh - 100)}{400} \right) \times 100 \text{ if } Wh \geq 100; \text{ else } ABCC = 100.$$

The index ranges from zero to 100. If everybody reports their correct age, ABCC has a value of 100.

d) Large-farm agriculture is often associated with a political economy in which large landowners prevented tax-financed public schooling, as they saw no need for serfs to be educated (and perhaps learn how to demand political rights) in schools financed by the taxes of the rich. Even after the abolition of serfdom, many agricultural laborers had no access to schooling.⁹

e) Similarly, the political leaders of the Russian Empire did not favor education, and the Polish population in Prussia received insufficient schooling in their native language (Baten, 2003). In Austria-Hungary, investments in education were also insufficient (Cvrcek, 2013). The Russian Empire was dominated by the nobility, who were landed property owners. Thus, investment in schooling was not very high on the national agenda in Russia until the Empire lost the Crimean War. This major defeat shined a spotlight not just on the military inferiority of Russia, but also on the Empire's backwardness in terms of productivity and human capital. Thereafter, the government implemented reforms which included investments in larger-scale schooling and literacy promotion (Mironov, 2012, p. 297).

f) The frequent wars, including civil wars, waged in eastern Europe might have fostered an attitude of risk aversion among the population regarding investments of any sort. The terrible damage of the mid-17th-century wars, followed by yet more carnage in the 1720s, led to the development of the manorial economy, and to land-labor ratios that

⁹ "Agricultural laborers" were only part of the socioeconomic landscape of eastern Europe for the period after the formal abolition of serfdom; they emerged quite early in Galicia due to the Josephinian reforms of 1780s, and then in the province of Greater Poland (Prussian Province of Posen) after the 1820s.

were not very different from those of the late 17th century; i.e., at the beginning of the agrarian change (Szołtysek 2008a).

d) Lower life expectancy may have had a similar effect. Life expectancy at birth, e_0 , was probably lower in this part of Europe than in the west, although our evidence for this assumption is still relatively weak. At least for historical Poland, this is only a tentative argument based on single case studies using different methodologies; it is widely believed that values of 27 for men and around 27-28 for women were typical in eastern Europe in the late 18th century, whereas some western European countries had values above 30 (Szołtysek 2014; also Kuklo 2009).

h) Low population density and the lack of a transport system made commuting to schools more costly and returns on schooling lower in much of the east. There were, however, regions in historical Poland with high population densities, such as Lesser Poland around Krakow, Galicia, and the regions close to the Baltic Sea controlled by Prussia. The 17th-century wars resulted in sharp reductions in population densities in some areas. The regions of western and central Poland, as well as of western Galicia, generally had the highest population numbers in the country. Moving to the eastern areas, we observe a gradual decrease in population density; late 18th-century Belarus had densities of well below 10 persons/km² (Szołtysek 2014; earlier Karpachev and Kozlovskij 1972).

i) Religion: Could the absence of Protestantism, or the lack of religious competition, have played a role (Baten and van Zanden 2008)? While most of Slavic Russia was Orthodox, the religious pattern was more mixed in the west. Calvinism emerged among the Polish nobility in the 16th century and dominated this political class

well into the mid-17th century, or even longer. Meanwhile, the western fringes of Poland were inhabited by religiously mixed communities, with a substantial share of Protestants (mainly settlers from different parts of Germany and the Netherlands). The degree of religious fractionalization was much lower in the Polish eastern “borderlands” throughout the early modern period (Szady 2010, 228-250). Even there, however, religiously mixed communities or regions could be found, with the peasantry dominated by Uniates (and later members of the Orthodox faith), and not seldom intermingled with islets of Polish-speaking minor nobility. There was also a substantial Jewish population. Indeed, the competition between the Uniates and the Orthodox believers on the one hand, and the Catholics and the Protestants on the other, may have been of equal importance in this context (Mitterauer 2003).

Some of these potential determinants will be discussed below using interregional comparisons, while others will be discussed using east-west comparisons. However, the aim of the present paper is not to perform a regression in which these potential explanations are systematically tested against each other.

2. Sources

Our sources are: (1) the “lists of souls” (either the Roman *Catholic Libri Status Animarum* or their Protestant *Seelenregister* equivalents); (2) the censuses of the Civil-Military Order Commissions 1790–1792 in the territories of Poland-Lithuania (which were occupied by the Russian Empire after the late 18th century); (3) the Russian *revizii* (tax-oriented censuses); (4) the censuses of 1880 in Prussia and Austria-Hungary (which had occupied other territories of modern Poland) and 1897 in Russia; and (5) other types

of household lists, including “communion books” and local administrative surveys, as well as private and Crown estate inventories.¹⁰

In Table 1, we report on the places and regions for which evidence is available. In the second column, we list the country in which the regional unit is situated today. In the following three columns, we indicate the county, the larger district, and the empire in which these places were situated towards the late 19th century. It should be noted that, of the places situated in Prussia or Austria-Hungary, we included only those with a large majority of Polish speakers. This was done to avoid placing German-speaking communities in the category of “today’s Poland,” as their descendants might later have fled or been moved to Germany after WWII. In the ethnic overlap between Poland, Lithuania, and Belarus, we have been less restrictive. There might be some migration biases later on. We also took care not to include any sources in which some cross-checking by priests or officials might have taken place. In those cases, there was almost no heaping present. We only included county-birth decade averages that were based on at least 50 observations (for the number of cases, see an appendix available from the authors).

¹⁰ All these sources, except for 19th-century censuses and the 17th/18th-century sources on Russia, are the part of the *CEURFAMFORM Database* developed by M. Szołtysek. The database development was supported by the Marie Curie Intra-European Fellowship project (FP6-2002-Mobility-5, Proposal No. 515065) at the Cambridge Group for the History of Population and Social Structure, Cambridge, UK, 2006-2008. More details in Szołtysek 2008, especially 2014.

3. Assessing human capital formation with the age-heaping indicator and other measures of human capital

Measuring the production factor “human capital” has never been simple, as advanced forms of skills are difficult to compare. All economists have therefore resorted to using proxy indicators, such as the share of people signing a marriage register. Grundlach (2001) noted that the empirical measurements of contributions to economic growth of the human capital factor and the productivity of education which have been performed in human capital research up to this point are still not completely satisfying. A comparison of different proxy indicators might be the best option for obtaining reliable insights. This is the rationale for using the age-heaping methodology (as well as comparisons of literacy and schooling, wherever they are available to us). We will explain the advantages and caveats in somewhat greater detail, as the application of this method in economic history is still relatively new.

This approach employs the set of methods that developed around the phenomenon of “age-heaping;” i.e., the tendency of poorly educated people to round their age erroneously. For example, less educated people are more likely than people with a greater endowment of human capital to state their age as “30,” even if they are in fact 29 or 31 years old (Mokyr 1985).¹¹ Crayen and Baten (2008) found that the relationship between illiteracy and age-heaping for LDCs after 1950 is very close. They calculated age-

¹¹ Among demographers, this specific type of age misreporting constitutes “one of most frustrating problems” (Ewbank 1981, 88). It is treated as a source of distortion in age-specific vital rates which needs to be removed, or at least minimized in order to study the family or household variables.

heaping and illiteracy rates for not less than 270,000 individuals who were organized by 416 regions, ranging from Latin America to Oceania. Their findings indicated that the correlation coefficient with illiteracy was as high as 0.7, and that the correlation with the PISA results for numerical skills was as high as 0.85. They therefore concluded that the age-heaping measure “Whipple Index” is more strongly correlated with numerical skills. A’Hearn, Baten, and Crayen (2009) used a large U.S. census sample to perform a very detailed analysis of this relationship. They subdivided the sample by race, gender, high and low educational status, and other criteria. In each case, they obtained a statistically significant relationship. It is also remarkable that the coefficients were found to be relatively stable between samples; i.e., a unit change in age-heaping was associated with similar changes in literacy across the various tests. These results are not only valid for the U.S.: in all of the countries studied so far that had substantial age-heaping, the correlation was found to be both statistically and economically significant.¹²

In order to assess the robustness of those U.S. census results and the similar conclusions which could be drawn from the less developed countries of the late 20th century, A’Hearn et al. (2009) also assessed age-heaping and literacy in 16 different European countries between the middle ages and the early 19th century. Again, they found a positive correlation between age-heaping and literacy, although the relationship was somewhat weaker than the relationship found in an analysis of data from the 19th and 20th centuries. It is likely that the unavoidable measurement error when using early

¹² On the regions of Argentina, see, for example, Manzel, Baten, and Stolz (forthcoming).

modern data induced the lower statistical significance.¹³

To the best of our knowledge, the widest geographical sample studied so far was created by Crayen and Baten (2010). This sample included 70 countries for which both age-heaping and schooling data (as well as other explanatory variables) were available. They found in a series of cross-sections between the 1880s and 1940s that primary schooling and age-heaping were closely correlated, with R-squares between 0.55 and 0.76 (including other control variables, see below). Again, the coefficients were shown to be relatively stable over time. This large sample also allowed them to examine various other potential determinants of age-heaping. To assess whether factors such as the degree of bureaucracy, birth registration practices, and government interaction with citizens influenced the likelihood that an individual would know his or her exact age, independent of personal education, Crayen and Baten used the number of censuses performed for each individual country up to the period under study as an explanatory variable for their age-heaping measure. Except for countries with a very long history of census-taking, all of the variations of this variable turned out to be insignificant, which would suggest that an independent bureaucracy effect was rather weak. In other words, it appears that societies in which a large number of censuses were conducted and birth registers were introduced early had a high degree of age awareness. But those societies also introduced schooling

¹³ Historical demographers often find that data from premodern times are very rough, imprecise, or fragmentary. Even the 18th-century statistical materials are a veritable jungle of uncertainties and traps, as they were sometimes collected haphazardly and analyzed without care; as a result, they often capture only a portion of the phenomenon they are seeking to describe, and are thus incomplete (Szołtysek 2011, 2014). The quality of the data on age may be particularly poor.

early, and this was the variable that clearly had more explanatory power than the independent bureaucracy effect. Crayen and Baten also tested whether the general standard of living had an influence on age-heaping tendencies (using height as well as GDP per capita as welfare indicators), and found a varying influence: in some decades, there was a statistically significant correlation, while in others there was none.

