Joerg Baten, Mikołaj Szołtysek and Monica Campestrini

Please note: This is not the final version of this paper. A later version has been published in the European Review of Economic History.

"Girl Power" in Eastern Europe? The Human Capital Development of Central-Eastern and Eastern Europe in the Seventeenth to Nineteenth Centuries and Its Determinants

Abstract: How did human capital develop in the Polish-Lithuanian Commonwealth and other east-central and eastern European countries? We trace the development of a specific human capital indicator during this period: numeracy. We draw upon new evidence for Poland, Belarus, Ukraine, Lithuania, and Russia, controlling for potential selectivity issues. Numeracy started at low levels, especially in Russia and, later, in Lithuania. In the mid-18th century, levels in Russia began to converge to Polish levels; later, the other regions followed. We test potential determinants such as serfdom, female autonomy, nutrition and geography. We find that female autonomy proxied by the share of young female singles had a particularly consistent positive effect. An instrumental variable regression suggests that the relationship could be causal.

Keywords: Central-Eastern Europe; Historical Demography; Eastern Europe; Human Capital; Numeracy; Age-Heaping; Census Microdata; Family Systems

JEL: N33 O40 I20

Contact: Joerg Baten, Univ. Tübingen/CEPR/CESifo. Email: joerg.baten@uni-tuebingen.de (corresponding author)

Mikołaj Szołtysek, Max Planck Institute for Social Anthropology, Halle (Saale); szoltysek@eth.mpg.de

Monica Campestrini,Univ. Tübingen, Email: Monica.Campestrini@uni-tuebingen.de

Acknowledgements: We thank Inna Vasylchuk and the ClioInfra project (NWO).

Introduction

Hanushek and Woessmann (2012) recently argued that math and science abilities are the most crucial determinants of economic growth. They employed a wide range of country panel and migration data to assess the causality of this relationship. Compared to numerical skills, their other educational indicators explained far less of the growth differences. While their research focused on the 1960–2006 period, Baten and Juif (2014) found that numeracy in the early modern period (ca. 1820) is strongly correlated with the later development of math and science skills. Some societies were able to enter a human-capital intensive path of development, and only a few of these economies later left this path. Taken together, these two studies suggest that basic numeracy during the early modern period can provide insights for long-term development. Eastern Europe is a particularly interesting case for understanding long-term development. It is geographically and culturally close to early industrial core regions, as it is situated on the same continent as Northwestern Europe.

In what follows, we trace the numeracy indicator for East-Central and Eastern Europe (ECE). “Numeracy” is defined here as the share of persons who were able to accurately state their own age (see section “Assessing...”). Our spatial focus is on the Polish-Lithuanian Commonwealth (and its successor states, i.e., Poland, Lithuania, Belarus, and Ukraine) and Russia.[1] Our sources include various types of census and census-like microdata from the territories of historical Poland-Lithuania and the Russian Empire as well as regional data for 1880 in Prussia and Austria-Hungary.

The application of age-heaping-based numeracy estimates to this newly available dataset is performed here for the first time by birth cohort (to estimate decadal trends) for a large part of early modern ECE. We carefully discuss the

---

[1] For simplicity, we use the Polish-Lithuanian Commonwealth and Poland-Lithuania interchangeably.

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 great challenge that requires special treatment in this study. To come to terms with these issues, we develop a specialized methodology. Given the regional character of our sources, we compare the places we are able to cover with representative nationwide census evidence of the early nineteenth century.[2] After measuring numeracy in several ECE regions, we will also compare the new evidence with estimates for western and southern Europe.

Measuring numeracy across regions and over time is the first major contribution of this study. Our second contribution is the assessment of the potential determinants of regional numeracy differences. In so doing, we link our analysis to the debate about family structure and economic growth. De Moor and van Zanden (2010) recently argued that differentials in female marriage ages and other components of the European Marriage Pattern (EMP)[3] might have been among the crucial causal factors of the 'Great Divergence' between Europe and China and the 'Little Divergence' between Northwestern Europe and the remainder of the continent. They also concluded that consensual marriage as promoted by Catholicism in the North Sea region and its surroundings in medieval times resulted in greater female autonomy and greater chances for skill premium and capital accumulation before marriage. They termed this effect "Girl Power". Widening the argument, several other scholars have suggested that the pattern of late marriage, high celibacy rates and nuclear household structure that was characteristic of Northwestern Europe in Early

[2] We decided to aggregate the numeracy estimates using current national borders rather than historical empires or other regional units. This will allow the estimates to be compared with other historical evidence in the future (such as GDP estimates, anthropometric welfare, and other indicators).

[3] Other components were female celibacy and neolocality.

Modern times might have been particularly conducive to economic growth. Voigtlaender and Voth (2013) argued that the West invented fertility restriction after the Black Death of the fourteenth century. Their case study of England showed that after the population loss, the country intensified its dairy farming. This production is usually associated with higher female contributions to family income and, hence, greater female autonomy, whereas the alternative grain agriculture requires more upper-body strength and, hence, more patriarchy. Girl power in England and its impact on human capital resulted in an early high-wage economy and early industrialization.

Dennison and Ogilvie (2014) recently criticized these theories and empirical analyses that argued that the EMP and girl power played a role in early growth success. They constructed a large data set of marriage ages in 39 countries and found that the highest marriage ages were not observable in England but in Scandinavia and German-speaking countries. They also questioned the exogeneity of the EMP and addressed a number of other conceptual issues.

In our study, we cannot consider all the aspects of this debate, but we contribute to one core element: the relationship between the share of unmarried women aged 20-29 years and numeracy. We will run a series of regressions to assess whether it is plausible that delayed marriage led to more labor experience of young women and, hence, higher human capital formation in the present and the following generation in both genders.

One advantage of our study for clarifying this question is its focus on ECE. If rich and poor regions are considered together, such as including the Northwestern and other regions of Europe in one sample, the results are often affected by unobservable heterogeneity issues. This can be largely avoided in our study. The less developed labor markets in ECE and EE allow to assess the exogenous role of female autonomy institutions more clearly. Endogeneity is less of an issue here than in the west and our findings are possibly more robust. Hence, important insights can be gained from Eastern European economic history. In addition, we test other potential determinants such as serfdom, political institutions, nutrition, population density, soil productivity, and other geographic variables. We will also perform an instrumental variable regression to obtain hints about the exogeneity of girl power. One source of exogeneity is relative soil quality for dairy farming, an activity in which women typically obtained a stronger position. Hence after discussing the literature and presenting estimates for the five ECE countries, we present an instrumental variable analysis of determinants of numeracy.

1. Review of potential determinants of numeracy

Almost no studies exist on numeracy trends in ECE or on their determinants. Some local evidence is presented by Mironov (1991), who looked at samples of Baltic peasants and other sources to assess the degree of age-heaping. Szołtysek (2015) considered evidence from Poland-Lithuania. In addition, Kaiser and Peyton (1993) studied the urban communities of Tula and Viatka around 1700. However, none of the four authors organized the data by birth cohorts of adults; hence, trends cannot be observed. A'Hearn et al. (2009) found that numeracy in ECE lagged behind that in the west throughout early modern times. There is slightly more evidence on the related educational indicators of literacy and schooling system, which we report in Appendix E.

An interesting debated has developed on the role of female autonomy. De Moor and van Zanden (2010) argued that in the North Sea region, women had more customary rights in the labor market and in other aspects of family economies. Late marriage might have resulted in more work experience for women, strengthening their position. Even after marriage and child-birth, when the traditional restricted role of women made them responsible for the children, the basic educational investment per child might have depended on the female human capital attained in the labor market before marriage. De Moor and Zanden described the potentially higher degree of female autonomy underlying the EMP, going back to John Hajnal. Hajnal famously argued that differences in demographic behavior existed between eastern and western Europe. In the East, marriage age was lower, nearly everyone married, the majority of households were of extended structure, and the newly married only rarely formed their households neolocally. Hajnal identified a border following a line between St. Petersburg and Trieste that might have left most of the Baltic and western Poland in the "western" part, with Ukraine, Russia, Belarus, and eastern Poland in the "eastern" part.[4] The discussion surrounding the potential familial institutions that might have mattered for growth differences focused on all components of the EMP (marriage age, neolocality, lifetime celibacy). However, we would argue that age at marriage is a particularly attractive component for indicating female autonomy (even if not without problems as discussed below) because of the microeconomic channel that runs from female autonomy institutions (reflected in the EMP) via labor experience to an increase in both women's own human capital as well as in that of their offspring. [5]

---

[4] Hajnal's observations have been rejected as too simplistic by Sklar (1970); for an even harsher criticism, see: Szoltysek 2007, 2008, 2012, 2015. We have very strong reservations about the Hajnal line.

[5]Sometimes, the hiring process itself sufficed to motivate skill development; this motivation was absent where girls married early and worked only within the family framework. Indeed, even simple occupations required more communication over simple numbers rather than just interacting with one´s

Women who did not gain as much work experience – because female autonomy institutions prevented them – also tended to provide less teaching and self-learning encouragement to their children. They were less able to teach important competences such as numeracy and other skills. They sometimes also valued these skills less because they did not "belong to their sphere".

Dennison and Ogilvie (2014) rejected the North Sea region's exclusive role in extreme forms of EMP. They found that Scandinavia and Germany showed even older ages at marriage (and similarly for other EMP components), even though industrial development was delayed by nearly a century relative to the UK. They also noted that parts of Slavic Europe (the Czech lands) showed older ages at marriage. They doubted that a relationship existed with numeracy (and other human capital) formation. Endogeneity is clearly an important issue in this part of the debate, as Dennison and Ogilvie, among others, have emphasized.

Carmichael et al. (2016) responded to the Dennison/Ogilvie critique by replacing the EMP indicators of demographic behavior with anthropological evidence from Murdock`s Atlas, assuming that the evidence on inheritance rights of women, multigenerational households, cousin marriage, etc., reported for the late 19th and early 20th century was constant over the early modern period. They argued that this might be an indicator of underlying institutions with higher "girl friendliness", whereas marriage age, for example, might also depend on poverty. If poverty was too discouraging, poor people did not marry early, but remained celibate, sometimes for their whole life. Carmichael et al. report a correlation between "girl friendliness" and

---

children and husband. Work in the family was normally less subject to competitive behavior encouraging skill formation (on labor during child ages, see Appendix F).

GDP per capita growth in the 1500-2000 period. Dennison and Ogilvie (2016) responded again by emphasizing the issue of endogeneity, leading them to conclude that non-familial institutions (property rights, market factors) had by far the stronger role in explaining growth differences.

Endogeneity is clearly a major issue in this exciting debate, assuming the well-developed labor market is the primary driving force behind better job opportunities and thus human capital growth for women. We argue here that female autonomy institutions and familial (or, gendered) labor market behaviour does play a role. We address the endogeneity question explicitly by considering gender-biased agricultural specialization. An important component is soil suitability, which in some regions favored cattle farming relative to grain agriculture (the two main components of the European agricultural sector). In cattle farming and dairy production, women often had a stronger role, whereas grain agriculture required more upper body strength and hence gave males a stronger role. This reduced female autonomy (Voigtlaender and Voth 2013, Alesina et al. 2013, Ogilvie 2003). Below, we assess whether this exogenous part of the variation could be an instrumental variable for the potentially endogenous share of unmarried singles, and find this confirmed.

In Figure 1 we clarify briefly the main differences of the debate; De Moor and van Zanden (2010) stress exogenous female autonomy institutions (reflected in EMP) acting via labor markets, and this had an impact on early development. Dennison and Ogilvie emphasize non-familial institutions as triggers (contracts etc.), which led to early labor market development, the rest being endogenous. Our model is closer to De Moor and van Zanden, as we also stress institutions of female autonomy (related to relative soil quality), although our study focuses on explaining numeracy. All three studies see a causal channel via labor markets, the main differences being what is emphasized as the exogenous trigger.

We will test the influence on numeracy exerted by the share of unmarried women aged 20-29, as data are available for more regions for this variable than for age at marriage. We assume that most of these singles marry between age 30 and 40 and have children. We calculate this share of unmarried young women for our regions, place by place, for each century. For the nineteenth century, we add the share of unmarried women for the same age bracket based on the Princeton fertility project database (http://opr.princeton.edu/Archive/pefp/).

Another main issue is that age at marriage depends not only on underlying female autonomy attitudes. As Carmichael et al. (2016) have argued, a decrease of age at marriage can also be an indicator of less poverty (as with Britain during the Industrial Revolution, for example). Hence we would need an adjustment for the latter. Unfortunately, GDP per capita, heights or even real wage estimates are not available for all regions of Eastern Europe during this early period. However, one possible way to adjust for differences and changes in poverty is to use the share of never married at age 50. In the early modern period until the late 19th century, not everybody could marry. Marriage was encouraged by the family, the village community and possibly other influential actors (feudal lords in some regions) only if a subsistence for both male and female partners and children could be expected. Voluntary celibacy played a substantial role in ECE from the late 19th century onwards only in the most urban and economically advanced regions. We show in Appendix D that the share of never married women at age 50 correlated strongly with poverty in ECE (proxied by low wages or low height values). Hence, we subtracted the share of never married women from the share of unmarried women aged 20-29 years as an adjustment. For the data set on the 17th-19th century, we can only calculate the never married women aged 40-49 years (who were not widows) as a proxy for those at age 50 years who were never married.

2. Serfdom. The second serfdom hypothesis is also commonly cited in the economic history literature as a determinant of slow development (Kula, 1976; Millward 1982; Cerman 2008; Ogilvie and Edwards 2000; also Sosnowska 2004; on a different perspective: Malinowski 2016a). Scholars have noted that historical Poland, and Russia in particular, were affected by noble landlordism and village subjection (Hagen 1998; Mironov 1996).[6] Eastern European landowners expanded their previously modest familial manor farms into large-scale domanial economies in the sixteenth century designed to produce surpluses for sale in the urban markets of western Europe.[7] This type of seigneurialism led landlords to demand from their peasant subjects not only rents in cash and kind but, above all, in labor services, which were essential to the very functioning of the demesne farms. Serfs, therefore, had relatively few incentives or opportunities to invest in the kind of basic education that would have enabled them to understand the concept of numeracy 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

---

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

7 Domar (1989) speculated about the emergence of the second serfdom, focusing on land-labor ratios as one core variable. On the trade between regions and the quasi "colonial" dependency during the early modern period, see Małowist (2010)

hereditary subjection and on their labor obligations (*corvée*) to the manors. In contrast, the region of Poland-Masuria, which became part of Prussia, had relatively mild serfdom. Quite early in history, most farmers were not serfs, and taxes and other obligations were not paid as *corvée* but rather in monetary units. Similarly, some of the Northeastern and Southeastern margins of Russia had a very low share of serfs, partly because the Russian Imperial government wanted to motivate Russian and other Western settlers to move to these regions (Figure 2). If we look at the distribution of the share of serfs in the Russian Empire during the mid-nineteenth century, a clear regional pattern emerges.[8] In a central corridor between Belarus (Minsk) and Pensa, the share of serfs was particularly large. By contrast, there were relatively few serfs in the thinly populated northeast region or that between the Black and Caspian Seas. During the late serfdom period, the southeast not only had a lower share of serfs but also 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 also literacy). The region of lowest numeracy was between Minsk and Pensa, whereas the Northeast and Southeast had better numeracy values (Figure 3).

Different perspectives on serfdom have recently emerged. In contrast to many studies that stress the negative impact of serfdom in each of its aspects, Nafziger (2010) argues that the institutions of land communes, which already developed under serfdom, were more efficient than previously thought. In a similar vein, Dennison and Nafziger (2013) argue that in spite of serfdom, the standard of living increased in Russia during the 19th century. The authors used micro level evidence on central

---

[8] Note, however, that no data are provided for Polish territories.

Russia and found that income and consumption, as well as human capital, improved even before the abolition of serfdom. Dennison (2011) discusses many institutional and contract-enforcement aspects of serfdom. Recently, Malinowski and van Zanden (2015) argued that income inequality was even lower in Poland under the serfdom system than in the Netherlands in the same period. We actually find below that serfdom variation was less significant for numeracy than the girl power factor. Was this perhaps caused by collinearity effects with female autonomy? We find that this was not the case, as the correlation coefficient is as low as -0.05 (p-value 0.62).

One of the measures used in the literature for approximating serfdom is the size of the manorial estate (see the literature cited above). A second related channel could hence be that large landowners prevented tax-financed public schooling, as they saw no need for serfs to be educated. The influence of large landowners in all ECE governments until the nineteenth century also kept national educational policies at comparatively low levels.

In our regressions below, we include dummy variables to identify the regions that had a “milder” serfdom (less than 30 percent serfs) (Pipes 1995). More exact measurements of serfdom do not exist for the period under study.

3. Nutrition. Recently, a number of studies have focused on the effect of childhood malnutrition on cognitive ability. If a person suffered from a severe nutritional deficiency— as many human beings did in the past— numerical abilities were difficult to develop and to preserve later in life (see Baten et al. 2015 and Appendix G for a review of the evidence). One of the bottleneck components of nutrition during the early modern period was animal protein, as substantial areas of land are required to produce sufficient meat and milk. For early modern ECE we do not have evidence of cattle per capita— the traditional indicator— for each region.

We can, however, include this variable using larger regions that represent later countries, as provided by the Clio-Infra project, even if we need to interpret this variable with considerable caution (clustering standard errors on the country/half-century level removes econometric problems).

4. Low population density. Sparse population and the lack of a transport system made commuting to schools costlier. 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. Moving to the eastern areas, we observe a gradual decrease in population density; late eighteenth-century Belarus had densities well below 10 persons/km² (Szołtysek 2015). Higher population density was also typically related to higher urbanization rates. We introduce the measurement of persons per square kilometer based on the ClioInfra database, which only presents large aggregated regions (modern countries) by decade.

5. 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 western part of the region. The western fringes of Poland were inhabited by religiously mixed communities. The degree of religious fractionalizationwas much lower in the Polish East and Russian "borderlands" throughout the early modern period (Szady 2010, 228-250). Indeed, the competition in the western regions between the Unites and the Orthodox believers on the one hand and between the Catholics and Protestants on the other may have been of some importance (Mitterauer 2003). We will capture this factor using dummy variables for mostly orthodox or mostly protestant, although we are well aware of the heterogeneity of the religious groups and regions.

6. Other factors: We tested, among other factors, policy and institution factors, but found other factors either less influential or not adequately measurable (Appendix K).

2. Sources

Our sources are as follows: (1) the “lists of souls” (*Libri Status Animarum, Seelenregister*); (2) the censuses of the Civil-Military Order Commissions 1790–1792 in Poland-Lithuania (which were later occupied by the Russian Empire); (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 of 1897 in Russia (regionally aggregated); and (5) other types of household lists, including “communion books” and local administrative surveys as well as private and Crown estate inventories.[9]

In Appendix A, Table A.1, we report on the places and regions for which evidence is available. Notably, of the places situated in Prussia or Austria-Hungary, we only included 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 have later fled or been moved to Germany after WWII. Regarding the ethnic overlap among Poland, Lithuania, and Belarus, we have been less restrictive. There might be some migration biases later on. We also took care to

[9] All these sources, except for nineteenth century censuses and the seventeenth/eighteenth century sources on Russia, are the part of the *CEURFAMFORM Database* developed by M. Szołtysek, 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, 2006–2008. Details: Szołtysek 2008, 2014.

not 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 Appendix A, Table A.2).

3. Assessing human capital formation with the age-heaping indicator

In both industrial and agricultural economies, numeracy was clearly a core component of human capital. In agricultural societies, individuals 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 (Baten 2016).

Measuring the production factor "human capital" has never been simple, as advanced forms of skills are difficult to compare. Economists have therefore resorted to using proxy indicators, such as the share of people signing a marriage register. 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, which is based on the tendency of poorly educated people to round their age erroneously. For example, less-educated people are more likely than people with greater levels of human capital to state their age as "30," even if they are in fact 29 or 31 years old.

The calculation of the ABCC Index of numeracy is shown here as a derivation of the Whipple Index (Wh):[10]

[10] "ABCC" comes from the authors' names A'Hearn, Baten and Crayen (2009), plus Greg Clarks, who inspired this simple linear transformation of the Whipple index with a comment. The underlying Whipple Index is the only one that fulfills the desired properties of scale independence and that ranks samples with different degrees of heaping reliably.

(1) $$Wh = \left(\frac{(Age25 + Age30 + Age35 + ... + Age60)}{1/5 * (Age23 + Age24 + Age25 + ... + Age62)} \right) \times 100$$

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

The correlation of numeracy and literacy illustrates that the numeracy indicator is informative. Crayen and Baten (2008) found that the relationship between illiteracy and age-heaping for less developed countries after 1950 is very close. They calculated age-heaping and illiteracy rates for no less than 270,000 individuals who were organized into 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 modern student test 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. This correlation was reported in many studies (Appendix B has more examples). Crayen and Baten (2010b) also examined a variety of other potential determinants of age-heaping, such as the degree of bureaucracy (proxied by the number of censuses performed for each individual country up to the period under study), birth registration practices, and government interaction with citizens that potentially influenced the likelihood that an individual would know his or her exact age, independent of personal education. Crayen and Baten found that bureaucracy indicators were mostly 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 where birth registers were introduced early on had a high degree of age awareness. However, those societies also introduced early schooling, and this was the variable that clearly had more explanatory power than the independent bureaucracy effect. They 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.

For the nineteenth century, we have evidence of regional differences in numeracy and literacy in the Russian Empire (Figure 3 and A.1). We observe a strong correlation between numeracy and literacy of 0.6025 (p-value = 0.000, Appendix L and Appendix A, Figure A.1).

We should clarify that the time costs incurred by those parents who let their children acquire basic numeracy are not extremely high. Some amount of explanation of numbers when playing children's games requiring number processing is often enough. Still, the costs were and are too high for many families. Families sometimes sent their children to work instead if it was necessary for family survival, with severe consequences for the children (Appendix F).

In conclusion, the correlation between age-heaping and other human capital indicators is well established, and the "bureaucratic" factor does not invalidate this relationship. One 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. However, excluding ages younger than 23 is an easy remedy for this. It also shows that most individuals knew their age as teenagers, but only those in well-educated societies were able to remember or calculate their exact age later in life. We will also exclude those above 72 because a number of distortions could affect groups composed of elderly individuals.

4. Potential selectivity issues and the adjustment of regional biases

Whenever a dataset is used that might not reflect the whole population of a given country or region, the following question concerning potential selectivity arises: 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 seventeenth or eighteenth centuries, but we should at least consider whether the dataset has systematic biases that distort the results in a significant way, and we should seek to minimize potential measurement error.

Social bias or labor market bias is not an important issue for our sample, given the census-type character of our sample. In principle, everyone in a given city or village should have been included in the dataset.[11] 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.3 in the appendix. Panel A of Table A.3 shows the new regional estimates for the period from the 1630s to the 1790s; Panel B shows estimates for the period from the 1800s to the 1900s. The latter set of estimates are based on the 1880 and 1897 population censuses as well as later censuses (except 1800s/1810s), while the former set of figures is 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

[11] Some of the Polish-Lithuanian registers excluded the clergy and nobility. This exclusion is not as severe because studies on ECE and Southern European countries have indicated that the numeracy of the nobility was usually not upwardly biased relative to the other population; see Baten and Sirbiladze (2015). In addition, in early modern times in the two constitutive parts of Poland-Lithuania, the nobility amounted to only approximately 8 percent of the total population in the Polish Crown and 6.5 percent in the Grand Duchy of Lithuania (see Kuklo 2009, 221).

representative? The fact that Moscow is included in the five documented regions suggests that there is probably upward bias. Hence, the following questions arise: Did the regions have ABCC values similar to the average values found for Russia in the 1897 census? How large was the upward bias? The analysis shows that four of the five regions had ABCC values that are 15-25 percent above the Russian average (Column “Adjustment factor”). For simplicity, we take the difference for the 1820s birth cohort and report the regionally adjusted values in Table A.4 in the appendix. For example, if Moscow had an 83 percent numeracy rate, which was 19 percent above the national average level of 64 in 1820, we adjusted the earlier values for Moscow downward by this amount. For the 1690s, 69 minus 19 equals 50, which is our estimate of the national Russian value based on the Moscow evidence (Table A.4). The estimates of all Russian places for one birth decade are then averaged. 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 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 Przemyshlany and the 1730s birth decade, the resulting value is clearly too low; moreover, the low value for the Warsaw region might have been either a measurement error or the result of the destruction of the city and the eighteenth century plague (Kuklo 1991). In the vast majority of cases, however, the regional adjustment procedure works relatively well. To remain consistent, we take all the values into account.

To assess the robustness of our regional adjustment method, we also applied a different assumption. Rather than using constant differences in 1820 from a national mean, we scaled the difference by the overall level of numeracy. The advantage of this second procedure is that it takes into account the national trends of the countries and makes the adjustment sensitive to possibly different trends (See Appendix C). The results of this second approach were actually very similar.

5. Estimates for the five ECE countries and international comparison

In the next step, we generate national estimates based on the regional values. In Appendix A, Figure A.2, we display the regional and national estimates for Russia. Some of the early estimates are above and others below the estimate for Russia, but the emerging trend seems to be relatively clear. Hence, we show national trends for all five countries in Appendix A, Figure A.3. We distinguish among (a) western and (b) eastern parts of today's Poland and (c) Poland-Masuria, which had different characteristics. The western part comprises Silesia and other parts of Prussia, as well as districts that were annexed by Prussia and Austria-Hungary in the eighteenth-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 east would yield similar estimates for the whole of Poland after being regionally adjusted to the national mean. In fact, the similarity found for all levels 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 with the LOWESS procedure (definition: notes to Figure A.4 in Appendix A).

ECE 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 4, data from ECE were plotted against the evidence from western and southern European countries. Stolz et al. (2012) 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 approximately an 80 percent numeracy rate in the early eighteenth century, but the northwestern region made more rapid progress and achieved a 95 percent numeracy rate around the year 1800. Numeracy in southern Europe stagnated at a quite high level of approximately 82 percent from the 1730s until the 1820s and then slowly converged with northwestern European levels.

Hence, the northwestern and southern European regions were clearly more numerate than all the ECE regions we are assessing here during the eighteenth and nineteenth centuries, although Poland-Masuria did not differ much from the European south during the seventeenth century. Moreover, the trends of convergence and the slowdown in the individual regions are interesting. Russia started at a low level (approximately 20 percent) in the early seventeenth century. The gap between Russia and Poland was less than five percent in the mid-eighteenth century. During the later nineteenth century, human capital once more accumulated quite rapidly, and the problem of basic numeracy was nearly solved by the year 1900, approximately.

Poland displayed mostly stagnant levels of numeracy during the early eighteenth centuries (approximately 60 percent),[12] but it grew in numeracy during the middle decades of the eighteenth century. Among the countries studied here, Belarus, Lithuania, and Ukraine lagged behind the most. During the early to mid-eighteenth century, numeracy still stood at approximately 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 reaching Russian levels during the nineteenth century.

We should note at this point that numeracy and literacy are not linearly correlated (i.e., a 20 percent numeracy increase does not equal a 20 percent literacy increase). Basic numeracy is usually achieved earlier. Hence, A'Hearn, Baten and Crayen (2009) estimated that an 80 percent numeracy rate corresponds to ca. 60 percent literacy, and a 60 percent numeracy rate corresponds to an approximately 20 percent literacy rate. Basic numeracy estimates are measures particularly suitable for tracing early stages of development because they have a great degree of variation in early modern Europe (and in nineteenth/early twentieth century developing countries). For rich countries in the twentieth century, other measures (such as tests of math skills) are more suitable. These more advanced measures would be useless in early stages of educational development because they are close to zero and have no variation that can be used for analysis.

---

[12] Kochanowicz (1991) describes how the conservative ideals of the Polish elites (mostly its nobility) prevented the development of markets and industrial production. Even the few industrial enterprises that were founded did not aim at making profit but at replacing imported goods with self-produced goods priced according to own cost. This could have reinforced the modest growth during the early 18th century.

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

6. Analysis: potential determinants of numeracy

Descriptive and regression results (seventeenth to nineteenth century ECE)

Based on the following econometric model equation, we regress numeracy by region and decade on the variables and factors discussed above.

$$N_{i,l,d,c} = \alpha + \beta_1 GP_{i,c} + \beta_2 S_{i,c} + X'\gamma + \mu_l + \delta_c + \varepsilon_{i,d}$$

Where $N_{i,l,d,c}$ stands for numeracy (in region i within land l, decade d of century c), $GP_{i,c}$ (the share of unmarried women aged 20-29) is our proxy for female autonomy, $S_{i,c}$ is our proxy for serfdom in region i and century c, μ_l and δ_c are country and century-fixed effects, $X'\gamma$ is the vector of additional explanatory variables that were described in the section above, α is a constant and ε is the error term. Including both indexes 'i' and 'l' (the region and later country, respectively) is necessary because some variables are not available for regions. Similarly, the index d stands for the decade, and c is the century (or half century for some variables) to which decade d belongs to. As noted above, as not all variables can be measured with regional and decadal resolution, clustered standard errors are needed to avoid econometric problems.

In a second step, we will also regress numeracy by Russian province based on similar variables. We will avoid speaking of a "causal influence" exerted by girl power or other factors on numeracy, as the direction of causality could run in the other direction (or both ways). Even if we cannot make strong claims about the direction of causality at this point, identifying common developmental factors of "girl power" and numeracy is already a considerable step forward in the analysis of early modern development in ECE. In a final step, we provide IV regression estimates that give some indication of causal relationships.

One question that we need to clarify is whether the share of unmarried women aged 20-29 could also be mechanically correlated with age-heaping because heaping on age 30 will affect the share of women being assigned as married or not married for the 20-29 age-group. It turned out that we can remove this potential source of endogeneity. We used only older age groups, aged 33 to 72.[13] This reduced N, but the results were remarkably robust compared to when ages 23 to 72 are included.

Descriptives of the variables for the seventeenth–nineteenth century panel are presented in Table 1. We were able to obtain values for all variables for 62 cases. The average numeracy rate is 66 percent, with a range between 28 and 88.

We calculate the share of single women aged 20 to 29 years at the century and region level, the average being 13 percent in early modern ECE. In some districts, the value was close to zero, such as in Bobrujski and Pinskij in Belarus, Lublinskij in Poland, and Gomel'skij in Russia. It was relatively high in the city of Warsaw, in Poland-Masuria, in Pomerania and, interestingly, in Kossow in Ukraine (later part of the Austrian-Hungarian Empire) and Tula in central Russia.