Was this correlation between numeracy and literacy also visible in eastern Europe? When comparing the log literacy in the Russian Imperial census of 1897 for the individuals born between 1825 and 1884 on the vertical axis, and their numeracy levels (see Figure 3, expressed by the ABCC index) on the horizontal axis, we can see that there is a clear correlation. The values found in the Baltic regions of Estonia and Livonia, as well as in the capital region of St. Petersburg, were very high, whereas the “serfdom”-intensive regions around Belarus had quite low values of both literacy and numeracy.¹⁴

Interestingly, the northeastern districts of European Russia—such as Archangelsk, Wologda, and Perm—had much higher numeracy than literacy levels. The previous literature has noted that the existence of schools is even more important for the development of literacy than for the development of basic numeracy, as families are more likely to teach children the fundamentals of numeracy than of literacy. In the thinly populated regions of the northeast, attending school was much more difficult than it was in the more densely populated areas farther south.

In conclusion, the correlation between age-heaping and other human capital indicators is quite well established, and the “bureaucratic” factor does not invalidate this

¹⁴ Please note that, as serfdom was abolished in all these territories in 1864, these effects must have arisen in the first four decades. The district of Kowno was an outlier.

relationship. A caveat relates to other forms of heaping (apart from the heaping on multiples of five), such as heaping on multiples of two, which was quite widespread among children and teenagers, and to a lesser extent among young adults in their twenties.¹⁵ This shows that most individuals knew their age as teenagers, but that only those in well-educated societies were able to remember or calculate their exact age later in life. At higher ages, this heaping pattern was mostly negligible, but it was, interestingly, somewhat stronger among populations who were numerate enough not to round on multiples of five. We will exclude those below age 23 and above 72, since a number of distortions could affect those specific age groups, leading to age reporting behavior that may have differed from that of the adult group in between. Many young males and females married in their early twenties or late teens, when they also had to register as voters, military conscripts, etc. On such occasions, they were sometimes subject to minimum age requirements, a condition which gave rise to increased age awareness. Moreover, individuals in this age group were physically growing, which made it easier to determine their age with a relatively high degree of accuracy. All of these factors contributed to a reduction in age-heaping among children and young adults relative to levels observed among older adults. Because the age-heaping patterns of very old individuals were subject to upward as well as downward bias for the reasons mentioned above, the very old should also be excluded.

¹⁵ It has been shown that, in some societies, in addition to the usual overrepresentation of five and zero, there was also a decided preference for figures ending on other digits, and that the avoidance of certain numbers was likely to occur in a patterned way as well (Stockwell 1966; Nagi, Stockwell and Snavley 1973; also Myers 1940).

Some uncertainty remains as to whether the age-heaping found in the sources reflects the numeracy of the responding individual, or rather the diligence of the reporting personnel who wrote down the statements (e.g., Szoltysek 2011, 2014). The age data of the relevant age groups of 23-72 were normally derived from statements from the person himself or herself. However, it is possible that a second party, especially the household head, the father, or the husband, may have made or influenced the age statement; or even that the enumerator estimated the age without asking the individual (especially for lodgers, inmates, or other temporary household members like unrelated servants). In such cases, we would not be able to measure the numeracy of the person interviewed. By contrast, if the enumerator asked the person for his or her age and obtained no response, a round age estimated by the enumerator would still measure basic numeracy correctly. A large body of literature has investigated the issue of how to handle cases in which individuals did not report their own information. Foldvari et al. (2010) speculated, for example, that a wife may appear to have been more numerate than she actually was because she improved her age statement with the help of her husband. They compared the numeracy of married and unmarried women, and found that the latter had significantly lower numeracy in some of their samples. However, de Moor and Zuijderduijn (2013) recently challenged these findings with a number of good arguments. Moreover, in the early modern period and the 19th century, marriage was often associated with higher educational and social status, as a number of studies have found (for example, Baten and Murray 1998). Our comparison of male and female numeracy in our sample indicated that women were sometimes more numerate than men, which would support the hypothesis that the accuracy of age reporting may have been higher among women than

among men. On the other hand, we found a correlation between the male and female numeracy of different households. Friesen et al. (2011) recently compared systematically the evidence of a gender gap in numeracy and in literacy for the late 19th and early 20th centuries, and found a strong correlation. They argued that there is no reason why the misreporting of literacy and age should have yielded exactly the same gap between genders. A more likely explanation is that the well-known correlation between numeracy and literacy also applies to gender differences. For our study, the question of whether the women answered themselves is slightly less important, because we only seek to estimate average numeracy.

Moreover, there is sometimes direct evidence in the sources that the wives themselves were asked. Manzel et al. (2011) reported finding sources on Latin American Indio women in which statements like this one were included: “She says that she is 30, but she looks more like 40.” Even for black female (and male) slaves in the Cape Colony in South Africa who were accused of crimes, the legal personnel created a separate column that indicated whether the person was guessing her age, or whether she actually knew it. We can speculate that if these Indio and African women—who probably were not shown much respect by colonial officers—were asked to report their age, then European women—who were likely treated with a greater level of respect—might have also have been asked to report their age.

The problem of different enumerators influencing the quality of age statements has also been studied in a 20th-century context. While a large part of age misreporting indeed arises because the respondents do not know their exact age, this problem is likely to be exacerbated by differences in the quality of the performance of the enumerators, as

some of them may have taken their duties more seriously than others (United Nations 1952, 59). Referring to the notorious hardships encountered in the surveying processes in contemporary developing countries, Ewbank observed: “In particular, the training of interviewers, their level of education, and their ability to understand and pursue the interests of the researcher will significantly affect the quality of data [on age]” (Ewbank 1981, 15). However, the difference between the behavior of 20th-century enumerators and the priests and officials of the 17th to 19th century is that the former had much easier access to sources that would enable them to cross-check age statements. Priests of the 18th century could have looked up birth years in birth registers, but because the registers were usually chronologically sorted, the cross-checking of ages would have required a substantial investment of time. But as some of the existing sources were clearly cross-checked (yielding ABCC values of around 100 very early), we used a historian’s judgment in excluding them. In addition, Szoltysek (2011, 2014 – Appendix 2) found that differences in the age-heaping patterns in historical Poland-Lithuania might be partly attributable to differences in the organizing principles of the enumeration process that are inherent to different types of listings, and not to differences in the capabilities of the individuals being surveyed.

Of course, a potential bias always exists if more than one person is involved in the creation of a historical source. For example, if literacy is measured by analyzing the share of signatures in marriage contracts, there might have been priests who were more or less interested in obtaining real signatures, as opposed to just crosses or other symbols. We are reassured in our assumptions by the findings of previous studies, which generally indicate that age-heaping was much more prevalent (and numeracy levels were lower)

among members of the lower social strata, and among the half of the sample population who had lower anthropometric values (Baten and Mumme 2010). Moreover, studies have shown that the regional differences in the prevalence of age-heaping were similar to the regional differences in illiteracy. We can therefore conclude that the method of age-heaping is a useful and innovative tool for assessing human capital.

4. Potential selectivity issues and the adjustment of regional biases

Whenever a new dataset is used that does not reflect the whole population of a given country or region, the logical next question is about potential selectivity: Is the sample a selective subset of the population we are interested in, or does it more or less cover this population in a representative way? Clearly, perfect representativeness is probably impossible for historical samples of the 17th or 18th centuries, but we should at least consider whether the dataset has systematic biases that distort potential results in a significant way, and we should seek to minimize potential measurement error.

Social bias or labor market bias is not a large issue in our sample, given the census-type character of our sample: in principle, everybody in a given city or village should have been included in the dataset, and it is likely that only a few emigrants were absent at the time of the census. However, regional composition is a major issue because we do not have surviving sources for all cities and villages. How can we deal with this potential regional bias in our sample?

We present the ABCC estimates for the individual regions in Table A.1 in the appendix. On the left side of Table A.1, the new regional estimates for the period from the 1630s to the 1810s are presented; while on the right side, estimates for the period

from the 1820s to the 1900s are given. The latter set of estimates are based on the 1880 and 1897 population censuses, as well as later censuses; while the former set of figures are based on the sources mentioned in the data section. For Russia, five regions can be documented, sometimes for very different periods. To what degree are those regions representative? The fact that Moscow is included in the five documented regions suggests that there is probably upward bias. Hence, the next logical question would be: Did the regions have ABCC values similar to the average values found for Russia in the 1897 census? Or: How large was the upward bias? The analysis shows that four out of the five regions had ABCC values 15-25 percent above the Russian average (Column “Adjustment factor”). Only the Eyskij location in the Kuban territories south of Rostov/Don had values similar to the Russian average for the birth decade of the 1820s. Hence, we need a regional adjustment. For simplicity, we take the difference for the birth cohort of the 1820s and report the regionally adjusted values in Table A.2 in the appendix. This adjustment is based on the assumption that the interregional bias was similar in the early period and for the birth decade of the 1820s. This might not have been the case for all of the regions, but in general the estimate will be closer to the true national average after the adjustment than before. The fact that we normally have four to five different regional datasets to compare allows us to gain an impression of the size of the measurement error implied by this procedure. For example, in the case of Przemyslan and the birth decade of the 1730s, the resulting value is clearly too low; moreover, the Warsaw region might have been underestimated for the 18th century. But in the vast majority of cases, the regional adjustment procedure works relatively well. In order to remain consistent, we take all of the values into account.