Twenty-three percent of our sample was characterized by milder serfdom (Poland-Masuria as well as the northeastern and southeastern margins of the Russian Empire). Cattle per capita has an average of 15, representing approximately one animal for every sixth inhabitant. Fifty-five percent of our sample came from regions that were at least partly orthodox (modern-day Russia, Belarus, Ukraine).

[13]As a robustness test, we also checked the standard grouping (age 23-72) and found out that the results are almost identical.

Table 2 reports the results of our panel regression. The variable girl power always shows a strong positive correlation with numeracy in all regression models. We also estimate robust standard errors. We also run a robust regression estimation in specification (4). We include time-fixed effects in most of the models except in specification (2). Even if we include country-fixed effects, the variable "Girl Power" is positively related to numeracy, which is a relatively strong indication that there is a relationship. Additionally, the coefficients are relatively similar across models. If we multiply one standard deviation of model 4, which has a coefficient of 5.19, with a standard deviation of girl power (in logs), which is 0.77, we arrive at a value of 4.00. With the coefficient of 9.48 from model (2), we arrive at a standard deviation effect for girl power of 7.30. These are economically significant coefficients. One additional standard deviation of girl power could have closed approximately 20-37 percent of the numeracy gap between Russia and western Europe (the gap was approximately 20 percent in the eighteenth century; later it was less).[14] It also represents a notable fraction of the standard deviation of numeracy (which was 17 percent).

When it comes to the "milder serfdom" dummy variable, the effect is consistently positive, although it is not statistically significant. Protein proximity has mixed results, either insignificant and negative (with country FE) or significantly and positively related to numeracy (no country FE). This will be further assessed below (section IV regression). Interestingly, Orthodox religion has a positive effect even after controlling for other variables such as protein proximity, girl power and serfdom. Hence, orthodoxy by itself does not seem to be negatively related if a number of other variables are controlled for. The adjusted R-square is relatively high.

---

[14] A one standard deviation effect of the explanatory variable is often taken as an indication of economic significance (not identical to statistical significance).

Finally, we are interested in the question of whether the girl power variable is statistically significant perhaps only because of some individual outliers. In the upper panel of Figure 5, we see the scattergram for region and decade units between residual numeracy and residual girl power. The positive regression slope indicates that there was in fact a positive effect and that it was not caused by outliers. In the lower panel, we show the same for the aggregation of half centuries (not decades as above) and added some district names as well. For example, the Ukrainian Kossow region in 1750 had a relatively high residual girl power (the share of unmarried women was relatively high), whereas Minskij in 1750 and rural Krakau in 1700 had low levels of unmarried women and very low numeracy.

Regression results for the Russian Empire in the nineteenth century

We also tested whether there was a girl power effect among the provinces of the Russian Empire (=Gubernias) in the nineteenth century as a robustness test (Tables 3 and 4). We used the numeracy rate of the early nineteenth century (birth decade 1820) and regressed this rate on the share of unmarried women aged 20-29 years using, once more, the poverty adjustment described above. We also ran a regression without adjusting for poverty (i.e. not subtracting those who never married until age 50), and the results were virtually identical (Models 4 and 7). The ABCC level was, on average, 63 percent for this decade and ranged from a little more than 43 percent up to 96 percent in the more educated provinces of nineteenth century Russia (Table 3; numeracy refers to both genders). The share of unmarried women was 18 percent, and, once more, there is considerable variation. We found 31 percent of the provinces characterized by milder serfdom (30 percent and less were serfs) in the northeast and southeast of the Russian Empire. The majority of provinces were predominantly

Orthodox; the exceptions were the Baltic provinces and some in the west of the empire. We added a cereal suitability variable to obtain a rough proxy of soil productivity. Index values for altitude and ruggedness were also included.

The results of the regression are shown in Table 4. Regressing numeracy in the 52 provinces, we find a significantly positive effect of girl power. The variable is statistically insignificant only in model (5), in which we focus only on the western part and in which some collinearity with the Protestantism variable exists (the latter causing the statistical insignificance); still, the coefficient there is of a meaningful size.

Taking the effect of a one standard deviation change in this variable, which is 0.64 (the log specification), we again find a substantial contribution of girl power, representing 27 percent of the standard deviation of numeracy. The “milder serfdom” variable always has a statistically significant and economically substantial influence of between 8 and 14 percent for the Russian provinces of the nineteenth century, depending on which other variables we include. Orthodoxy again has a positive effect here, and even more so Protestantism (relative to Catholicism, which represents the constant). Cereal suitability does not have a positive effect, perhaps because the population density was higher in regions with more fertile soil. We also included some geographic variables such as altitude and ruggedness, but they were not statistically significant. Overall, we can explain quite a substantial part of the numeracy variations using these regressions. As a robustness check, we also performed regressions of literacy and obtained results of similar sign (see Appendix H).

Instrumental variable estimation

To assess potential endogeneity in our regressions, we perform a Two-Stage Least Squares (2SLS) regression. We use an instrumental variable that may allow for a consistent estimate of the relationship between female autonomy and numeracy, even if endogeneity would be present. We base the construction of our instrument on the hypothesis of Alesina et al. (2013) and of Voigtlaender and Voth (2013), who agree that agricultural specialization and gender-specific labor demand influences gender roles. Alesina et al. (2011) argue that in areas where plough cultivation was widespread, women had a relative disadvantage with respect to men because this cultivation requires more upper body strength. The disadvantage of plough cultivation for women also increases due to its low compatibility with other activities, such as childcare. In Europe, the alternative to grain oriented agriculture (using ploughs) was cattle farming, which was typically associated with a more active role for women. Voigtländer and Voth (2013) similarly suggest that the relative prevalence of animal husbandry over grain cultivation might be an important determinant of differences in female autonomy. Animal husbandry benefits the relative bargaining position of women in their society because in this activity, upper body strength is of smaller relevance. Skills that were transferred from mother to daughter (disease prevention, hygienic behavior) were also a comparative advantage. As an instrument, we take the ratio between the pasture suitability of a region (=good for animal husbandry) over its cereal suitability. The first stage regression indicates that our instrumental variable is highly correlated with girl power (Table 5, F-stat much higher than 10). In the second stage, the relevance of girl power is confirmed.

However, the exclusion restriction is always an issue in IV estimation. For example, we could imagine that the instrument might affect numeracy not only via the girl power channel but also via nutritional benefits from cattle farming. To assess this potential issue, we include in our regressions two additional explanatory variables that can capture the direct effects of nutritional standards on numeracy. We include a height variable, as height is a proxy for nutritional quality. Even better, height is also employed in the literature as a proxy for general levels of welfare (Baten et al. 2015). Even correlates of the IV with omitted variables related to welfare are controlled for. By including height, we can also cope with this potential problem, increasing the robustness of our 2SLS regression results (the coefficient for height itself is not statistically significant, but we have controlled for contemporary nutrition effects.) As a second variable for direct nutrition effects, we include cattle per capita. Again, this does not render the second stage estimates of the girl power coefficient insignificant. While nobody can be perfectly sure that the exclusion restriction does not cause problems, we can at least substantially reduce the likelihood of a problem with the inclusion of these two variables.

It should be emphasized that the relative soil quality instruments long-run attitudes and institutions of female autonomy that were formed over centuries. In contrast, height and cattle per capita are short-run, nutrition-related variables that react to population growth, for example, which does not impact immediately on female autonomy institutions.

The results of this section suggest that there is an exogenous component of the girl power factor. Our interpretation is that relative soil suitability influenced underlying gender institutions in the long run, based on the relative bargaining position of women. In the short run, nutrition in districts with good pasture suitability was not necessarily higher, as the growing population might have lowered its nutritional quality in spite of soil suitability.

Conclusion

This study is the first to trace numeracy trends for ECE with a large sample, going back to the seventeenth century for some countries. We provide an account of initial underdevelopment, especially during the seventeenth and early eighteenth century, although regional differences were substantial. Some regions with high female autonomy ("girl power") had values comparable to Southern Europe. In general, ECE had low numeracy values during the seventeenth and eighteenth centuries and did not develop as rapidly as other European regions. This result is quite compatible with recent real wage studies on the region (Malinowski 2016b).

Although evidence on potential explanatory variables has to be used with strong caveats (for example, cattle per capita and other variables are only available for large regional units and by half century), we provide econometric evidence of a correlation between "girl power" and numeracy. An instrumental variable regression suggests that the relationship might actually be causal. Milder forms of serfdom also had a consistently positive sign and large coefficient, although not statistically significant. In contrast, geography, religion, population density and other variables might have mattered less. The fact that female autonomy and labor experience mattered even for ECE is a remarkable result that places historical demography and gender-specific histories at the core of European development (cf. Gruber and Szołtysek 2015 on Patriarchy Index).

References

ACEMOGLU, D., S. JOHNSON and J. A. ROBINSON (2002). "Reversal of fortune: geography and institutions in the making of the modern world income distribution", *Quarterly Journal of Economics* 117:4, 1231-79.

A'HEARN, B., J. BATEN AND D. CRAYEN (2009). "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital", *Journal of Economic History* **69**,783-808.

ALESINA, A., GIULIANO P. and NUNN N. (2013). "On the Origins of Gender Roles: Women and the Plough", *Quarterly Journal of Economics* 128:2, 469-530.

BATEN, JOERG (ed.) (2016). *A History of the Global Economy: 1500 to the Present,* Cambridge University Press

BATEN, J., D. CRAYEN AND J. VOTH AND (2015) "Numeracy and the Impact of High Food Prices in Industrializing Britain, 1780-1850", *Review of Economics and Statistics* 96:3, 418-430.

BATEN, J. AND D. Juif (2014) "A Story of Large Land-Owners and Math Skills: Inequality and Human Capital Formation in Long-Run Development, 1820-2000" *Journal of Comparative Economics* 42-2, 375-401.

BATEN, J. AND T. SIRBILADZE (2014). "War der Adel eine Elitegruppe? Die 'Numeracy' verschiedener Sozialgruppen im Georgien des 19. Jahrhunderts", *Vierteljahrsschrift fuer Sozial- und Wirtschaftsgeschichte* 101-2, 125-138.

BATEN, J. AND J. L. VAN ZANDEN (2008). "Book Production and the Onset of Early Modern Growth", *Journal of Economic Growth* **13-3**, 217-235.

CARMICHAEL, S. DE PLEIJT, A. VAN ZANDEN, JL and DE MOOR, T. (2016): "The European Marriage Pattern and Its Measurement", *Journal of Economic History* 76:1, *196-204.*

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_M.

CRAYEN, D. AND J. BATEN (2010b). "Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth", *Explorations in Economic History* **47-1**, *82-99.*

CRAYEN, D. AND J. BATEN (2008). “Global Trends in Numeracy 1820-1949”, *CESifo WP 2218*. [=Extended WP version of 2010b]

DE MOOR, T. AND J. L. 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, 1-33.

DENNISON, T. (2011) *The Institutional Framework of Russian Serfdom*. Cambridge: Cambridge University Press.

DENNISON, T AND OGILVIE, S. (2014). *Does the European Marriage Pattern Explain Economic Growth? JEH 74-3, 651-693.*

DENNISON, T AND OGILVIE, S. (2016). *Institutions, Demography and Economic Growth? JEH 76-1, 205-217.*

DOMAR, EVSEY D. *Capitalism, socialism, and serfdom: essays*. New York: Cambridge UP, 1989.

GRUBER, S., & SZOŁTYSEK, M. (2015). The patriarchy index: a comparative study of power relations across historical Europe. *The History of the Family*, 1-42.

HAGEN, W.W. (1998). “Village life in East-Elbian Germany and Poland, 1400-1800: subjection, self-defense, survival”, T. Scott (ed.), *The peasantries of Europe from the fourteenth to the eighteenth centuries*. London: Longman, 145-190

HANUSHEK, E.A., AND WOESSMANN, L. (2012). “Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation”, *Journal of Economic Growth*, 17(4): 267-321.

KAISER, D. AND PEYTON, E. (1993). “Time- and Age-Awareness in Early Modern Russia.” Comparative Studies in Society and History 35:4, 824–39.

KIRK, D. (1946). *Europe’s Population in the Interwar Years.* Princeton: Princeton University Press

KOCHANOWICZ, J. (1991), “Backwardness and Modernization: Poland and Eastern Europe in the 16th – 20th Centuries” Ashgate, Aldershot: 933-942.

KUKLO, C. Rodzina w osiemnastowiecznej *Warszawie*, Białystok 1991.

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.

MALINOWSKI, M. (2016a). Serfs and the City: Market Conditions, Surplus Extraction Institutions, and Urban Growth in Early Modern Poland, *European Review of Economic History*, 20-2, 123-146.

MALINOWSKI, M. (2016b). Little Divergence Revisited: Polish Weighted Real Wages in a European Perspective, 1500-1800, *European Review of Economic History*, 20, 1-23.

MALINOWSKI, M. AND J.L. VAN ZANDEN (2015). "National income and its distribution in preindustrial Poland in a global perspective." EHES Working Papers in Economic History No. 76.

MAŁOWIST, MARIAN (2010). Western Europe, Eastern Europe and World Development, 13th-18th Centuries: Collection of Essays of Marian Małowist 16. Brill, 2010.

MILLWARD, R. (1982). "An economic analysis of the organization of serfdom in Eastern Europe." *Journal of Economic History*, 42, 513-48.

MIRONOV, B. N. (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.N. (1996). "When and why was the Russian peasantry emancipated?" M.L. Bush (Ed.), *Serfdom and slavery: studies in legal bondage*, Harlow : Longman, 324-347.

MITTERAUER, MICHAEL (2003):*Warum Europa? Mittelalterliche Grundlagen eines Sonderwegs*. München.

NAFZIGER S. & DENNISON, T. (2013). "Micro Perspectives on Russian Living Standards", 1750-1917. *The Journal of Interdisciplinary History* 42.3 : 397-441.

NAFZIGER S. (2010) "Peasant Communes and Factor Markets in Late Nineteenth Century Russia", *Explorations in Economic History* 47.4 : 381-402.

OGILVIE, S. (2003) *A bitter living: women, markets, and social capital in early-modern Germany*. Oxford UP.

OGILVIE, S., J. S. S. EDWARDS (2000). "Women and the Second Serfdom: Evidence from Early Modern Bohemia", *Journal of Economic History* 60, 961-994.

PIPES, R. (1995). *Russia under the old regime*. Penguin: London.

SOSNOWSKA, A. (2004). *Zrozumieć zacofanie. Spory historyków o Europę Wschodnią (1947-1994)*. Warszawa: Wydawnictwo Trio.

STOLZ, Y., BATEN J. AND REIS J. (2013) "Portuguese Living Standards 1720-1980 in European Comparison", *Economic History Review* 66-2, pp. 545-578.

SZOŁTYSEK, M. (2007). "Central European household and family systems, and the

"Hajnal-Mitterauer" line: the parish of Bujakow (18th-nineteenth centuries)”, *History of the Family* 1, 19-42.

SZOŁTYSEK, M. (2008). “Threekinds of preindustrialhousehold formation systeminhistorical Eastern Europe: Achallenge to spatial patterns of the European family”, *The History of the Family, 13 (3), 223-257.*

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. (2015). “Rethinking East-central Europe: family systems and co-residence in the Polish-Lithuanian Commonwealth.” Peter Lang: Bern et al.

VOIGTLAENDER, N. AND J.VOTH (2013) “How the West ‘Invented’ Fertility Restriction.” *American Economic Review* 103:6: 2227–64.

Table 1: Descriptives of our decadal panel sample of regions, birth decades 1630s-1820s

Variable	Obs.	Mean	Std. Dev.	Min	Max
Numeracy	62	65.52	17.14	27.62	88.25
Girl Power	62	0.13	0.08	0.02	0.36
ln(Girl Power)	62	-2.25	0.77	-4.11	-1.03
Milder Serfdom	62	0.23	0.42	0.00	1.00
Cattle per capita	62	15.49	13.66	4.57	33.64
ln(Cattle p. c.)	62	2.32	0.91	1.52	3.52
Religion (orth.)	62	0.55	0.50	0.00	1.00
Pop. Density	62	12.48	8.60	1.26	27.04
ln(Pop. Density)	62	2.13	1.04	0.23	3.30

Notes: The 62 cases refer to the combination of variables in Table 2, Column 1 for which we have all explanatories.For numeracy (both genders, aged 3 - 72) and girl power (i.e., the share of single women aged 20-29, subtracting single women aged 40-49, adjusted for poverty) sources see text. “Milder serfdom” was assigned if Pipes (1995, p. 145) classified the region as having less than 30 percent proprietary serfs per population on the eve of emancipation, or if the region was Poland-Masuria, for which other literature indicates mild serfdom. Apart from this region, in our data set mild serfdom characterized Viatka in NE Russia, Charkov in modern-day Ukraine, and Eyskij (Gouv. Kuban territories) in South Russia. We defined all regions as (partly) orthodox except the Baltic and Poland. Cattle per capita is expressed as 100 animals per person for expository purposes (Source: Clio-Infra (2015), this refers to modern country borders). Population density: Inhabitants per square kilometer per half century (Source: Clio-Infra (2015), this also refers to modern country borders).Altitude is defined as elevation above sea level (in meters). The discussion of the latter 3 variables can be found in Hippe and Baten (2015). ‘Keep them ignorant.’ Did Inequality in Land Distribution Delay Regional Human Capital formation? Working Paper Univ. Tuebingen..

Table 2: Least-Squares Dummy Variable (LSDV) Regressions of numeracy, using our decadal sample, birth decades 1630s-1820s

Model	(1)	(2)	(3)	(4)	(5)	(6)
Estimation	OLS	OLS	OLS	Robust	No C-FE	OLS
Girl Power	5.19***	9.48***	4.48**	5.19**	7.23***	4.62**
	(0.005)	(0.000)	(0.016)	(0.027)	(0.003)	(0.019)
Milder Serfdom	5.38	5.58		4.80	8.40	5.18
	(0.161)	(0.275)		(0.254)	(0.137)	(0.181)
Protein proximity			-9.62		25.21***	
(log cattle p.c.)			(0.420)		(0.001)	
Pop. Dens.						11.38
						(0.269)
Religion (orth.)					27.03**	
					(0.033)	
Time FE	YES	NO	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	NO	YES
N	62	62	62	62	62	62
Rsq	0.62	0.51	0.61	0.64	0.54	0.62

Note: Robust p-values are indicated in brackets. Dependent variable is numeracy (age 33 – 72). In addition, we did a robust regression (less sensitive for outliers) in model (4). The robust specification in column 4 gives less weight to outlying observations. Time FE are dummies referring to centuries, country FE are dummies for (later) countries. We use clustering by region and half-century to cope with any possible aggregation and serial correlation problems.

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Descriptives of our Russian Empire evidence: cross-section of provinces during the nineteenth century

Variable	Obs	Mean	Std. Dev.	Min	Max
Numeracy	49	63.26	11.46	43.24	95.62
Girl power	49	0.18	0.12	0.05	0.47
ln(Girl power)	49	-1.91	0.64	-3.09	-0.75
Milder Serfdom	49	0.31	0.47	0.00	1.00
Ln(Cattle p. c.)	49	-3.04	0.34	-4.30	-1.09
Religion (orth.)	49	0.90	0.31	0.00	1.00
Cereal suitability	49	62.73	19.97	10.60	94.90
Altitude	49	1.50	0.60	-0.13	3.22
Ruggedness	49	0.46	0.29	0.14	1.89

Notes: The 49 cases refer to model 4 of table 4. For sources on numeracy and girl power (i.e., the share of single women aged 20-29, subtracting the share of never married women at aged 50), see text. "Milder serfdom" was assigned if Pipes (1995, p. 145) classified the region as having less than 30 percent proprietary serfs per population on the eve of emancipation. We defined all regions as (partly) orthodox except the Baltic and Poland.

Population density: Inhabitants per square kilometer per half century (Source: Hippe and Baten (2015), who refer to population aged 23-72).

The suitability of land for a crop (in this case cereals) is estimated by "comparing likely attainable yields with the maximum biological yield for that crop in ideal environmental conditions. Land where attainable yields are very close to the maximum potential yield is classified as very suitable for that crop, whereas land where attainable yields are far below the potential maximum is classified as only marginally suitable or not suitable" (Van Velthuizen et al. 2007, p. 2). The values cannot be directly interpreted as numbers, given their index nature. Data on altitude (median) and ruggedness (standard deviation of altitude) are ESRI grid raster data with a resolution of 30 arc-seconds provided by Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and A. Jarvis (2005). Very high resolution interpolated climate surfaces for global land areas, *International Journal of Climatology*, 25: 1965-1978. Regional data have been derived from this dataset.

Table 4: Regression of numeracy: cross-section of provinces of the Russian Empire (nineteenth century)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Poverty adj.	Yes	Yes	Yes	No	Yes	Yes	No	Yes
Region omitted	None	None	None	None	Russia	None	None	None
Estimation techn.	OLS	Robust	OLS	OLS	OLS	OLS	OLS	OLS
Girl power	4.82*	5.85***	4.67*	6.22**	4.90**	3.37	6.79**	5.41**
	(0.053)	(0.005)	(0.064)	(0.018)	(0.032)	(0.482)	(0.021)	(0.047)
Milder Serfdom	11.55**	8.05***	14.08***	14.48***	10.10***	8.39	10.11***	9.89**
	(0.014)	(0.010)	(0.000)	(0.000)	(0.006)	(0.214)	(0.002)	(0.015)
Cereal suit.	11.90*				-0.29***	0.10	-0.26**	-0.34**
	(0.085)				(0.004)	(0.518)	(0.013)	(0.025)
Protein Proximity	-0.17*	-0.24***	9.82*	9.29*	5.81**	11.07***	4.63*	5.73**
	(0.065)	(0.001)	(0.067)	(0.078)	(0.022)	(0.008)	(0.062)	(0.030)
Pop. Density					1.52			2.07
					(0.352)			(0.259)
Protestant						30.14**		
						(0.036)		
Orthodox						8.52*		
						(0.057)		
Altitude							1.52	1.56
							(0.648)	(0.643)
Ruggedness							0.58	-1.85
							(0.914)	(0.765)
Constant	114.81***	104.77***	96.85***	96.29***	101.91***	79.82***	98.83***	102.53***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Observations	52	52	53	53	49	20	49	48
Adjusted R-squared	0.42	0.40	0.36	0.40	0.49	0.64	0.51	0.45

Note: Robust p-values are indicated in brackets. The standard errors are clustered by (later) country.
The dependent variable refers to the 1820s birth cohort (age group 63-72).
*** p<0.01, ** p<0.05, * p<0.1

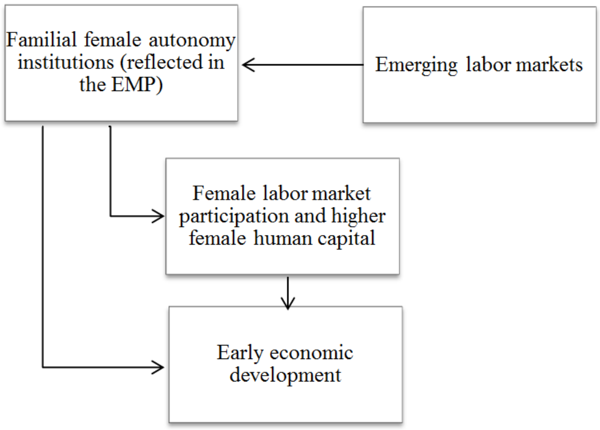
Table 5: Two-Stage-Least Square Regressions of numeracy: cross-section of provinces of the Russian Empire (nineteenth century)

	(1)	(2)	(3)	(4)	(5)	(6)
Single 20-29 adjustment	Yes	Yes	Yes	Yes	Yes	No
Girl power	9.68**	14.63**	19.34***	6.14*	8.70**	9.54**
	(0.010)	(0.017)	(0.001)	(0.073)	(0.020)	(0.013)
Milder serfd.	13.40***		16.37***	10.54***	14.32***	14.37***
	(0.000)		(0.000)	(0.001)	(0.000)	(0.000)
Protein proximity	4.52*		5.55***	5.09**		
	(0.072)		(0.009)	(0.047)		
Altitude				-3.26		
				(0.220)		
Ruggedness				8.26*		
				(0.067)		
Pop. Dens.				-1.77		
				(0.141)		
Height					-0.95	-0.71
					(0.449)	(0.545)
Constant	91.23***	90.88***	103.66***	92.57***	228.42	188.70
	(0.000)	(0.000)	(0.000)	(0.000)	(0.262)	(0.320)
Observations	48	48	18	48	46	46
Second stage adj. R-squared	0.27	n.a.*	0.44	0.43	0.29	0.35
First stage						
Relative pasture suitability	1.37***	1.39***	1.14***	1.40***	1.35***	9.17***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-statistic	119.98	139.81	72.28	98.88	107.69	15.81

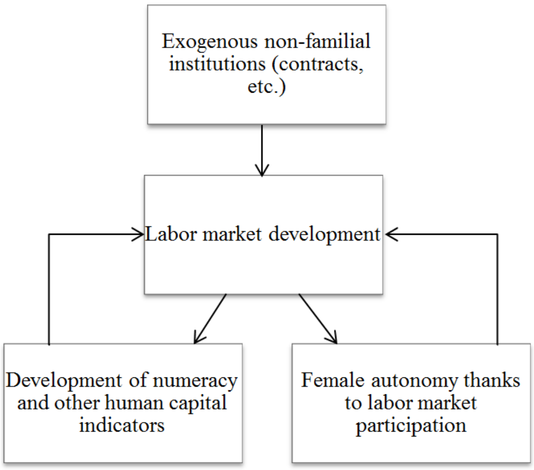
Note: * Adj. R-square cannot be computed with only no additional exogenous variables.
Instrument "Relative pasture suitability" refers to the ratio between the pasture suitability of a region over its cereal suitability. Robust p-values are indicated in brackets. The standard errors are clustered by country and half century. The dependent variable refers to the 1820s birth cohort (age group 63-72).
*** p<0.01, ** p<0.05, * p<0.1
Pasture and cereal suitability data, altitude, ruggedness provided by FAO and IIASA (2007). *Suitability of global land area for rainfed production of cereals (intermediate level of inputs) (FGGD)*, online, last accessed 5 December 2012, dataset downloadable at http://www.fao.org:80/geonetwork/srv/en/resources.get?id=14077&fname=cereal_int.zip&access=private, see also documentation at http://www.fao.org/geonetwork/srv/en/metadata.show?id=14077). For more details, see Van Velthuizen, V., Huddelston, B., Fischer, G., Salvatore, M., Ataman, E., Nachtergaele, F., et al. (2007). *Mapping biophysical factors that influence agricultural production and rural vulnerability*, Rome: FAO.

Figure 1: A comparison of theoretical models on “Girl Power”, human capital and economic development.

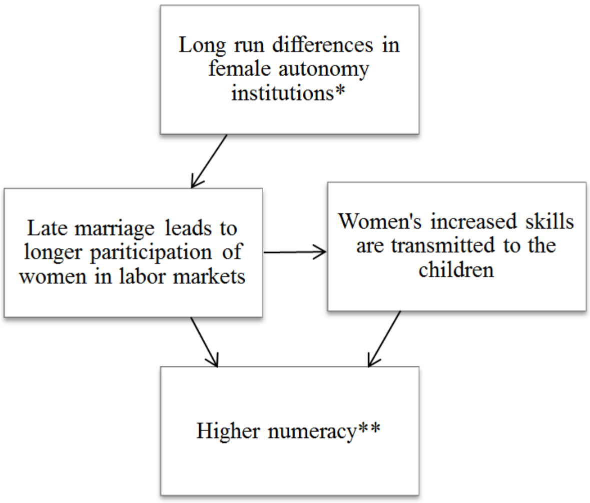
De Moor and van Zanden (2010)



Dennison and Ogilvie (2014)



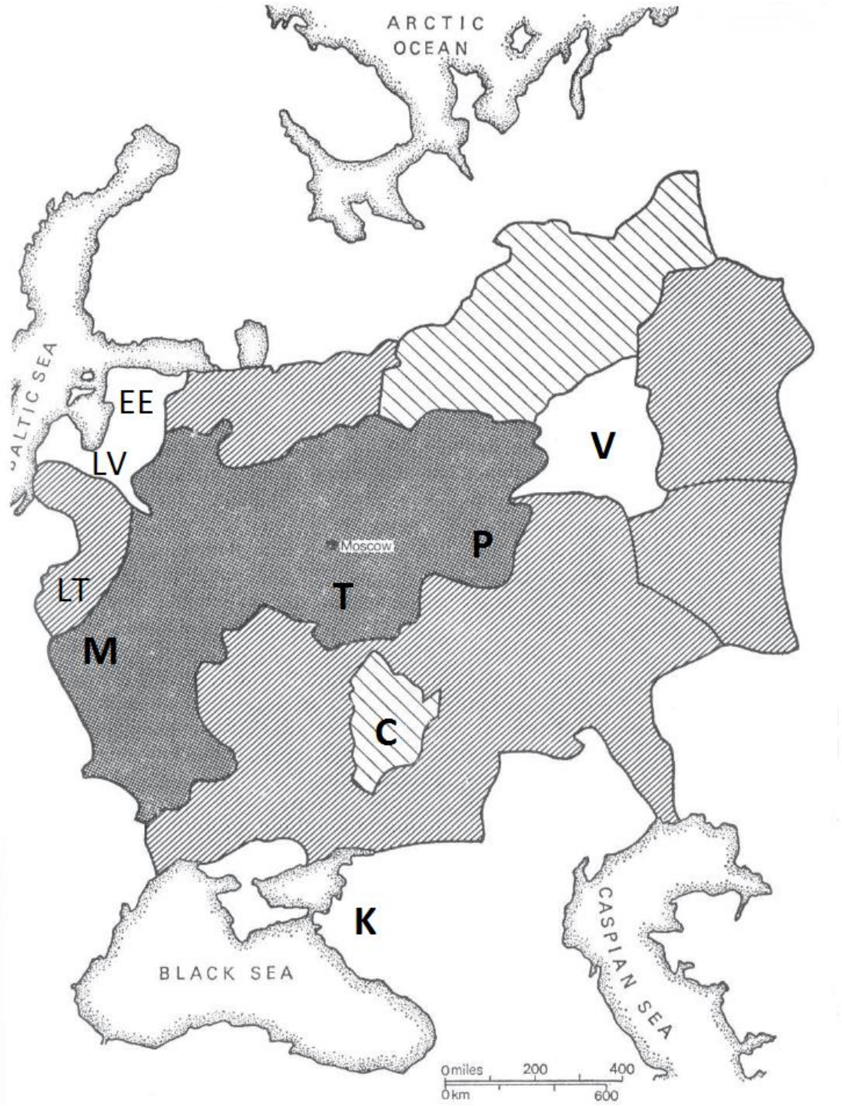
Baten, Szołtysek, Campestrini (2016)



* Caused by geographic differences (among other factors): relative suitability for pasture (dairy farming) relative to grain agriculture is a proxy for one important component (see text).