5 Estimates for the five eastern European countries and international comparison

In a next step, we generate national estimates based on those regional values. In Figure 4, we display the regional and national estimates for Russia. Some of the early estimates are above and others are below the estimate for Russia, but the emerging trend seems to be relatively clear. Hence, we show the national trends for all five countries in Figure 5. We distinguish between the western and the eastern parts of today's Poland. The western part is made up of Silesia (part of Sudetenland, as well as Upper Silesia), which included German-speaking areas; other parts of Prussia; and districts that were annexed by Prussia and Austria-Hungary in the 18th-century partitions of Poland. The eastern part consists of the regions that were occupied by the Russian Empire. We were curious about whether the west and the east would yield similar estimates for the whole of Poland after being regionally adjusted to the national mean. In fact, the similarity of the levels found suggests that this division does not affect the estimates for Poland significantly, even if the variation over time is not identical.

Finally, our aim was to make those series graphically comparable with estimates for other European regions. This was achieved by using the LOWESS procedure, which was previously used by Manzel et al. (2011). In order to make the comparison, the eastern and western parts of Poland were considered together (Figure 6).

Eastern Europe in international comparison

What broad trends could be identified using this procedure, and how do they compare with those of other European regions? In Figure 7, data from eastern Europe were plotted

against the evidence from western and southern European countries, which we derived from Stolz et al. 2012 (see also Tollnek and Baten 2011). The authors assessed the northwestern and central European region (Austria, Germany, France, Sweden, and the UK), for which relatively continuous evidence from the 1730s is available, and the southern European region (Italy, Spain, Portugal). Both series start at around 80 percent numeracy in the early 18th century, but the northwestern region made more rapid progress, and achieved 95 percent numeracy around 1800. The northwest had solved the basic numeracy problem by around the middle of the 19th century. Numeracy in southern Europe stagnated at a quite high level of around 82 percent from the 1730s until the 1820s, and then slowly converged with northwestern European levels.

Earlier evidence suggests that during the 15th century, numeracy levels varied across Europe from 72 percent ABCC in the Netherlands, to 55 percent in northern Italy, to 40 percent in Germany, and down to 18 percent in southern Italy (A'Hearn et al. 2009). Juif and Baten (2011) found that Spain and Portugal had numeracy levels of around 60 percent in both the early and the late 17th century.

Hence, the northwestern and southern European regions were clearly more numerate than all of the eastern European regions we are assessing here during the 18th and 19th centuries, although Poland did not differ very much from the European south during the 17th century (values for the south from Juif and Baten 2011). Moreover, the trends of convergence and the slowdown in the individual regions are interesting. Russia started at a much lower level than Poland, or at around 20 percent in the early 17th century; but the gap between Russia and Poland had declined to less than five percent in

the mid-18th century. During the 19th century, human capital again started to accumulate, and the problem of basic numeracy was almost solved by around 1900.

Poland displayed stagnant levels of numeracy throughout much of the 17th and the early 18th centuries (around 60 percent), whereas the European south grew by some 20 ABCC points during this period. Levels of basic numeracy continued to increase in Poland during the middle decades of the 18th century, and, like in Russia, the problem was solved by around 1900. During the 19th century, a steady upward trend can be discerned in all of the eastern European regions.

However, among the countries studied here, Belarus, Lithuania, and Ukraine lagged behind the most. During the early to mid-18th century, numeracy still stood at around 20 percent in Lithuania, 40 percent in Belarus, and 50 percent in Ukraine. Ukraine then began to develop rapidly, which resulted in Ukrainian numeracy levels overtaking Russian levels during the 19th century. It would be interesting to investigate whether the migration of Jewish people from the Polish-Lithuanian regions to Ukraine also stimulated this surge in Ukrainian numeracy. Belarus and Lithuania experienced the most rapid growth in their numeracy levels during the 19th century.

The relatively large discrepancy between Polish and Russian levels early on, and the much larger gap during the 18th century between Russia and the territories of Belarus, Lithuania, and Ukraine, are among the major findings here.

Results and Conclusion

Serfdom seems to have played a key role in limiting human capital development in eastern Europe, as is apparent in the regional patterns we discussed in the digression from

the literature review. The earliest evidence we have on western Poland suggests that in the early 17th century, the region was not very far behind other regions of Europe. For example, it displayed a numeracy level similar to those of Portugal and Spain in 1600-49 and 1650-99 (Juif and Baten 2011). By contrast, Russia was probably at a much lower level during this period, whereas Belarus, Ukraine, and Lithuania started at very low levels when our evidence becomes available in the 18th century. The fact that western Poland was still doing relatively well during the early 17th century, but was not able to reach western European levels during the 17th, 18th and early 19th centuries—and even fell back relative to southern Europe during this period—might support a second serfdom hypothesis.

The reasons why the numeracy levels of eastern Europeans lagged behind those of western Europeans are complex. Like much of central and southern Europe, eastern Europe was negatively affected by a series of wars. In addition, educational progress in eastern Europe might have been hindered by the second serfdom process, which continued in the 17th and 18th centuries.

Finally, the observed differences in numeracy patterns between western and eastern Poland-Lithuania may be indicative of broader developmental disparities between the two parts of the country. Our findings regarding the differences in numeracy patterns correspond well with broader demographic differences between Poland proper on the one hand, and Belarusian and Ukrainian lands on the other, in terms of age at marriage, home-leaving behavior, and household formation and structure (Szołtysek 2014). Thus, the next step in the analysis might be to place Polish-Lithuanian and Russian numeracy patterns in a wider eastern European comparative perspective.

An explosion of new microdata sources which contain information on historical family composition, as well as the data needed to analyze numeracy patterns in Lithuania, Hungary, Romania, Bulgaria, Latvia, Serbia, and eastern Ukraine, is opening up unprecedented opportunities for exploring a much broader range of human capital developments (see Szołtysek and Gruber 2014; also Ruggles 2012)¹⁶.

¹⁶ The Mosaic Project, which began in 2011 at the Max Planck Institute for Demographic Research (MPIDR) in Rostock (www.censusmosaic.org), will be particularly valuable to researchers pursuing these questions. The project draws upon the experiences of a global community of researchers involved in global-scale international data infrastructure projects like IPUMS and NAPP. The project gathers, harmonizes, and distributes (openly) surviving census and census-like material from historic Europe, and from eastern Europe in particular. Currently, Mosaic provides data for 91 regions of historic Europe containing over 700,000 individuals living in 143,000 domestic groups from the Atlantic to the Urals (Szołtysek and Gruber 2014).

References

- Allen, R. C. (2001). The great divergence in European wages and prices. *Explorations in Economic History*, 38, 411–447. doi:10.1006/exeh.2001.0775.
- A'Hearn, Brian, Baten, Jörg, and Crayen, Dorothee and “Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital” *Journal of Economic History* 69-3 (Sept 2009), pp.783-808.
- Baten, Jörg. "Anthropometrics, Consumption and Leisure: The Standard of Living," in Sheilagh Ogilvie/Richard Overy (eds.), *Germany: A New Social and Economic History*, vol. III: 1800-1989. London: Edward Arnold, 2003, pp. 383-422
- Baten, Jörg and Christina Mumme: “Globalization and Educational Inequality in Long-Run Development during the 17th to 20th Centuries: Latin America and other Developing World Regions”, *Journal of Iberian and Latin American Economic History* 28-2 (2010), 279 -305. doi:10.1017/S021261091000008X.
- Baten, Jörg and John Murray (1998). "Women's Stature and Marriage Markets in Pre-Industrial Bavaria", *Journal of Family History* 23-2, pp. 124-135
- Baten, Jörg and Jan Luiten van Zanden (2008): “Book Production and the Onset of Early Modern Growth”, *Journal of Economic Growth* 13-3 , pp. 217-235,
- Cerman, M. (2008). Social structures and land markets in late medieval central and eastern-central Europe. *Continuity and Change* 23(1), 55-100.
- Cerman, M. (2012). *Villagers and Lords in Eastern Europe, 1300-1800*. Basingstoke: Palgrave Macmillan.
- Crayen, Dorothee and Baten, Jörg "New Evidence and New Methods to Measure Human Capital Inequality before and during the Industrial Revolution: France and the U.S. in the 17th to 19th Centuries " *Economic History Review* 53-2 (2010), pp. 452-478.
- Crayen, Dorothee and Baten, Jörg “Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth”, with Dorothee Crayen, *Explorations in Economic History* 47-1 (2010), pp. 82-99 doi:10.1016/j.eeh.2009.05.004.
- Crayen, Dorothee and Joerg Baten (2008): Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth, *CESifo Working Paper 2218*. [Extended Working Paper version of Crayen and Baten 2010]