** Higher numeracy influences long run development (and not necessarily early development). For example, the late development of Scandinavia and Germany might be rooted in early numeracy development. But this is not the scope of this article (see Baten 2016 for an overview).

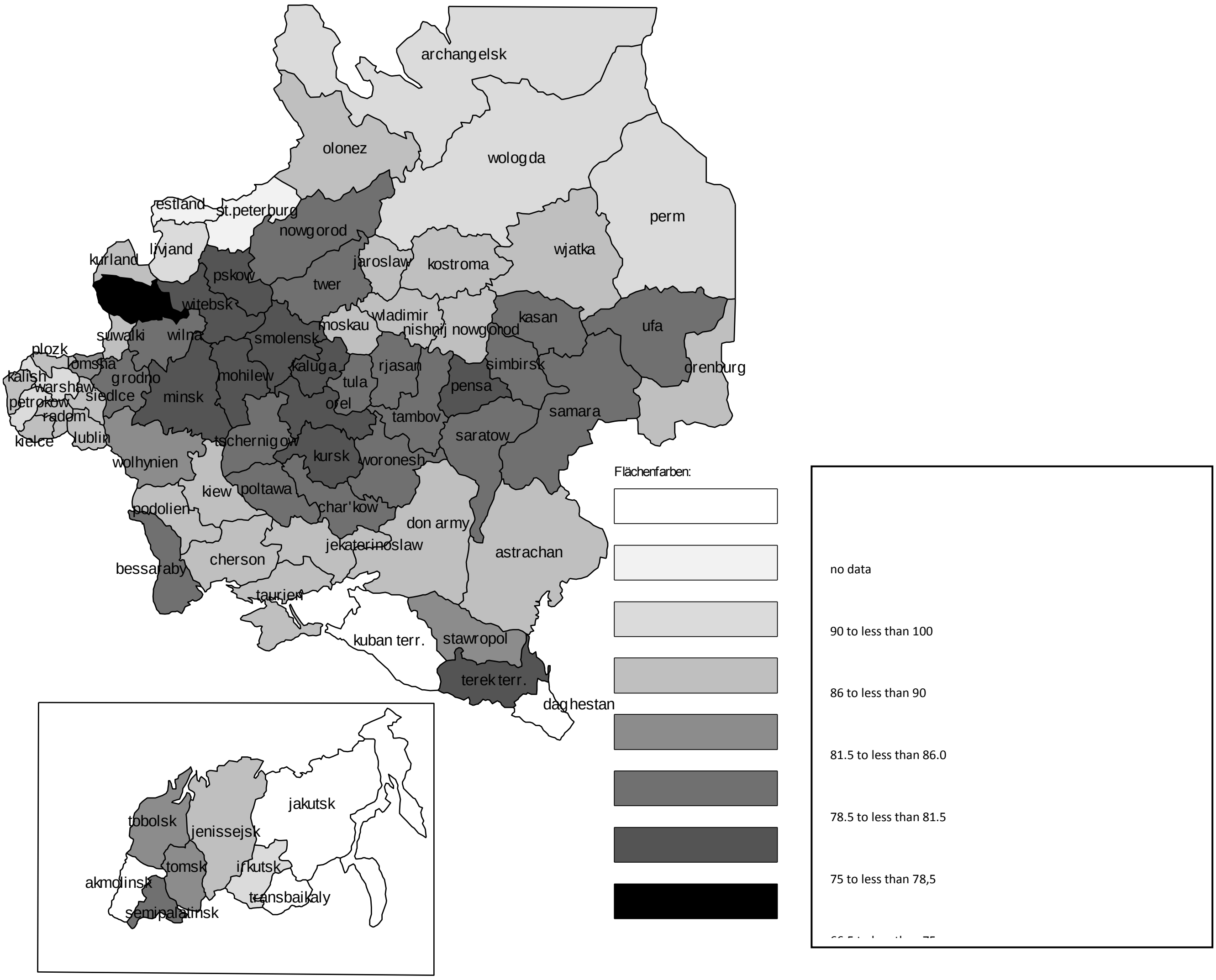
Figure 2: Serfdom in the Russian Empire



Source: Pipes (1995)

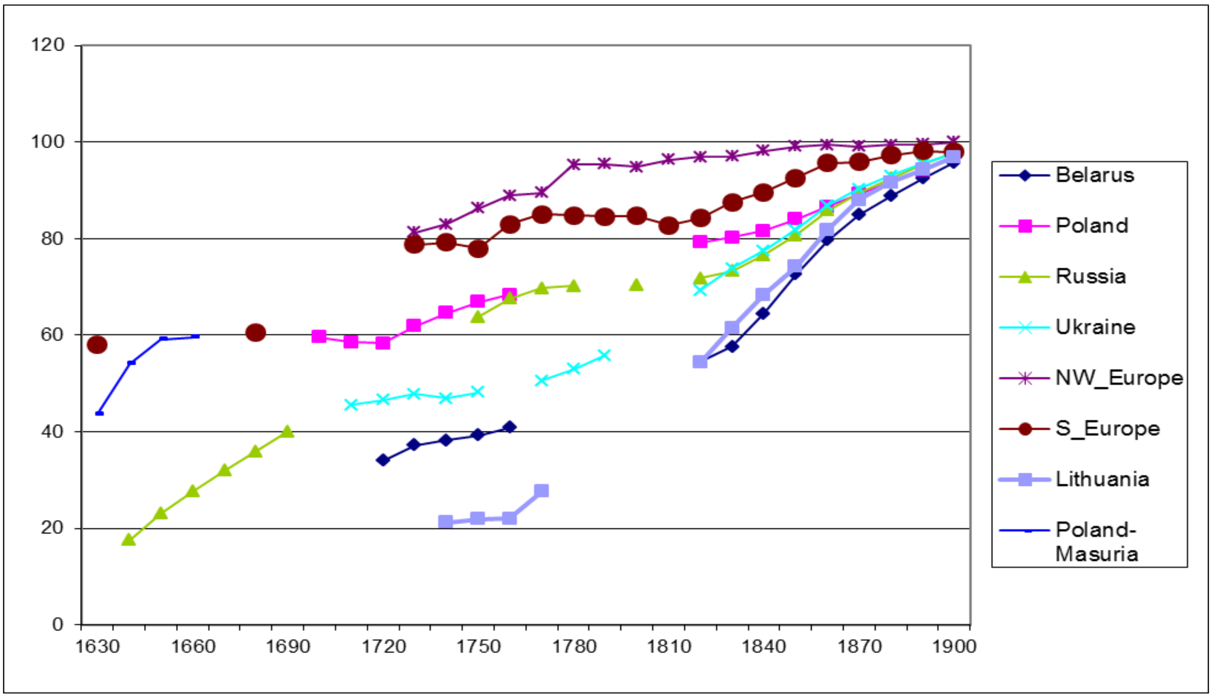
Note: EE – Estonia; LV – Livonia; LT – Lithuania; M – Minsk; T – Tula; P – Pensa; V – Viatka; C – Char'kow; K – Kuban Territories. In the white areas less than 10% of the total population consisted of proprietary serfs. The light areas with diagonal lines indicate a percentage share of 10-30% and the light grey-shaded areas a share of 30-50%. In the dark grey-shaded areas, 50-70% of the total population were proprietary serfs.

Figure 3: Numeracy in the governments of the Russian Empire (ABCC index)



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

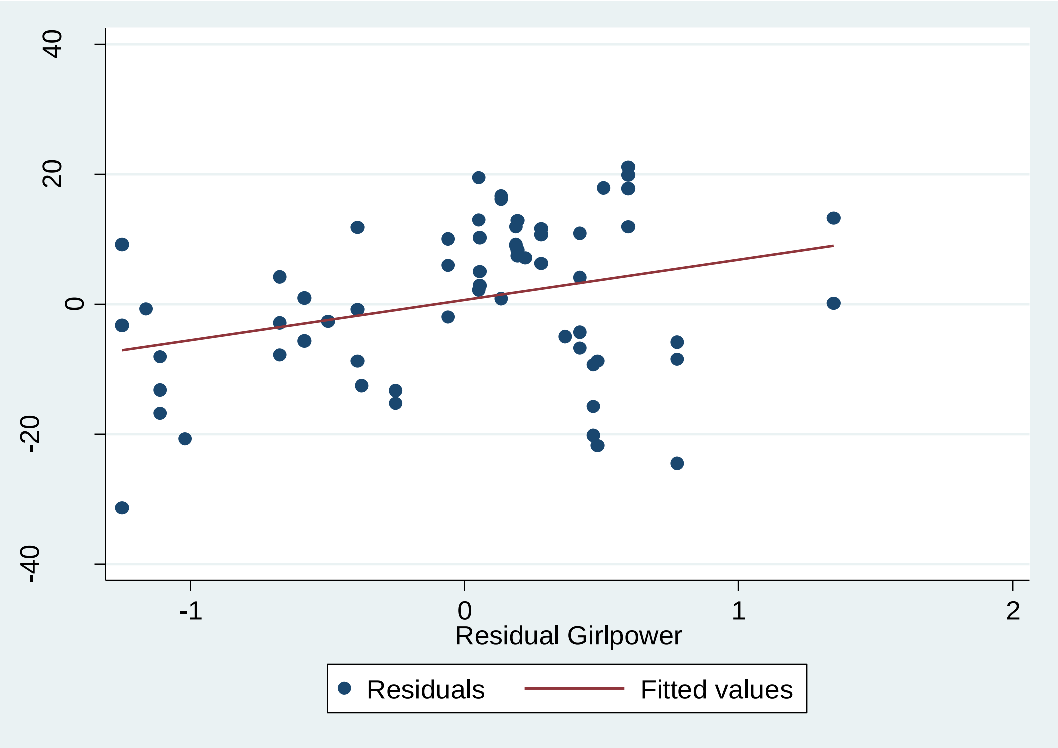
Figure 4: Eastern European ABCC estimates in comparison



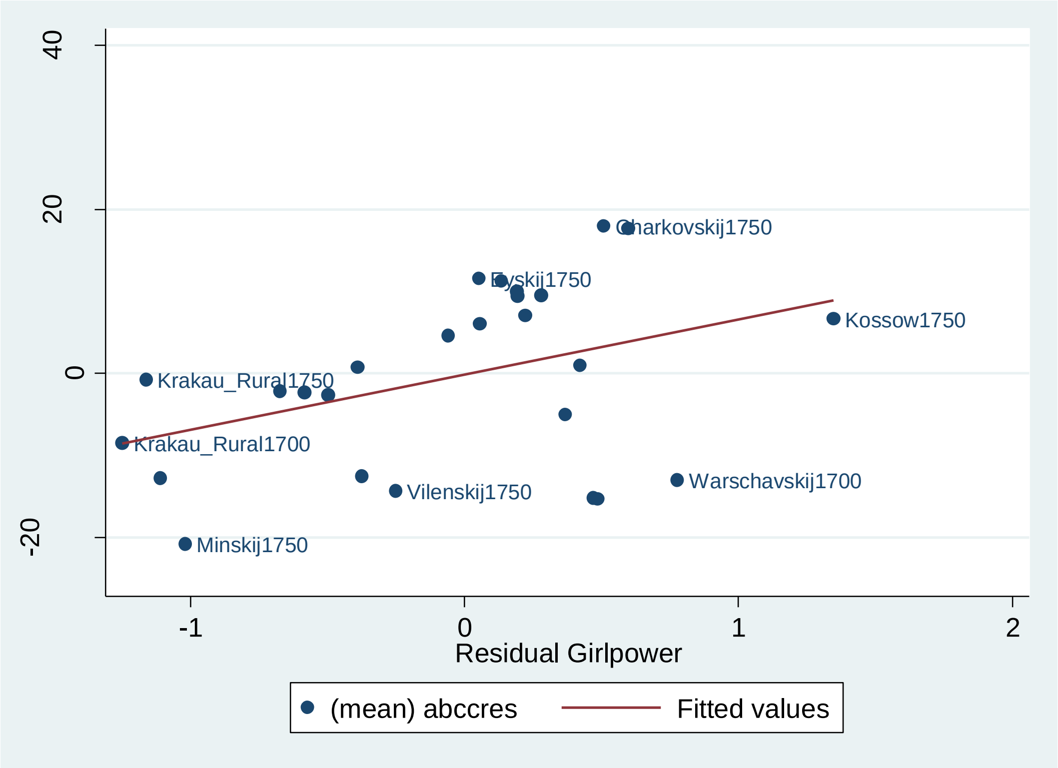
Sources for NW and S Europe: Stolz et al. (2012), southern Europe in 1630 (1680) refers to the average value of Spain and Portugal during the period 1600-49 (1650-99), references in Baten (2016). Note: “Poland” refers to the whole country (including “Poland-E” and “Poland-W.”).

Figure 5: Residual "Girlpower" and numeracy (upper: by region and decade; lower: by region and half century, names added for selected observations)

By decade and region:



By half century and region:



Appendix A: raw and adjusted regional values

Table A.1: Places and Provinces Included (Period before 1880/1897)

Note: Country abbreviations follow ISO-2 standard codes ("by" = Belarus, "lt"= Lithuania etc.). 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 toward the late nineteenth century.

Place/County	Country today	County 1880/1897	Adm. Gouv./Province 1880/1897	Empire
Bobruisk	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-Masuria	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	Przemyshlany (Bezirks-Hauptmannschaft)	Galizien	Austria-Hungary
Wilenski	lt	Vilenskij	Vilenskaja	Poland/Russia
Kcynski	pl	Warschavskij	Warschavskaja	Poland/Russia
Eyskij	ru	Eyskij	Kuban territory	Russia
Moskovskij	ru	Moskovskij	Moskovskij	Russia
Orenburgskij	ru	Orenburgskij	Orenburgskij	Russia
Tulskij	ru	Tulskij	Tulskij	Russia
Vjatskij	ru	Vjatskij	Vjatskij	Russia

Table A.2: The Number of Cases of Our Samples, Birth Decades 1630s-1820s (for the period after 1820, N was always larger than in 1820)

Country	**1630**	**1640**	**1650**	**1660**	**1670**	**1680**	**1690**	**1700**	**1710**	**1720**
Russia		430	1034	1808	2396	2745	915			
Poland-East									144	457
Poland-West/Masuria	69*	110*	97*	196*				57	105	181
Belarus										888
Lithuania										
Ukraine									158	489

Country	**1730**	**1740**	**1750**	**1760**	**1770**	**1780**	**1790**	**1800**	**1810**	**1820**
Russia			119	203	262	341		78		72685
Poland-East	1175	2031	2671	3781						42031
Poland-West	334	722	787	1076					118545	158170
Belarus	1654	2839	4903	7492						27067
Lithuania		68	149	227	268					12030
Ukraine	890	1542	2225		81	384	127		3254	18885

Notes: Asterisk (*) refers to Poland-Masuria. Included are all age groups.

Table A.3, Panel A: ABCC by region (raw values), and regional adjustment factors. 1630-1790

Region	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790
Russia																	
Eyskij													62	73	79	78	
Moskovskij					66	68	69										
Orenburgskij																	
Tulskij		45	48	59	61	57	51										
Vjatskij		26	35	32	33	38											
Russia																	
Poland-East (later Russian)																	
Chenstohovskij											66	60	73	74			
Kaliskij										76	77	72	71	88			
Warschavskij									57	41	50	67					
Poland-East																	
Poland-West (19th C Austria/Prussia)																	
Ermland-Masuria	62	49	84	83													
Cracow_County										34	62	75	75	81			
Posen										66	71	82	83	86			
Oppeln												83	93	97			
Pomerania								83	77	85	87	93					
Poland-West (19th C Austria/Prussia)																	
Belarus																	
Bobrujskij										24	34	37	35	28			
Gomelskij										42	41	48	51	45			
Minskij										43	34	37	40	34			
Pinskij										38	50	51	54	48			
Belarus										34	37	41	42	36			
Lithuania																	
Vilenskij												31	33	35	26		
Lithuania																	
Ukraine																	
Charkovskij									42	44	53	60	67			84	
Kossow															60	73	94
Przemyshlany											28	41	51				
Ukraine																	

Note: The value for the countries is not the average calculated from the sub-regions mentioned in the table, but the average of the whole country for the period after the 1820s. It is based on the census from 1897.

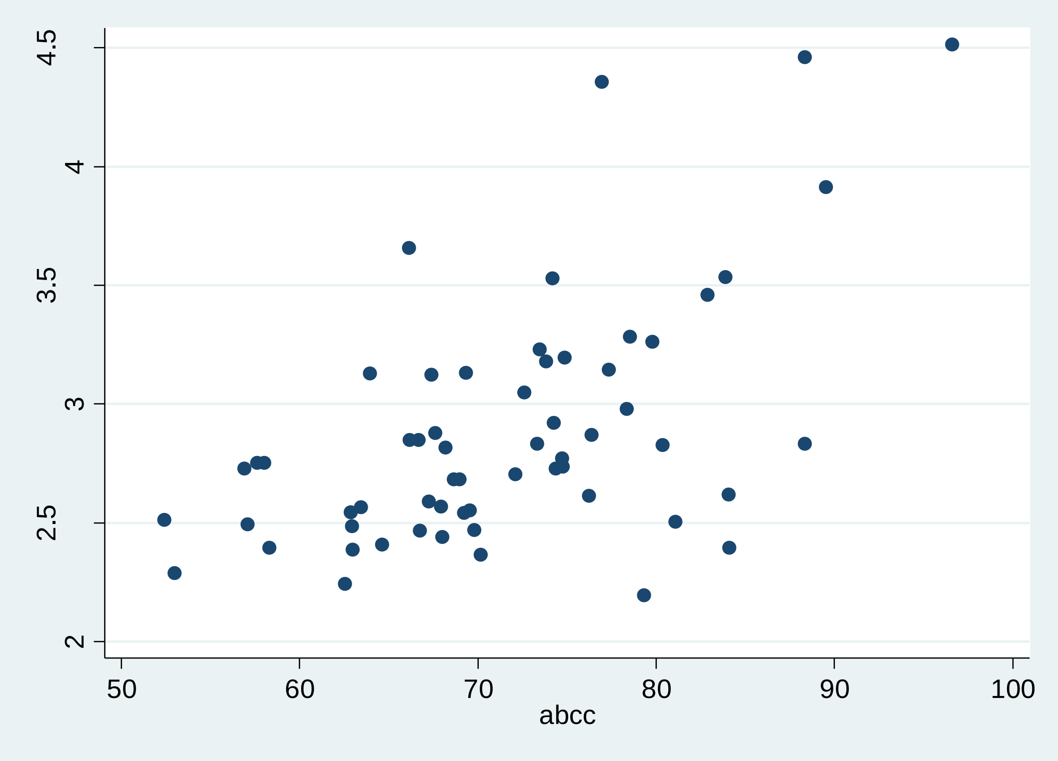
Table A.3, Panel B: ABCC by region (raw values), and regional adjustment factors. 1800-1900

Region	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	Adj.factor (1820-national)
Russia												
Eyskij			65	72	80	87	88					1
Moskovskij			83	84	88	93	96					19
Orenburgskij	90		87	80	75	81	84					23
Tulskij			87	81	81	87	92					23
Vjatskij			78	80	85	90	90					14
Russia			64	69	77	84	86	90	93	95	98	
Poland-East (later Russian)												
Chenstohovskij			76	78	78	89	88					-1
Kaliskij			74	77	77	90	87					-3
Warschavskij			91	90	91	96	98					14
Poland-East			77	78	80	86	87	88	92	94	97	
Poland-West (19th C Austria/Prussia)												
Ermland-Masuria		98	88	89	90	95	91					11
Cracow_County		92	97	95	100	98						20
Posen		92	92	93	95	97						15
Oppeln		93	93	93	95	98						16
Pomerania		98	97	98	98	98						20
Poland-West (19th C Austria/Prussia)			77	78	80	86	87	88	92	94	97	
Belarus												
Bobrujskij			48	54	63	75	81					-2
Gomelskij			50	59	70	82	84					0
Minskij			55	60	66	77	81					5
Pinskij			57	62	68	80	80					7
Belarus			50	55	63	75	82	84	90	92	96	
Lithuania												
Vilenskij			69	68	71	82	87					10
Lithuania			59	61	62	74	86	89	92	94	97	
Ukraine												
Charkovskij			67	72	79	87	91					1
Kossow		98	97	96	96	97						31
Przemyshlany		90	92	97	93	92						26
Ukraine			66	72	78	85	87	91	94	95	98	

Table A.4: 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
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	
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					
Poland-West (19th C Austria/Prussia)																			
Ermland-Masuria	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					
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					
Lithuania																			
Vilenskij												21	23	25	16				
Lithuania												21	23	25	16				
Ukraine																			
Charkovskij									41	43	52	58	66			71			
Kossow															29	42	63		
Przemyshlany											1	14	25						
Ukraine									41	57	42	48	55		29	56	63		

Figure A.1: Comparison of Literacy and Numeracy (ABCC) in the governments of the Russian Empire



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

Figure A.2: Regionally adjusted numeracy (ABCC) of places in Russia

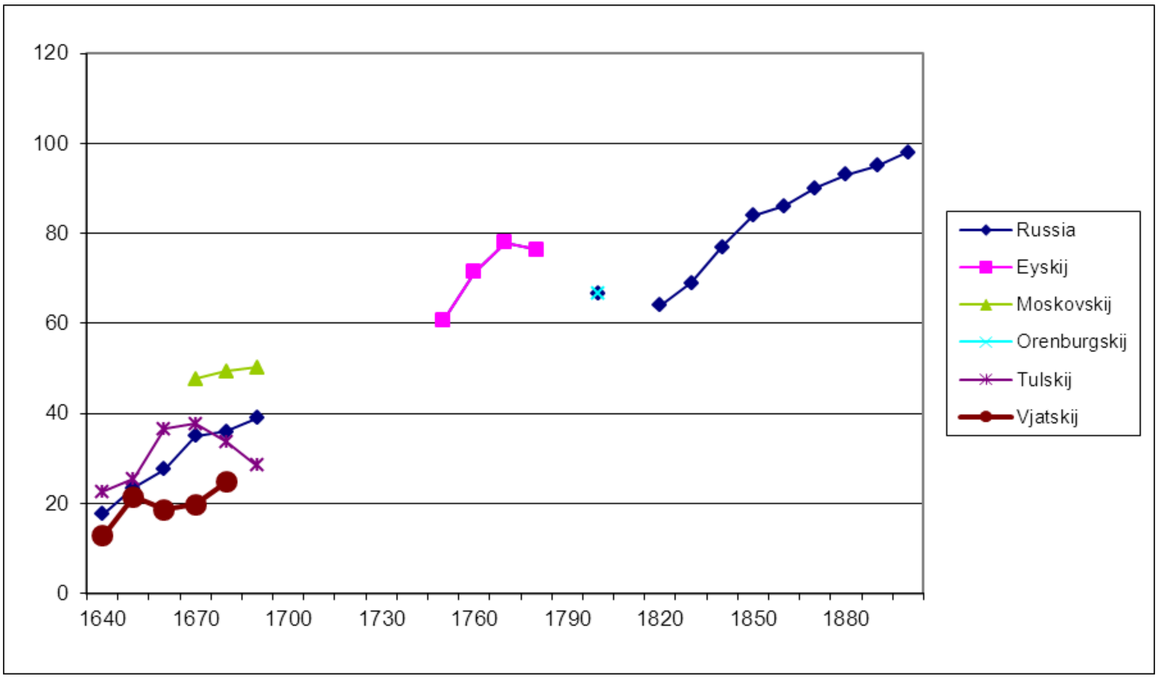
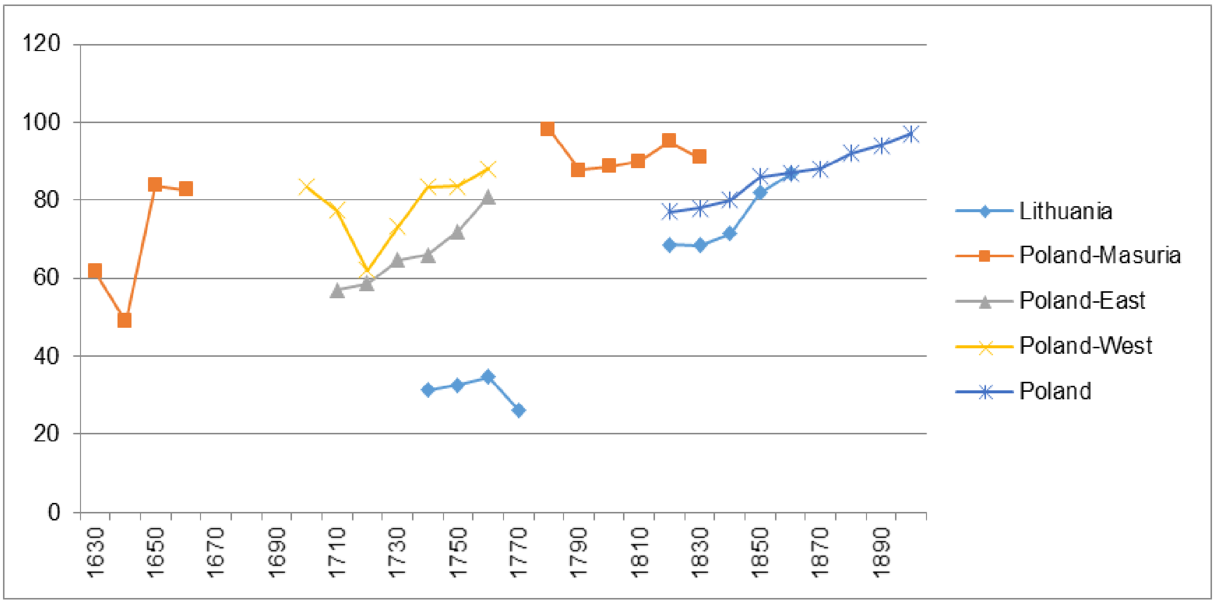
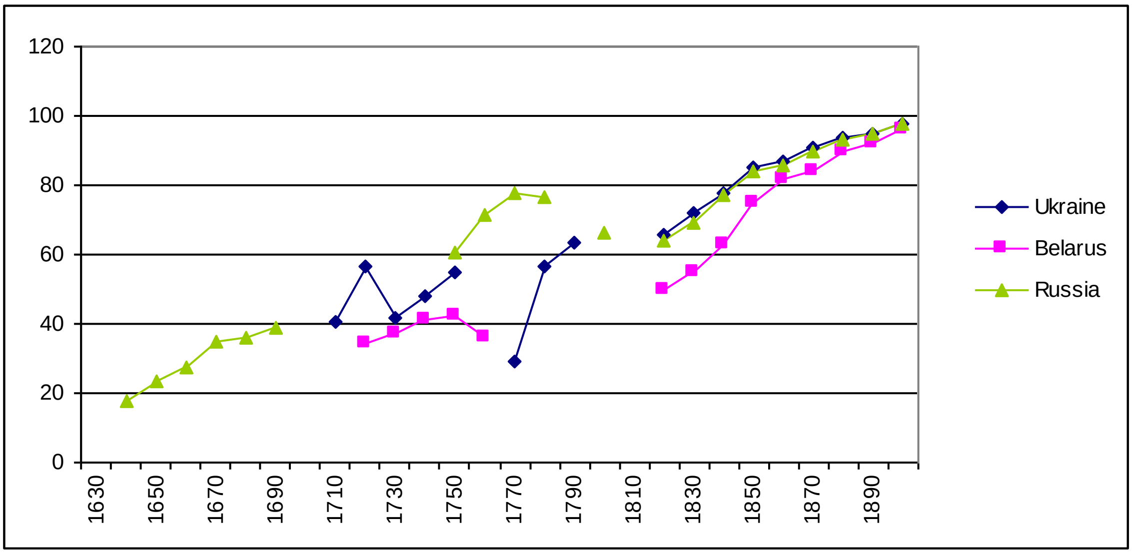