- Cvrcek, Tomas and Miroslav Zajicek (2014). "School, what is it good for? Useful Human Capital and the History of Public Education in Central Europe: Paper Summary". *Journal of Economic History* 74-2, 592-3.
- De Moor, Tine and Jaco Zuiderduijn (2013). The Art of Counting: Reconstructing Numeracy of the Middle and Upper Classes on the Basis of Portraits in the Early Modern Low Countries, 41-56. *Historical Methods* 46-1, 41-56.
- De Moor, Tine and Jan Luiten van Zanden (2010) Girl power: the European marriage pattern and labour markets in the North Sea region in the late medieval and early modern period, *Economic History Review* 63-1, pp. 1-33.
- De Moor, Tine, and Van Zanden, J.-L., 'Uit fouten kun je leren. Een kritische benadering van de mogelijkheden van 'leeftijdstapelen' voor sociaal-economisch historisch onderzoek naar gecijferdheid in het pre-industriële Vlaanderen en Nederland', *Tijdschrift voor Economische en Sociale Geschiedenis*, 5 (2008), pp. 55-86.
- Dennison, T. (2011). Household formation, institutions, and economic development: Evidence from imperial Russia. *The History of the Family* Volume 16, Issue 4, 456-465.
- Dennison, T. K. and Ogilvie, Sheilagh, Does the European Marriage Pattern Explain Economic Growth? (May 31, 2013). CESifo Working Paper Series No. 4244. Available at SSRN: <http://ssrn.com/abstract=2274606>.
- Ewbank, Douglas C. (1981), Age misreporting and age-selective underenumeration: sources, patterns, and consequences for demographic analysis. Committee on Population and Demography, Report No. 4, Washington, D.C.: National Academy Press.
- Friesen, Julia, Joerg Baten and Valeria Prayon: "Women count. Gender (In-)Equalities in the Human Capital Development in Asia, 1900-1960", Working Paper Univ. Tuebingen.
- Földvári, Peter, Bas van Leeuwen, and Jieli van Leeuwen-Li. "How did women count? A note on gender-specific age heaping differences in the sixteenth to nineteenth centuries". Oxford: Blackwell Publishing. 2010.
- Guzowski, P. (2010). European Marriage Pattern in Central Europe: the case of late medieval and Early modern Poland. Paper presented at the 35th Annual Meeting of the

- Social Science History Association (SSHA), Palmer House Hilton, Chicago, Illinois (USA), November 2010.
- Hagen, W.W. (1998). Village life in East-Elbian Germany and Poland, 1400-1800: subjection, self-defense, survival. In T. Scott (ed.), *The peasantries of Europe from the fourteenth to the eighteenth centuries* (pp. 145-190). London: Longman.
- Juif, Dácil-Tania and Joerg Baten (2011) "On the Human Capital of 'Inca' Indios before and after the Spanish Conquest" Working Paper Univ. Tuebingen
- Kaiser, Daniel and Peyton Engel. "Time- and Age-Awareness in Early Modern Russia." *Comparative Studies in Society and History* 35, no. 4 (1993): 824–39.
- Karpachev, A. M., and Kozlovskij, P. G. (1972). 'Dinamika chislennosti naselenia Belorussii vo vtoroi polovine XVII –XVIII vv.', *Ezhegodnik po agrarnoj istorii Vostochnoj Evropy za 1968 god*, 81–94.
- Koepke, Nikola and Joerg Baten (2005), "The Biological Standard of Living in Europe During the Last Two Millennia," *European Review of Economic History* 9-1 (2005), pp. 61-95
- Koepke, Nikola and Jörg Baten (2008) "Agricultural Specialization and Height in Ancient and Medieval Europe", *Explorations in Economic History* 45 (2008), pp. 127-146.
- Kozlovskij, P.G. (1961). 'Zemelnyj nadel i povinnosti krestyan v magnatskich vladenyach Belorussii vo vtoroj polovine XVIII v.' *Ezhegodnik po agrarnoj istorii Vostochnoj Evropy*, 218-231.
- Kuklo, C. (2009). *Demografia Rzeczypospolitej przedrozbiorowej*. Warszawa: DiG.
- Kula, W. (1976). *An economic theory of the feudal system: towards a model of the Polish economy, 1500-1800*. London: N.L.B.
- Lyashchenko, Peter I. (1949). *History of the National Economy of Russia to the 1917 Revolution*, tr. Leon M. Herman. New York: Macmillan.
- Malinowski, M. (2013). East of Eden: Polish living standards in a European perspective, ca. 1500-1800, CGEH WP No. 43.
- Manzel, Kerstin, Stolz, Yvonne, Baten, Jörg "Convergence and Divergence of Numeracy: The Development of Age Heaping in Latin America, 17th to 20th Century", *Economic History Review* (2012, forthcoming).

- Millward, R. (1982). "An economic analysis of the organization of serfdom in Eastern Europe." *Journal of Economic History*, 42, 513-48.
- Mironov, Boris N. (1991). The Development of Literacy in Russia and the USSR from the Tenth to the Twentieth Centuries. *History of Education Quarterly*, Vol. 31, No. 2. (Summer, 1991), pp. 229-252.
- Mironov, B.N. (1996). When and why was the Russian peasantry emancipated? In M.L. Bush (Ed.), *Serfdom and slavery: studies in legal bondage* (pp. 324-347). Harlow : Longman.
- Mironov, Boris N., 'Novaya Istoricheskaya Demografiya Imperskoy Rossii: Analicheski obsor Corremennoy Istoriografii', *Vestnik Sankt-Peterburgskovo Universiteta*, 4 (2006), pp. 62-90.
- Mironov, Boris N. "The Standard of Living and Revolutions in Russia, 1700-1917.", Gregory L. Freeze (Ed.), New York: Routledge, 2012, p.297.
- Mitterauer, Michael: Warum Europa? Mittelalterliche Grundlagen eines Sonderwegs. München 2003
- Moon, D. (1999). *The Russian peasantry, 1600–1930: the world the peasants made*. London & New York: Longman.
- Myers, R. J. (1940). 'Errors and Bias in the Reporting of Ages in Census Data', *Transactions of the Actuarial Society of America*, 41 (104), 395–415.
- Nagi, M. H., E. G. Stockwell, and L. M. Snavley (1973). Digit Preference and Avoidance in the Age Statistics of Some Recent African Censuses: Some Patterns and Correlates. *International Statistical Review* Vol. 41, No. 2, pp. 165-174.
- Nafziger, Steven "Communal Institutions, Resource Allocation, and Russian Economic Development: 1861-1905," *The Journal of Economic History* (June 2008): 570-575.
- Nafziger, Steven "Peasant Communes and Factor Markets in Late Nineteenth-Century Russia," *Explorations in Economic History* 47.4 (2010): 381-402.
- Ogilvie, S., J. S. S. Edwards (2000). Women and the Second Serfdom: Evidence from Early Modern Bohemia. *Journal of Economic History*, 60, 961-994.
- Plakans, A. & Wetherell, C. (2005). The Hajnal line and Eastern Europe. In T. Engelen & A.P.Wolf (Eds.), *Marriage and the family in Eurasia. Perspectives on the Hajnal hypothesis* (pp. 105-126). Amsterdam: Aksant.

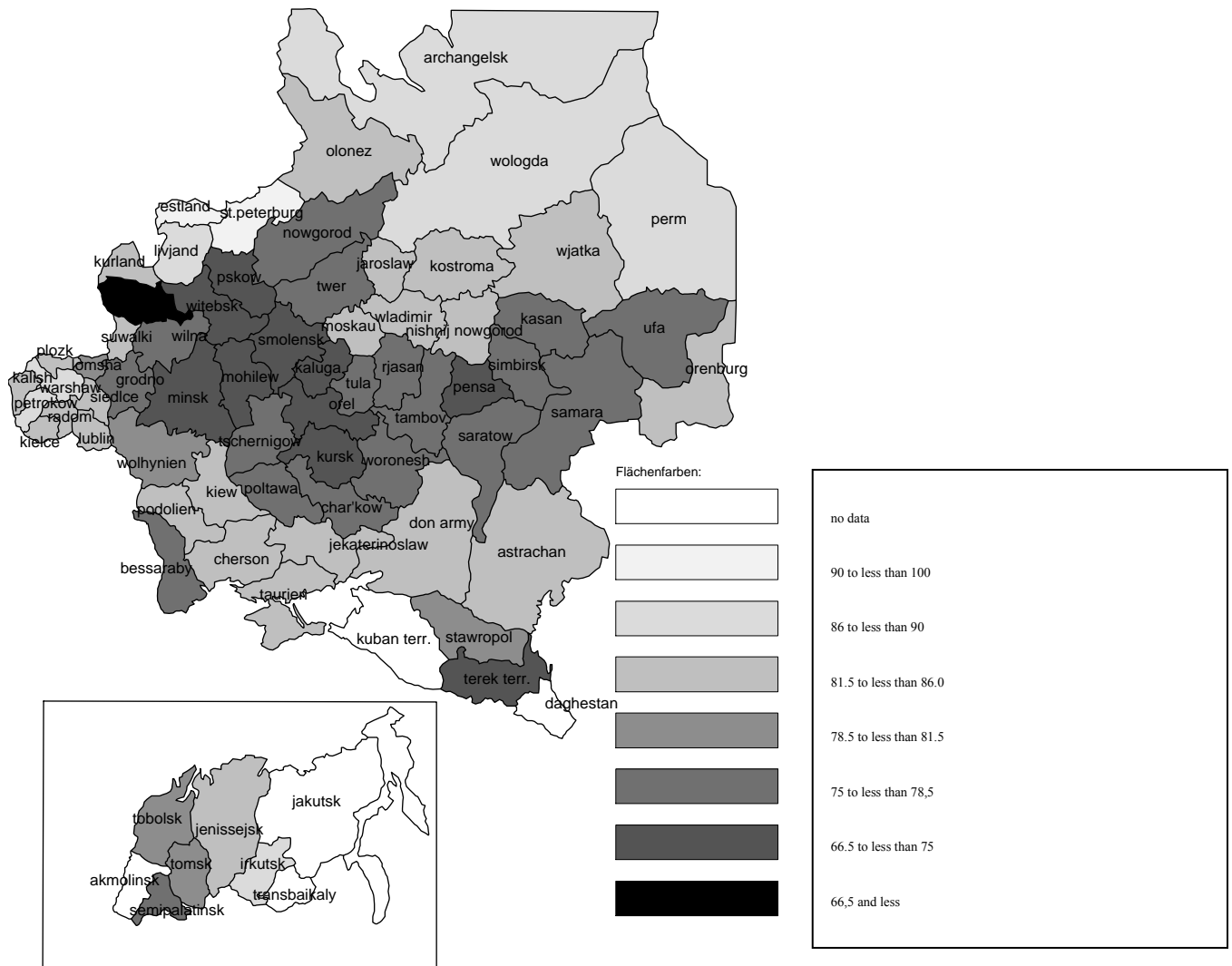
- Reis, Jaime. "Economic Growth, Human Capital Formation, and Consumption in Western Europe Before 1800." In *Living Standards in the Past*, edited by Robert C. Allen, Tommy Bengtsson, and Martin Dribe, 195–225. Oxford: Oxford University Press, 2005.
- Ruggles, S. (2012). The Future of Historical Family Demography. *Annual Review of Sociology*. Vol. 38: 423-441.
- Rutkowski, J. (1986 [1921]). Poddaństwo włościan w XVIII wieku w Polsce i niektórych innych krajach Europy. In J. Rutkowski, *Wieś europejska późnego feudalizmu*, selected and edited by J. Topolski (pp. 25-215). Warszawa: PIW.
- Sosnowska, A. (2004). *Zrozumieć zacofanie. Spory historyków o Europę Wschodnią (1947-1994)*. Warszawa: Wydawnictwo Trio.
- Stockwell, Edward G. (1966). Patterns of digit preference and avoidance in the age statistics of some recent national censuses : a test of the Turner hypothesis. *Eugenics quarterly* 13(3), 205-208.
- Stolz, Yvonne, Joerg Baten and Jaime Reis "Portuguese Living Standards 1720-1980 in European Comparison – Heights, Income and Human Capital", *Economic History Review* (2012, forthcoming)
- Szołtysek, M. (2007). Central European household and family systems, and the "Hajnal-Mitterauer" line: the parish of Bujakow (18th-19th centuries). *The History of the Family*, 1, 19-42.
- Szołtysek, M. (2008). *Three kinds of preindustrial household formation system in historical Eastern Europe: A challenge to spatial patterns of the European family. The History of the Family*, 13 (3), 223-257.
- Szołtysek, M. (2011). „A sheer jungle of uncertainties and traps”: uses and misuses of premodern census microdata. Paper presented at the Seminar of Economic History, University of Tübingen, Department of Economics (Germany), June 2011.
- Szołtysek, M. (2012a) Spatial construction of European family and household systems: promising path or blind alley? An Eastern European perspective. *Continuity and Change* 27(1), 11-52.
- Szołtysek, M. (2012b). 'The genealogy of Eastern European difference: an insider's view', *Journal of Comparative Family Studies*, 43(3), 335-371.