Figure A.3: ABCC country trends

Panel A

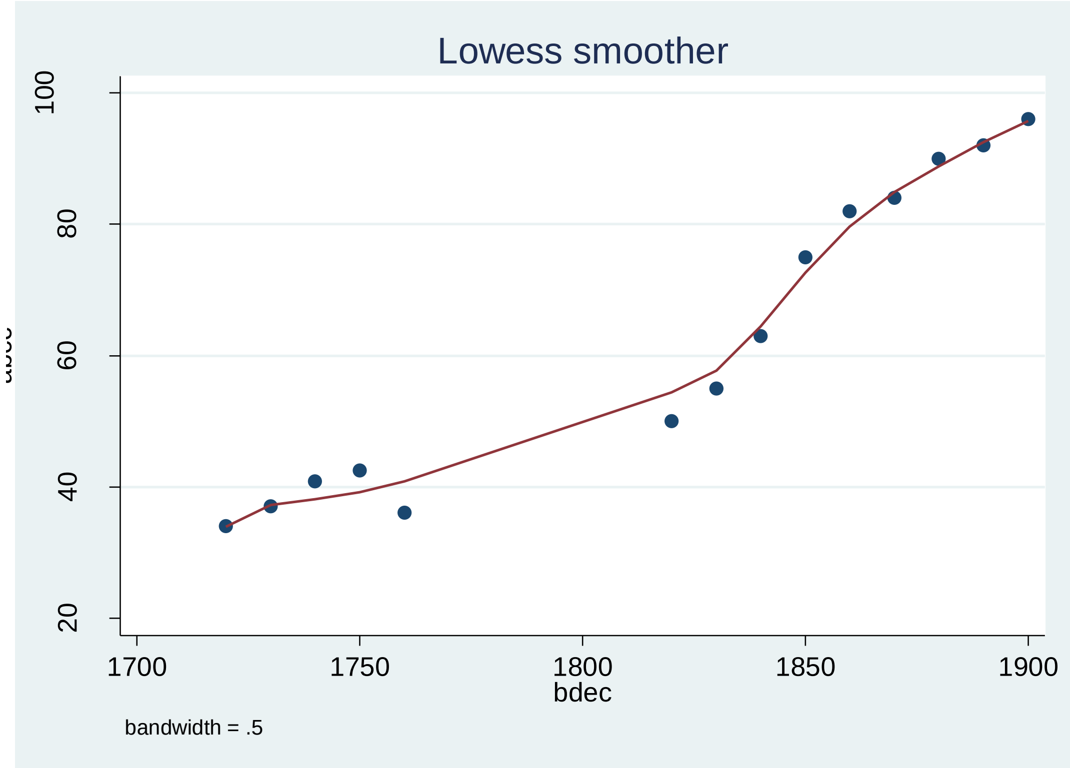


Panel B



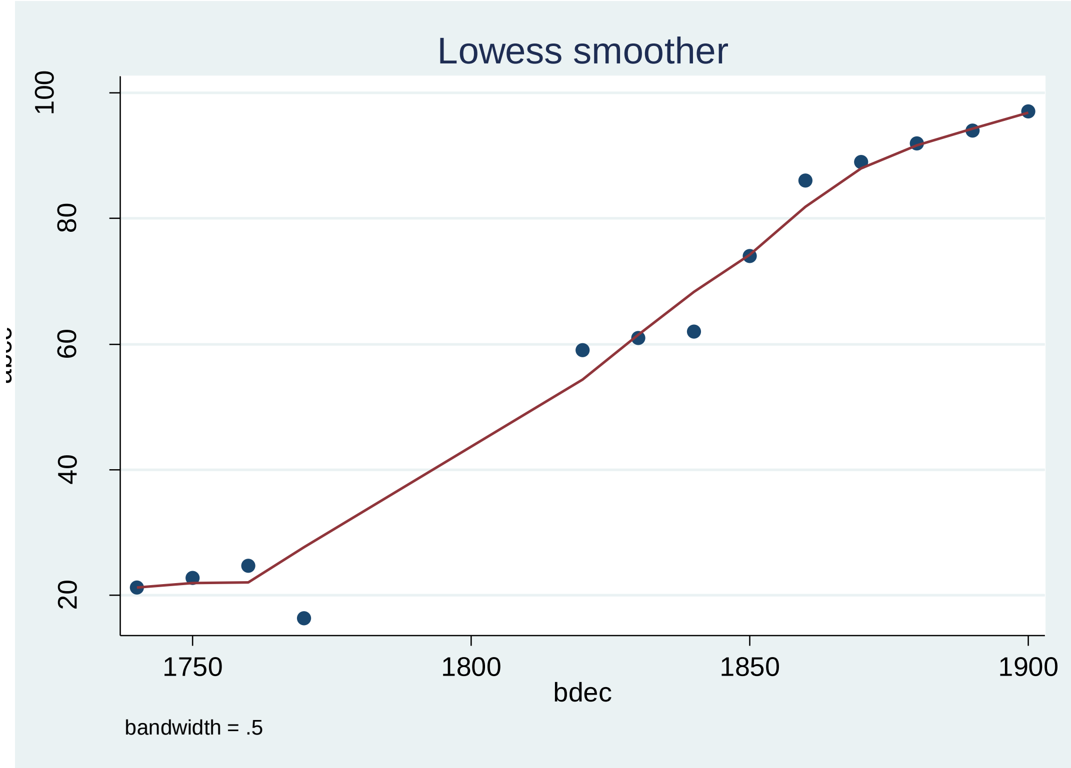
Note: “Poland” refers to the whole country (including “Poland-E” and “Poland-W.”)

Figure A.4: LOWESS-smoothed ABCC trends: Panel A -- Belarus:

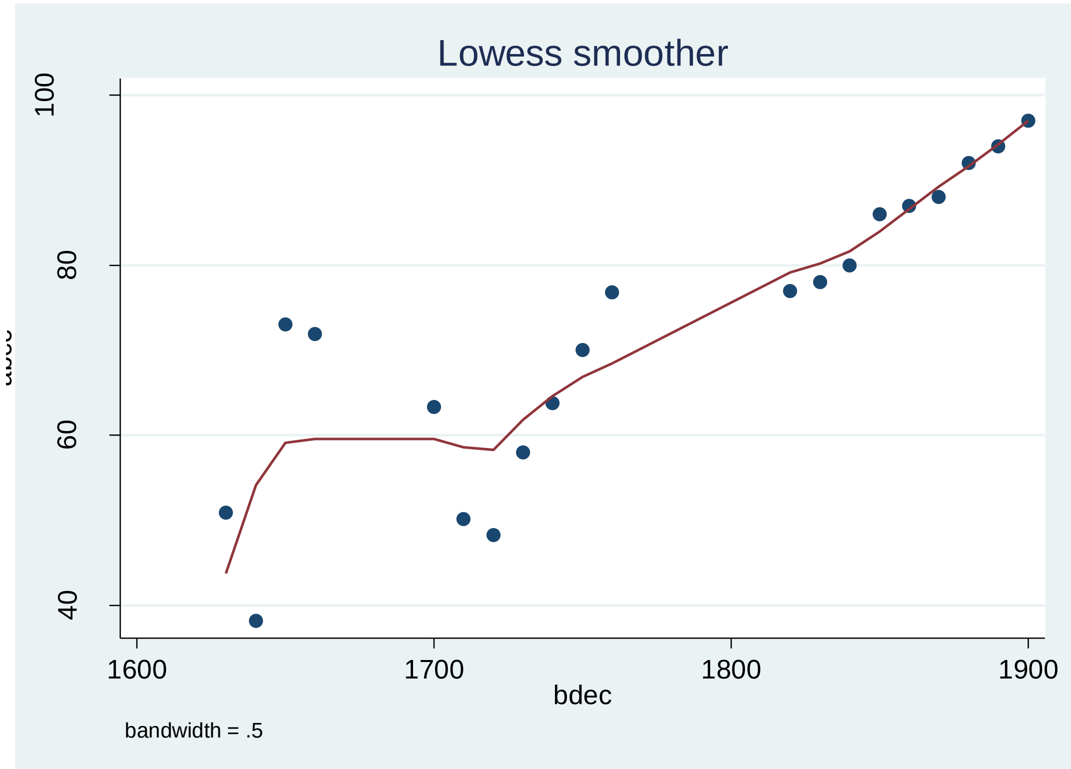


Note: LOWESS estimates a LOcally WEighted Scatterplot Smoothing. The graphical display resembles a moving average, except that the normal moving average assigns the same weight to surounding observations, whereas LOWESS assigns a stronger weight to closer surrounding values and a lower weight to the more distant surrounding values. It fits a local polynomial (least squares fit, outlier-resistant strategy), hence it gives a better impression due to its smoothing technique.
In order to make the comparison, the eastern and western parts of Poland were considered together.

Panel B -- Lithuania:

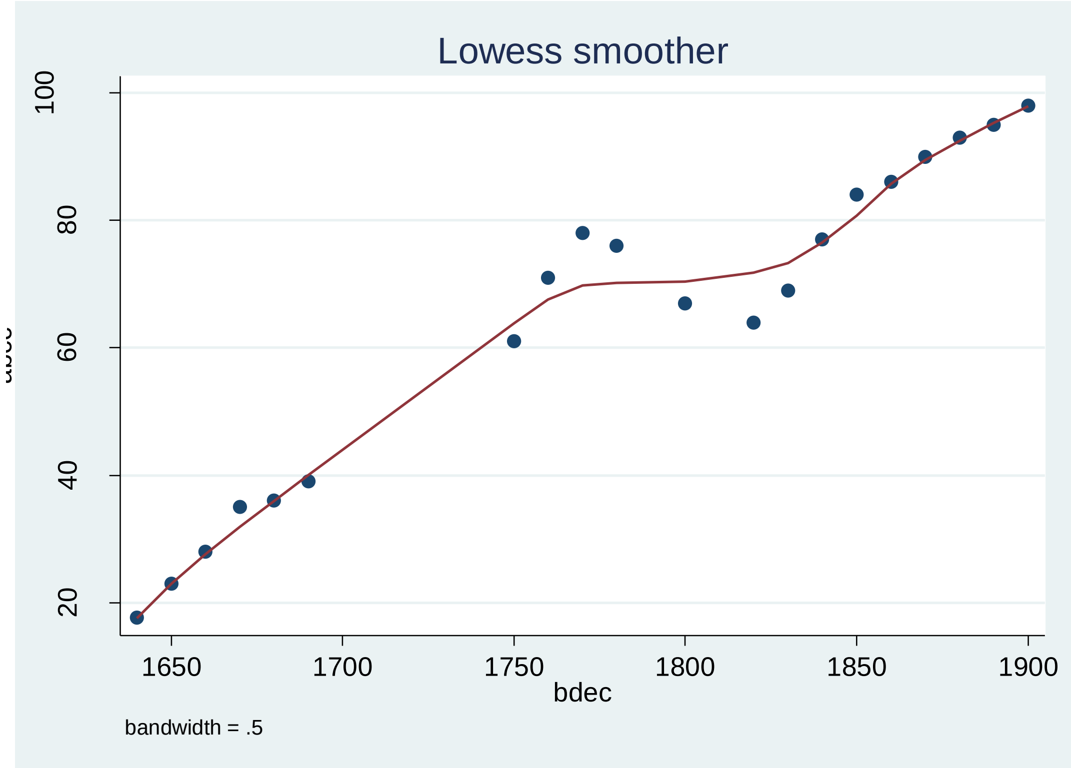


Panel C -- Poland:



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

Panel D -- Russia:



Panel E -- Ukraine:

(Internet) Appendix B: More details on the age-heaping method

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 that 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 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.

A question that needs careful study is 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., Szołtysek 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. It is even possible that the enumerator estimated the age without asking the individual (especially for lodgers, inmates, or other temporary household members such as 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. In the early modern period and the nineteenth 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. Friesen et al. (2011) recently systematically compared the evidence of a gender gap in numeracy and in literacy for the late nineteenth and early twentieth 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.

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 such as the following 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 also have been asked to report their age. For our study, the question of whether the women answered themselves is slightly less important because we only seek to estimate average numeracy.

The problem of different enumerators influencing the quality of age statements has also been studied in a twentieth 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 twentieth century enumerators and the priests and officials of the seventeenth to nineteenth centuries is that the former had much easier access to sources that would enable them to cross-check age statements. Priests of the eighteenth 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. However, as some of the existing sources were clearly cross-checked (yielding ABCC values of approximately 100 very early), we used a historian’s judgment in excluding them. In addition, Szołtysek (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.

A'Hearn, Baten, and Crayen (2009) used a large U.S. census sample to perform a very detailed analysis, and they confirmed a 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.

To assess the robustness of those U.S. census results and the similar conclusions that could be drawn from the less-developed countries of the late twentieth century, A'Hearn et al. (2009) also assessed age-heaping and literacy in 16 different European countries between the middle ages and the early nineteenth century. Again, they found a positive correlation between age-heaping and literacy.

The widest geographical sample studied so far was created by Crayen and Baten (2010b). 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.

References to Appendix B (if not in main references)

BATEN, J. and MURRAY J. "Women's Stature and Marriage Markets in Pre-Industrial Bavaria", *Journal of Family History* 23-2 (1998), pp. 124-135.

BATEN, J. and MUMME C."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, 279 -305.

EWBANK, D. C. (1981). *Age Misreporting and Age-selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis*. Washington, D.C: National Academic Press, 1981.

FRIESEN, J., BATEN J. and PRAYON V. (2011) "Women count. Gender (In-)Equalities in the Human Capital Development in Asia, 1900-1960, Tübingen Working Papers in Economics and Finance 29.

MANZEL, K., BATEN J. and STOLZ Y. (2011) "Convergence and Divergence of Numeracy: The Development of Age Heaping in Latin America, 17th to 20th Century", *Economic History Review* 65, 3 (2012), pp. 932–960. Older Version: CEPR Working Paper. V.

SZOŁTYSEK, J. (2011). *Kreowanie mobilności mieszkańców miast*. Wolters Kluwer.

United Nations Statistics Division (2003). *Nonsampling Errors in Surveys.* New York: United Nations Secretariat.

(Internet) Appendix C: Alternative regional adjustment

To assess the robustness of our regional adjustment method, we also applied a different assumption. Rather than using constant differences in 1820 from a national mean, we scaled the difference by the overall numeracy rate. We apply a two-step procedure, the first step being the original estimate for the countries. The second step follows this strategy. If, for example, a place was 10 percentage points above the national mean of 50 percent in 1820 (i.e., 60 percent) but the national estimate was only 25 percent in the year 1700 according to our first stage estimate, we worked with only half the difference (5 percentage points) to adjust local values in 1700 to represent the national mean, i.e.,

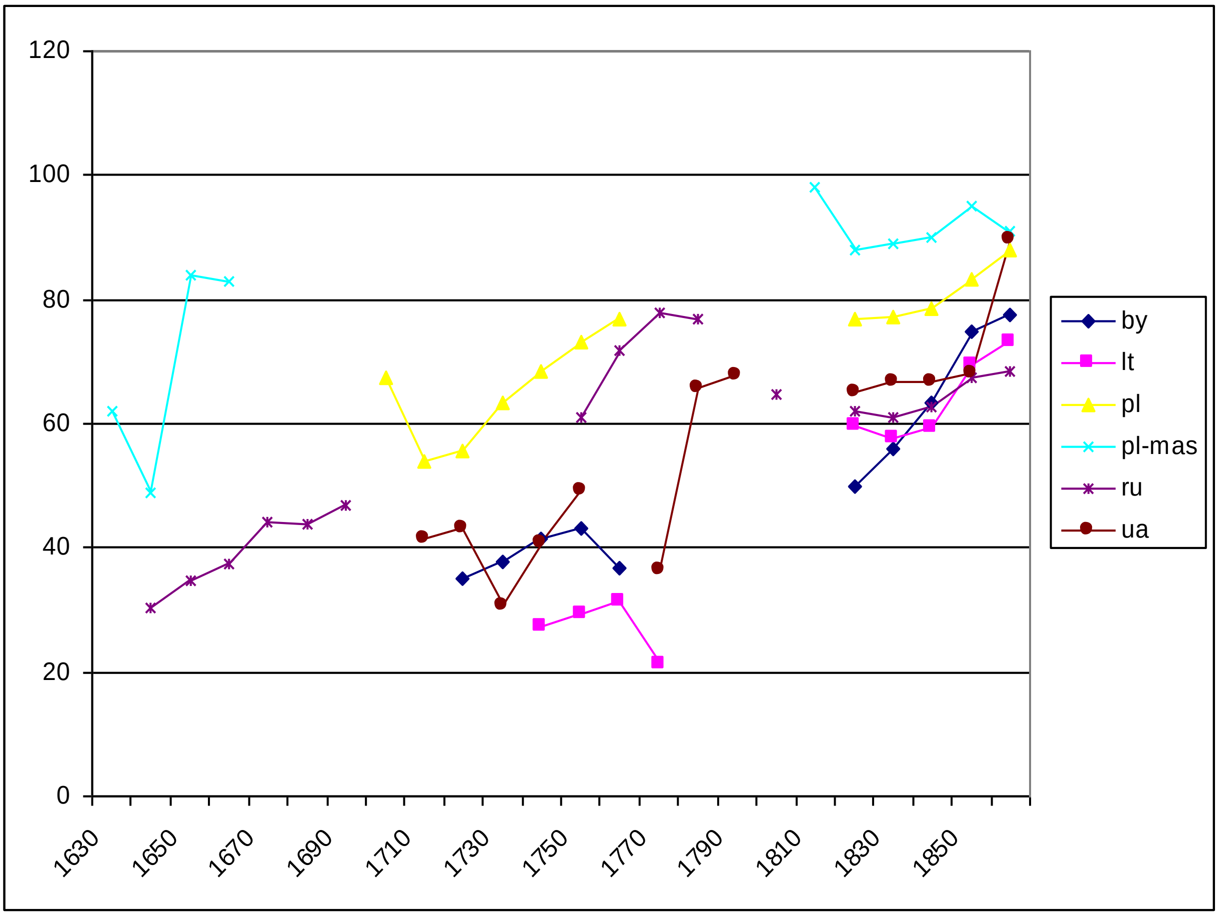
$$N_e^t = L_o^t - \frac{(L_o^{1820} - N_o^{1820})}{N_o^{1820}/N_{efirst}^t}$$

$$Example: \qquad 27 = 32 - \frac{(60 - 50)}{50/25}$$

L_o = *local observed (a)*	$L_o^{1820} - N_o^{1820}$ = *adjustment factor*
N_e = *national estimate*	N^t_{efirst} = *first national estimate*