- Szołtysek, M. (2014). Rethinking East-central Europe: family systems and co-residence in the Polish-Lithuanian Commonwealth. Habilitation thesis submitted to the Philosophische Fakultät I der Martin-Luther-Universität Halle-Wittenberg.
- Szołtysek, M.; Gruber, S. (2014). The past, present and future of the Mosaic Project. The 10th European Social Science History Conference, Vienna, Austria, April 2014.
- Tollnek, Franziska and Joerg Baten (2011). "The Farmer's U: Which Occupational Group Inherited Human Capital in Early Modern Europe and Latin America?" Working Paper Univ. Tuebingen
- Van Zanden, Jan Luiten (1999). "Wages and the Standard of Living in Europe, 1500-1800" *European Review of Economic History*, August 1999, v. 3, iss. 2, pp. 175-97

Table 1: Places and provinces included (period before 1880/1897)

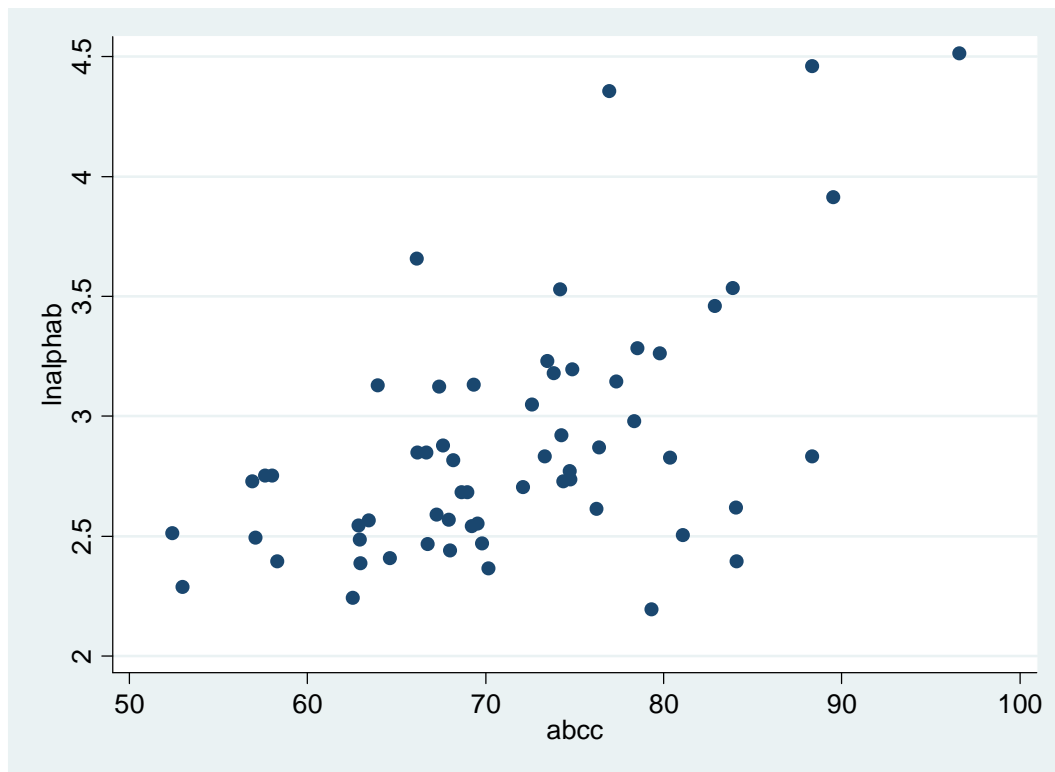
Place/County	Country today	County 1880/1897	Adm. 1880/1897	Gouv./Province	Empire
Bobrujski	by	Bobrujskij	Minskaja		Poland/Russia
Lelowski	pl	Chenstohovskij	Petrokovskaja		Poland/Russia
Charkov	ua	Charkov	Charkov		Russia
Crakow	pl	Cracow (Bezirks-Hauptmannschaft)	Galizien		Austria-Hungary
Proszowski	pl	Cracow (Bezirks-Hauptmannschaft)	Galizien		Austria-Hungary
Olsztynski	pl	Ermland-Masuren	Koenigsberg		Prussia
Mozyrski	by	Gomel'skij	Mogilevskaja		Russia
Koscian	pl	County in Greater Poland	Posen		Prussia
Kaliski	pl	Kaliskij	Kaliskaja		Poland/Russia
Kruszwicki	pl	Kaliskij	Kaliskaja		Poland/Russia
Ostrzeszowski	pl	Kaliskij	Kaliskaja		Poland/Russia
Wielunski	pl	Kaliskij	Kaliskaja		Poland/Russia
Kossow	ua	Kossow (Bezirks-Hauptmannschaft)			Austria-Hungary
Radziejowski	pl	County in Greater Poland	Posen		Prussia
Sepólno	pl	County in Greater Poland	Posen		Prussia
Wyrzysk	pl	County in Greater Poland	Posen		Prussia
Olesnicki	pl	Limanowa (Bezirks-Hauptmannschaft)	Galizien		Austria-Hungary
Leczycki	pl	Lodzinskij	Petrokovskaja		Poland/Russia
Krasnystaw	ua	Lublinskij (East)	Lublinskaja		Poland/Russia
Minski	by	Minskij	Minskaja		Poland/Russia
Nieswieski	by	Minskij	Minskaja		Poland/Russia
Nowogrodzki	by	Minskij	Minskaja		Poland/Russia
Slucki	by	Minskij	Minskaja		Poland/Russia
Wilejka	by	Minskij	Minskaja		Poland/Russia
Bytomski	pl	Opole	Opole		Prussia
Siewierski	pl	Opole	Opole		Prussia
Dawidgrodecki	by	Pinskij	Minskaja		Poland/Russia
Malborski	pl	Pomerania	Koeslin		Prussia
Przemyslany	ua	Przemyslany (Bezirks-Hauptmannschaft)			Austria-Hungary
Wilenski	lt	Vilenskij	Galizien		Poland/Russia
Kcynski	pl	Warschavskij	Vilenskaja		Poland/Russia
Eyskij	ru	Warschavskaja	Kuban territory		Russia
Moskovskij	ru	Eyskij	Moskovskij		Russia
Orenburgskij	ru	Moskovskij	Orenburgskij		Russia
Tulskij	ru	Orenburgskij	Tulskij		Russia
Vjatskij	ru	Tulskij	Vjatskij		Russia
		Vjatskij			Russia

Figure 1: Numeracy in the governments of the Russian Empire (ABCC index)



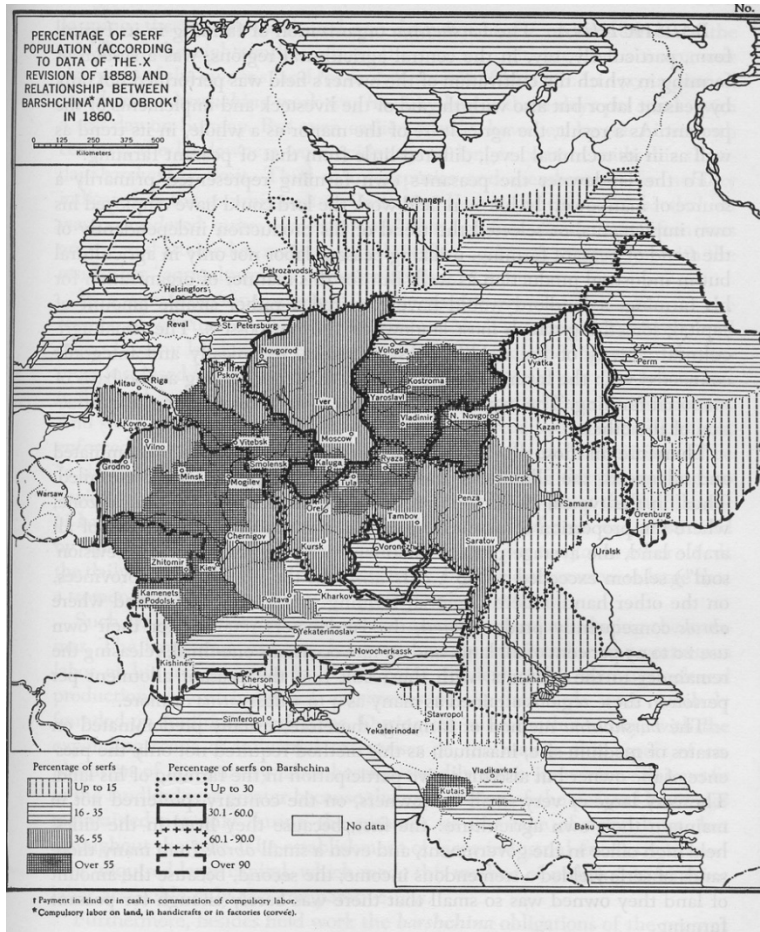
Note: It refers to individuals born between 1825 and 1874. The black value of Kovno is probably an outlier.

Figure 2: Comparison of Literacy and Numeracy in the governments of the Russian Empire (ABCC)



Note: It refers to individuals born between 1825 and 1884.

Figure 3: Serfdom in the Russian Empire



Source: Lyashchenko (1949)

Figure 4: Regionally adjusted numeracy (ABCC) of places in Russia

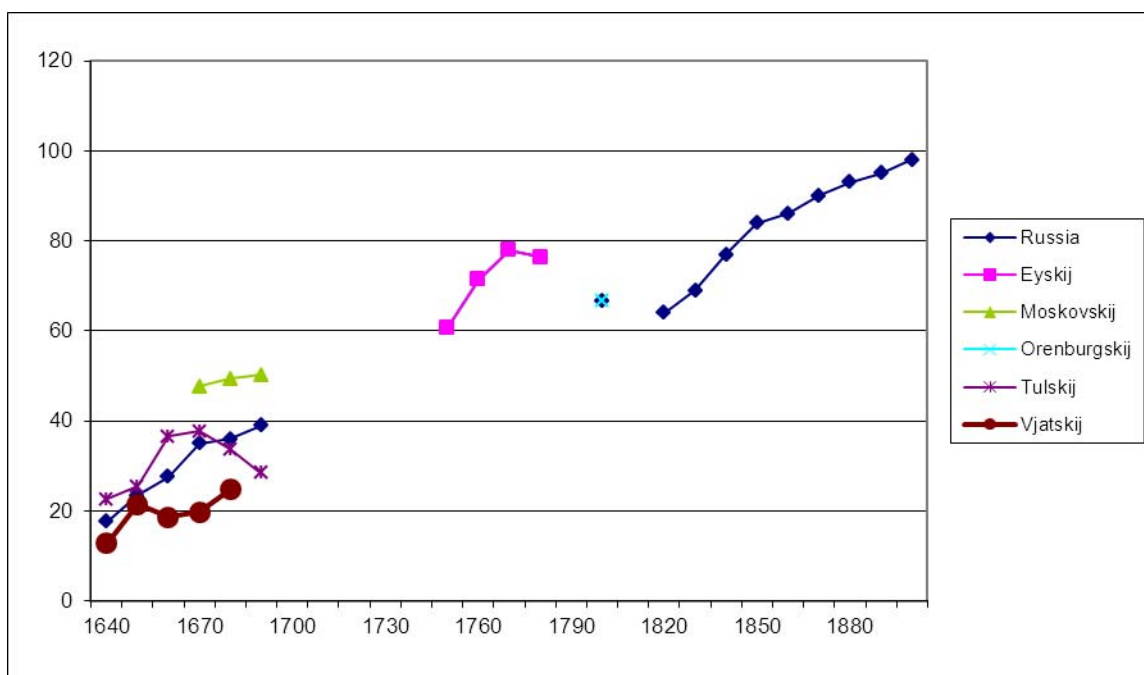
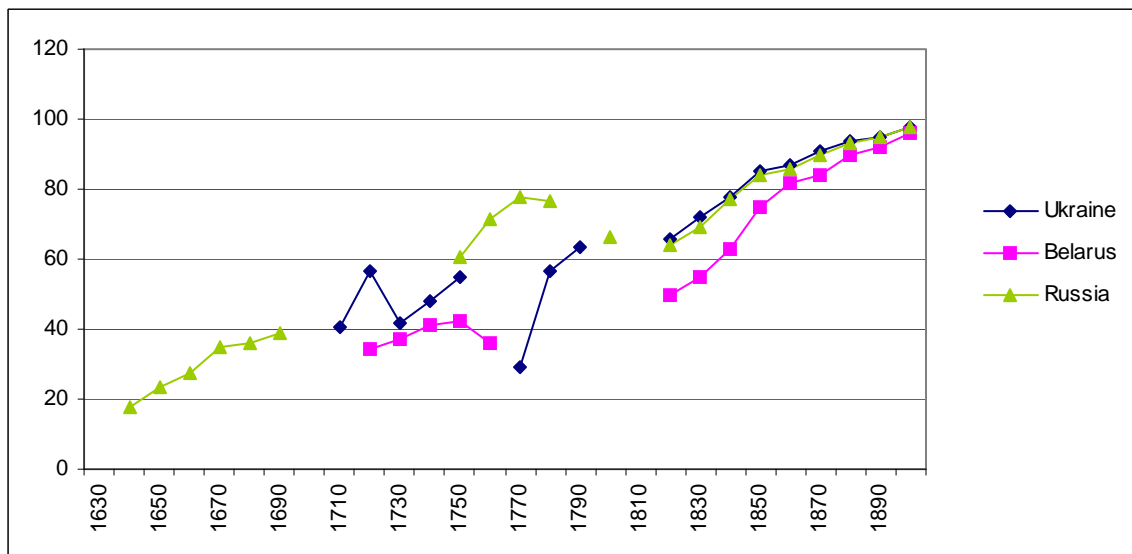
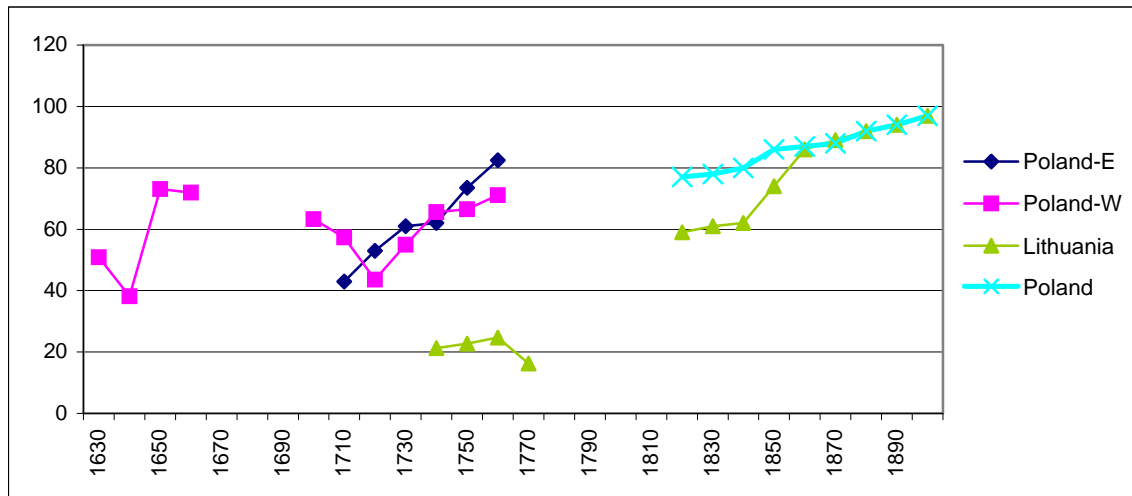
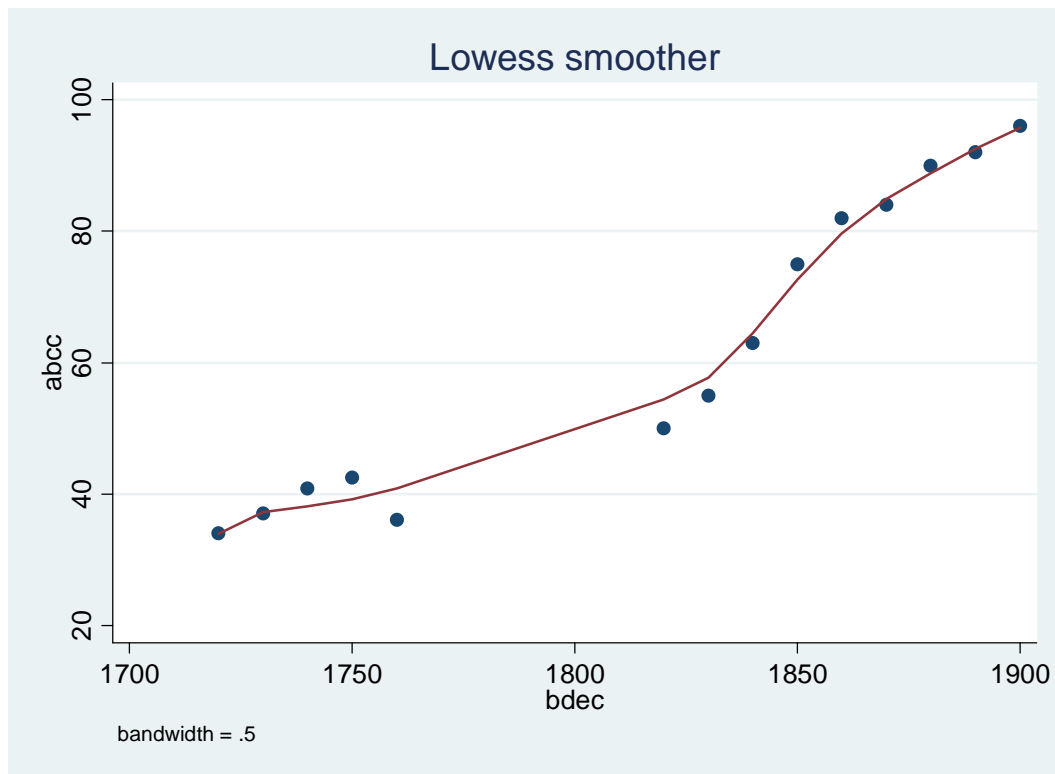


Figure 5: ABCC country trends

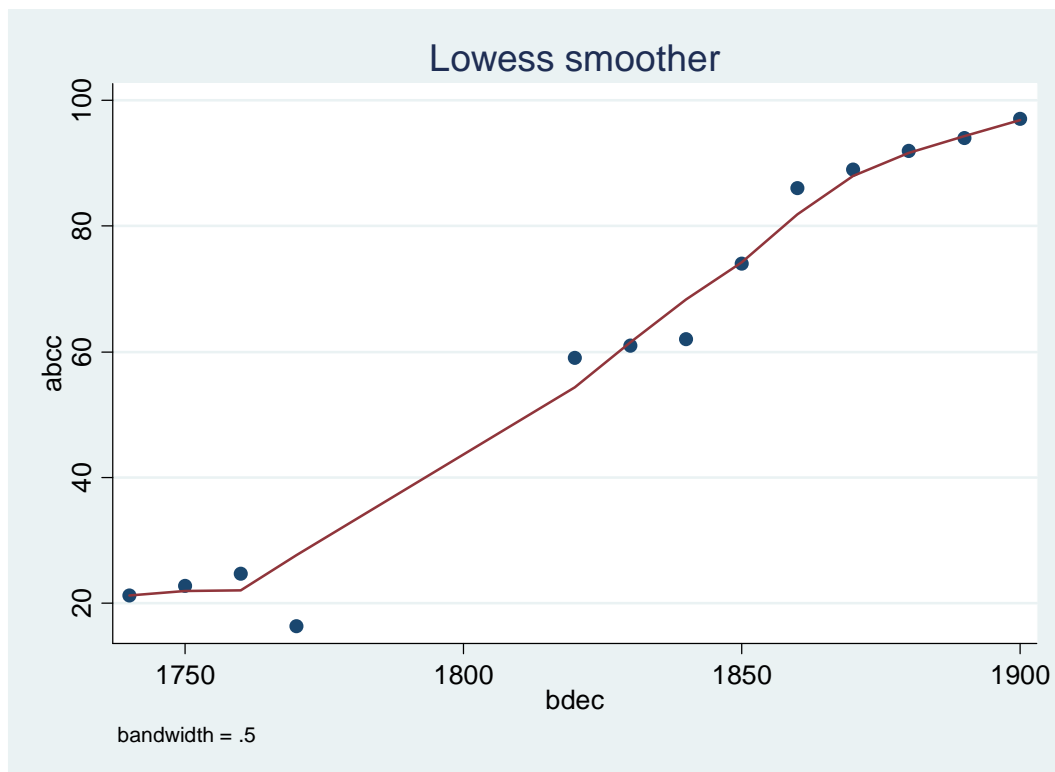


Note: "Poland" refers to the whole country (including "Poland-E" and "Poland-W.")

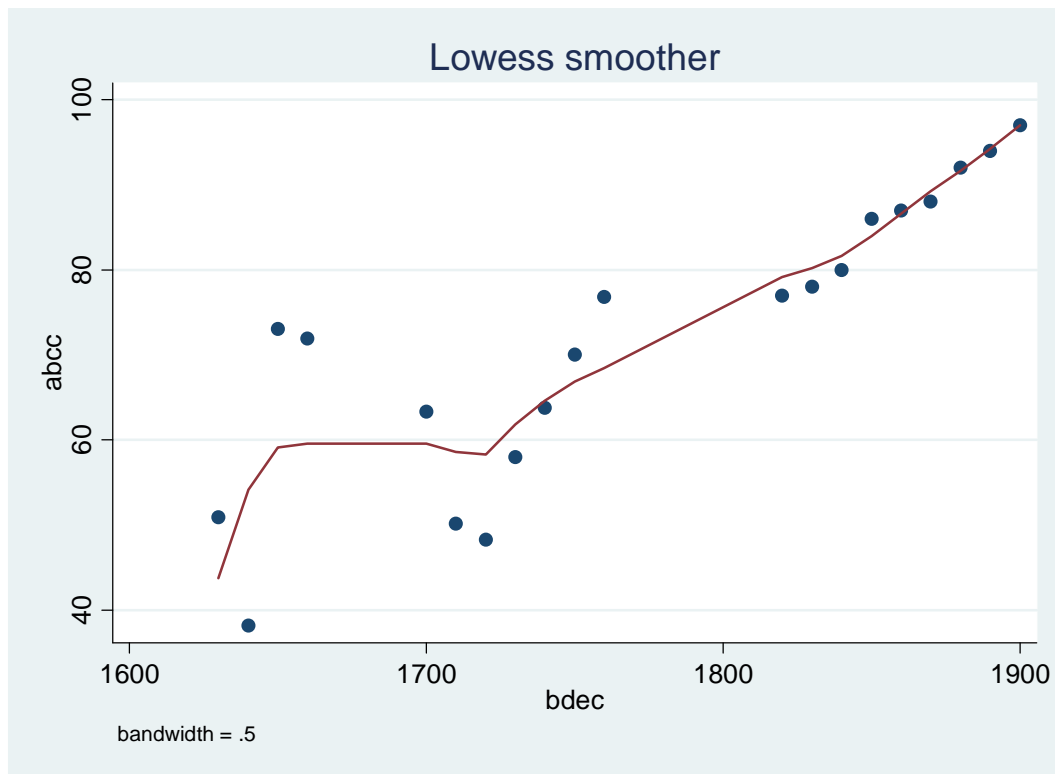
Figure 6: LOWESS-smoothed ABCC trends: Panel A -- Belarus:



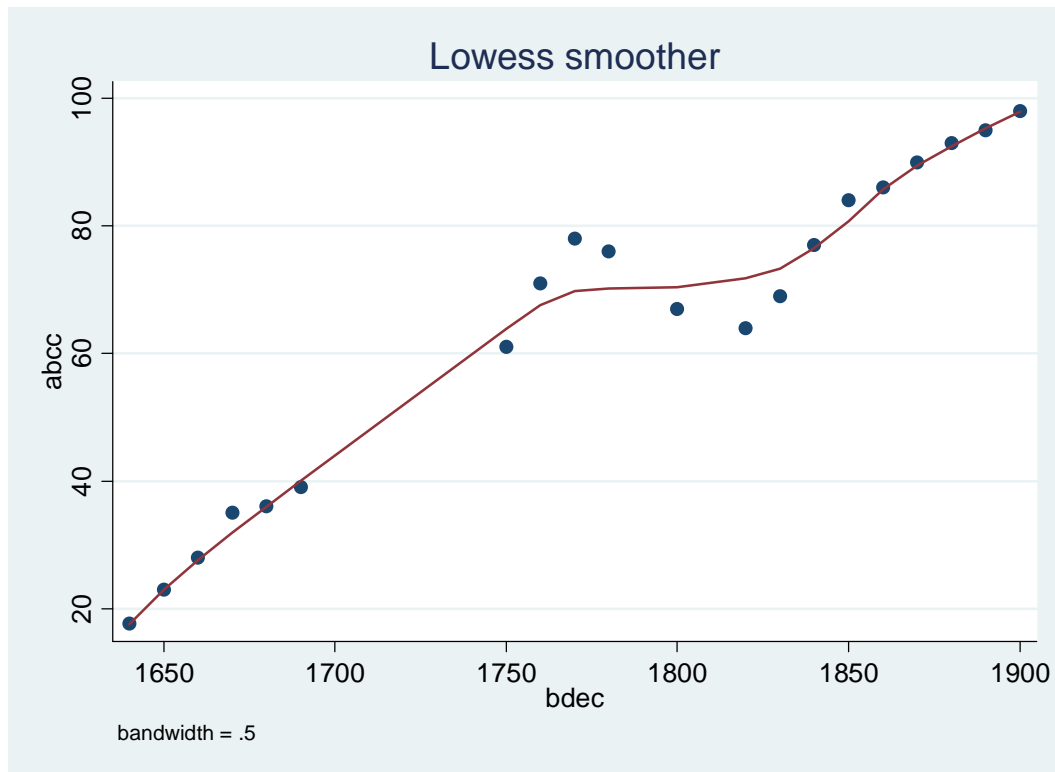
Panel B -- Lithuania:



Panel C -- Poland:



Panel D -- Russia:



Panel E -- Ukraine:

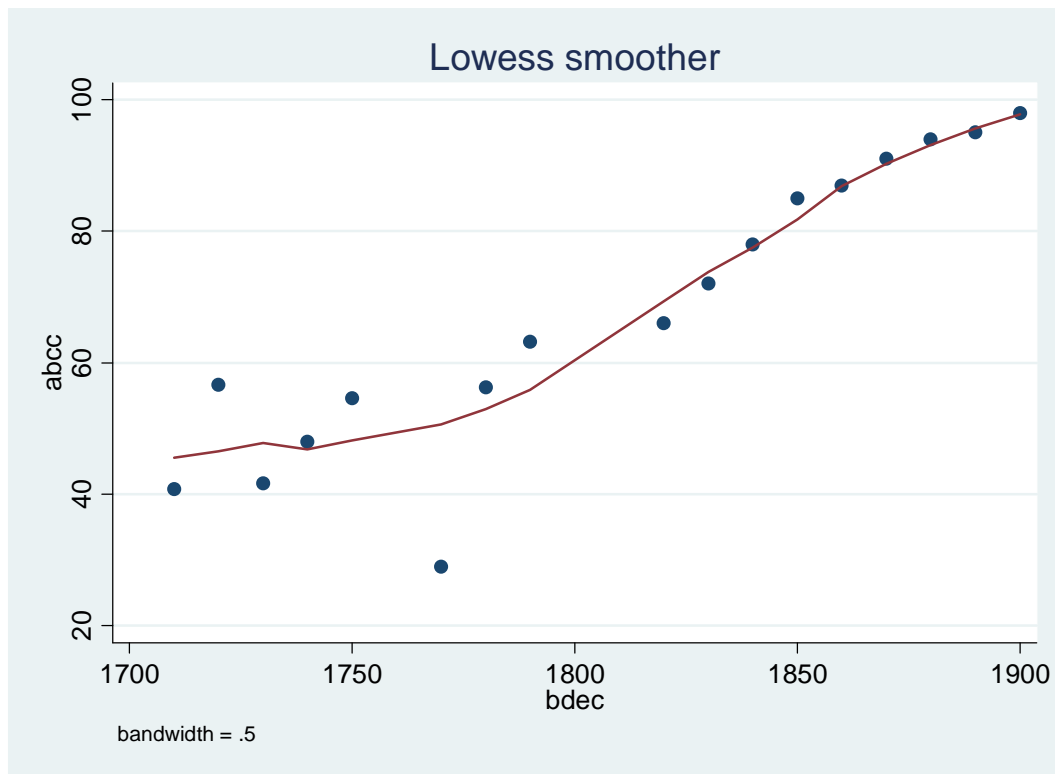
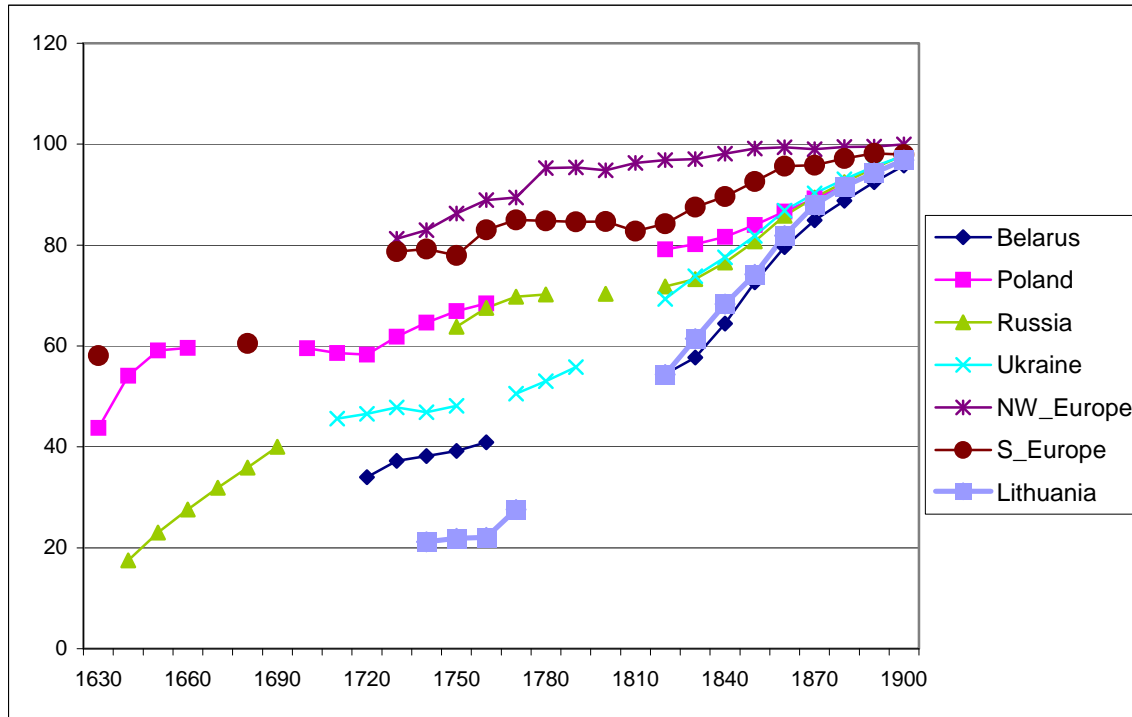


Figure 7: Eastern European ABCC estimates in comparison



Sources for NW and S Europe: Stolz et al. (2012), see also Tollnek and Baten (2011);
 | Southsouthern Europe in 1630 (1680) refers to the average value of Spain and Portugal
 during the period 1600-49 (1650-99), see Juif and Baten (2001).

Table A.1: ABCC by region (raw values), and regional adjustment factors

Region	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	Adj.factor (1820-national)
Russia																													
Eyskij													62	73	79	78				65	72	80	87	88					1
Moskovskij					66	68	69													83	84	88	93	96					19
Orenburgskij																		90		87	80	75	81	84					23
Tulskij		45	48	59	61	57	51													87	81	81	87	92					23
Vjatskij		26	35	32	33	38														78	80	85	90	90					14
Russia																				64	69	77	84	86	90	93	95	98	
Poland-East (later Russian)																													
Chenstohovskij											66	60	73	74						76	78	78	89	88					-1
Kaliskij											76	77	72	71	88					74	77	77	90	87					-3
Warschavskij									57	41	50	67								91	90	91	96	98					14
Poland-East																				77	78	80	86	87	88	92	94	97	
Poland-West (19th C Austria/Prussia)																													
Ermland-Masuren	62	49	84	83														98	88	89	90	95	91						11
Cracow_County										34	62	75	75	81					92	97	95	100	98						20
Posen										66	71	82	83	86					92	92	93	95	97						15
Oppeln												83	93	97					93	93	93	95	98						16
Pomerania								83	77	85	87	93							98	97	98	98	98						20
Poland-West (19th C Austria/Prussia)																				77	78	80	86	87	88	92	94	97	
Belarus																													
Bobrujskij										24	34	37	35	28						48	54	63	75	81					-2
Gomelskij										42	41	48	51	45						50	59	70	82	84					0
Minskij										43	34	37	40	34						55	60	66	77	81					5
Pinskij										38	50	51	54	48						57	62	68	80	80					7
Belarus										34	37	41	42	36						50	55	63	75	82	84	90	92	96	
Lithuania																													
Vilenskij											31	33	35	26						69	68	71	82	87					10
Lithuania																				59	61	62	74	86	89	92	94	97	
Ukraine																													
Charkovskij								42	44	53	60	67				84				67	72	79	87	91					1
Kossow															60	73	94		98	97	96	96	97						31
Krasnystaw										83	85	84	86							79	80	81	91	92					13
Przemyslany											28	41	51						90	92	97	93	92						26
Ukraine																				66	72	78	85	87	91	94	95	98	

Table A.2: ABCC by region (regionally adjusted values)

Region	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900		
Russia																														
Eyskij													61	71	78	76														
Moskovskij					48	49	50																							
Orenburgskij																		67												
Tulskij		23	25	37	38	34	28																							
Vjatskij		13	21	19	20	25																								
Russia		18	23	28	35	36	39						61	71	78	76		67		64	69	77	84	86	90	93	95	98		
Poland-East (later Russian)																														
Chenstohovskij											67	60	74	75																
Kaliskij										79	80	74	73	90																
Warschavskij								43	27	36	52																			
Poland-East								43	53	61	62	74	83						77	78	80	86	87	88	92	94	97			
Poland-West (19th C Austria/Prussia)																														
Ermland-Masuren	51	38	73	72																										
Cracow_County										14	42	55	55	61																
Posen										51	56	67	68	71																
Oppeln												67	77	81																
Pomerania								63	57	65	67	73																		
Poland-West	51	38	73	72				63	57	44	55	66	67	71					77	78	80	86	87	88	92	94	97			
Belarus																														
Bobrujskij										26	36	39	37	30																
Gomelskij										42	41	48	51	45																
Minskij										38	29	32	35	29																
Pinskij										31	43	44	47	41																
Belarus										34	37	41	42	36					50	55	63	75	82	84	90	92	96			
Lithuania																														
Vilenskij												21	23	25	16															
Lithuania												21	23	25	16				59	61	62	74	86	89	92	94	97			
Ukraine																														
Charkovskij								41	43	52	58	66				71														
Kossow															29	42	63													
Krasnystaw									70	72	71	73																		
Przemyslany										1	14	25																		
Ukraine								41	57	42	48	55			29	56	63			66	72	78	85	87	91	94	95	98		