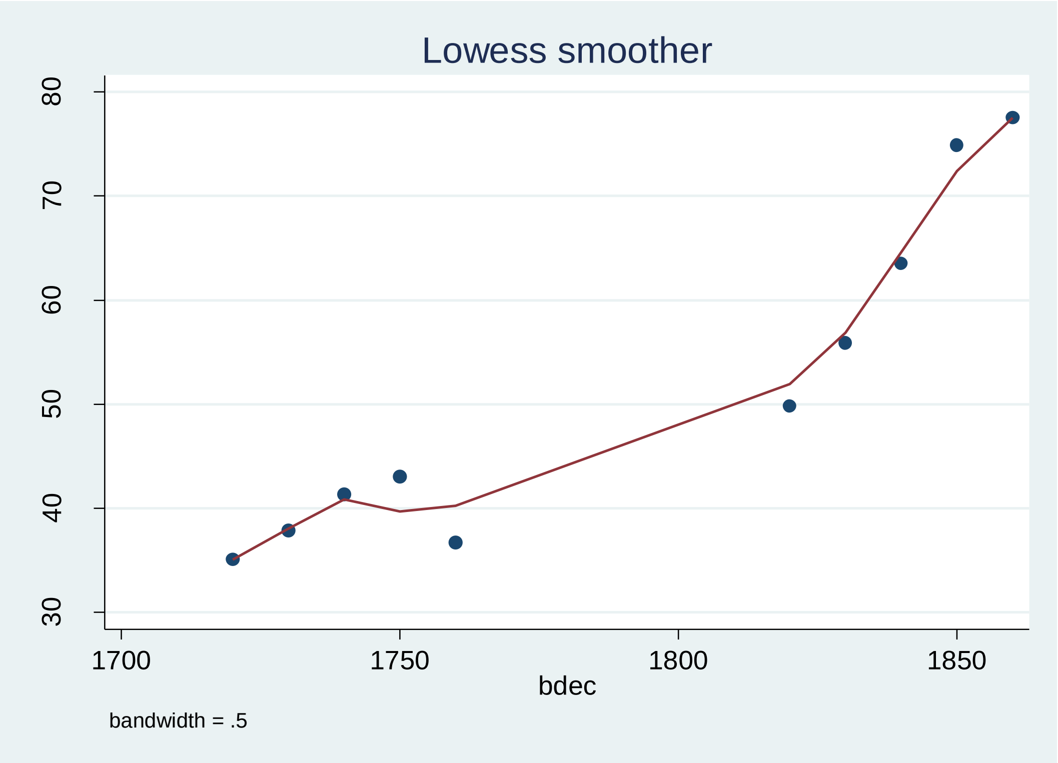
The advantage of this second procedure is that it takes into account the national trends of the countries and makes the adjustment sensitive to possibly different trends. The results are shown in Figure C.1 and C.2. They are very similar to the results in the paper using the simpler method.

Figure C.1: regionally adjusted values, alternative method



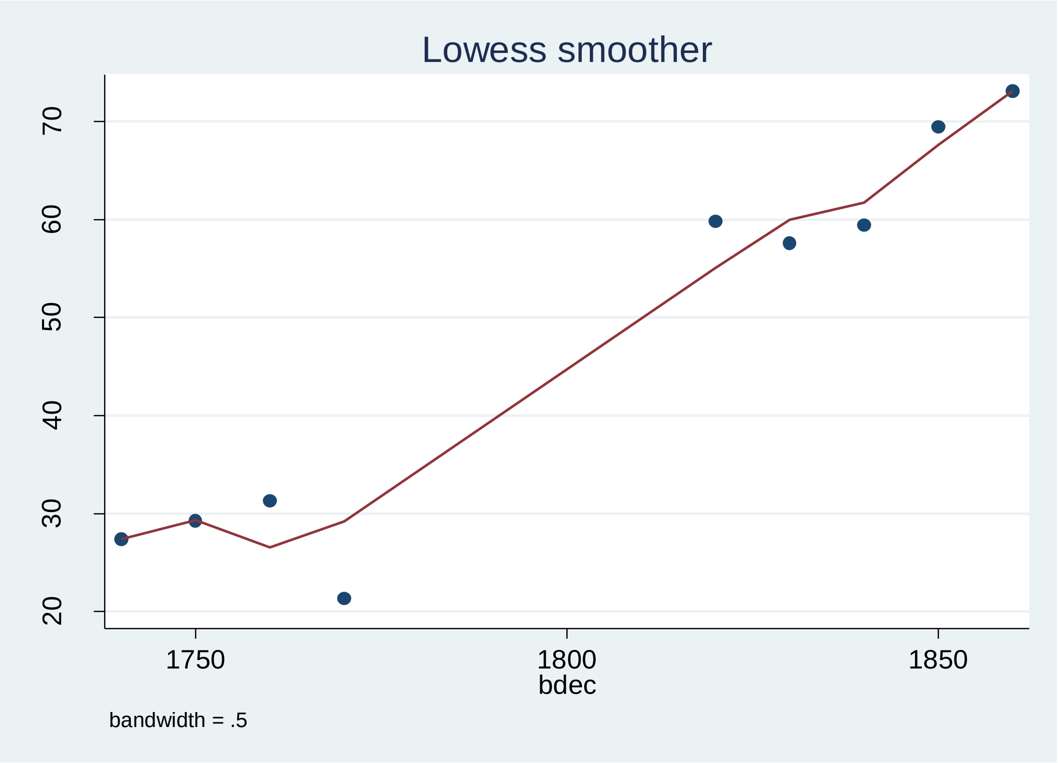
Notes: by = Belarus; lt = Lithuania; pl = Poland; pl-mas= Poland-Masuria; ru = Russia; ua = Ukraine.

Figure C.2: Belarus



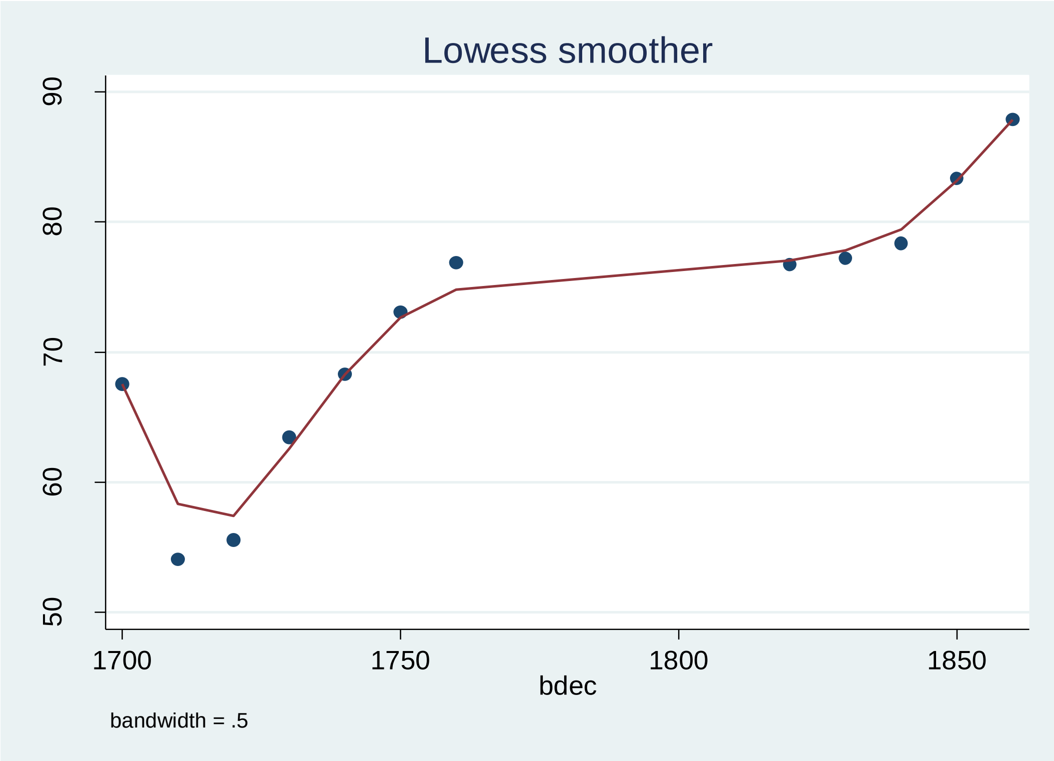
Note: bdec= birth decade

Figure C.3: Lithuania



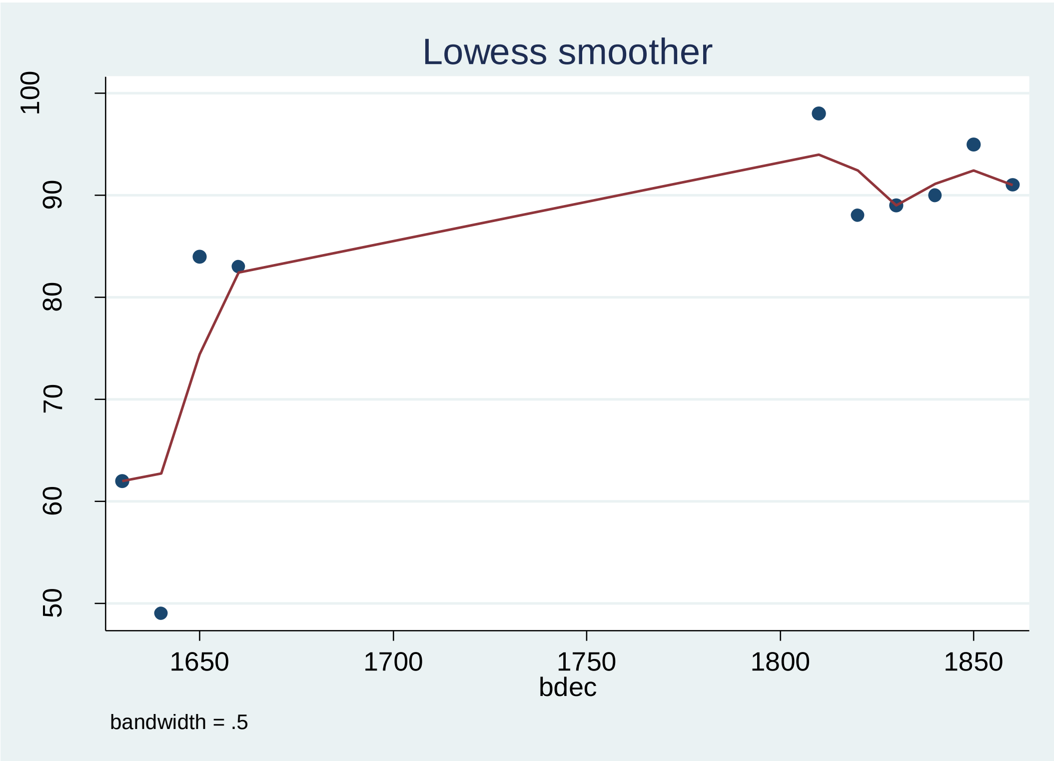
Note: bdec= birth decade

Figure C.4: Poland (excl. PL-Mas.)



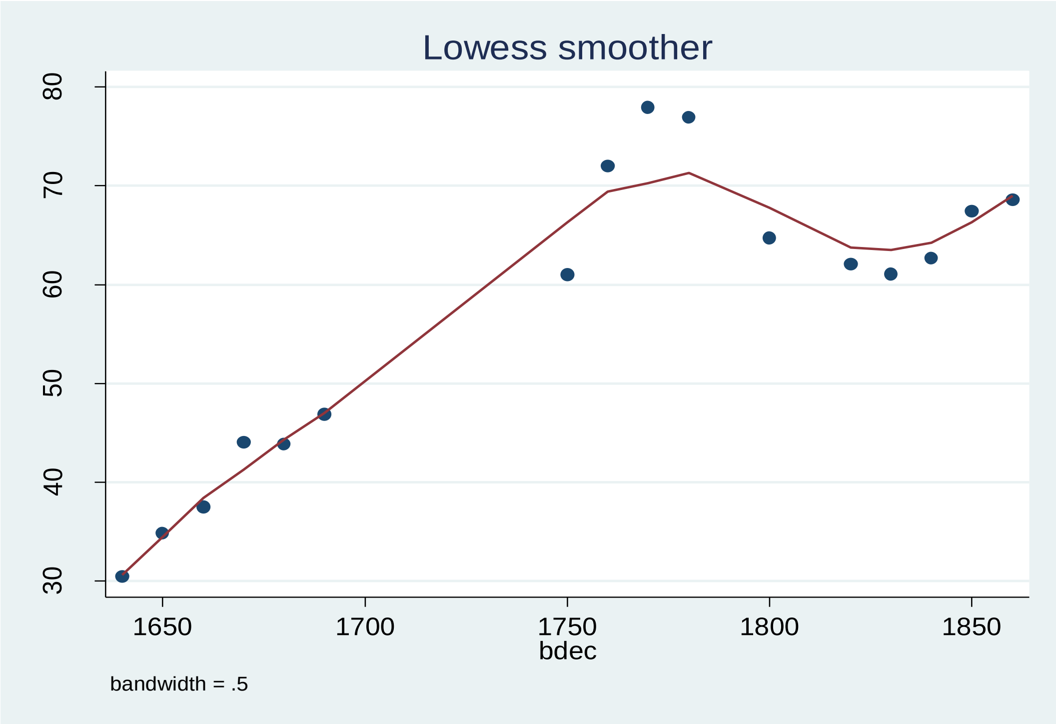
Note: bdec= birth decade

Figure C.5: Poland-Masuria



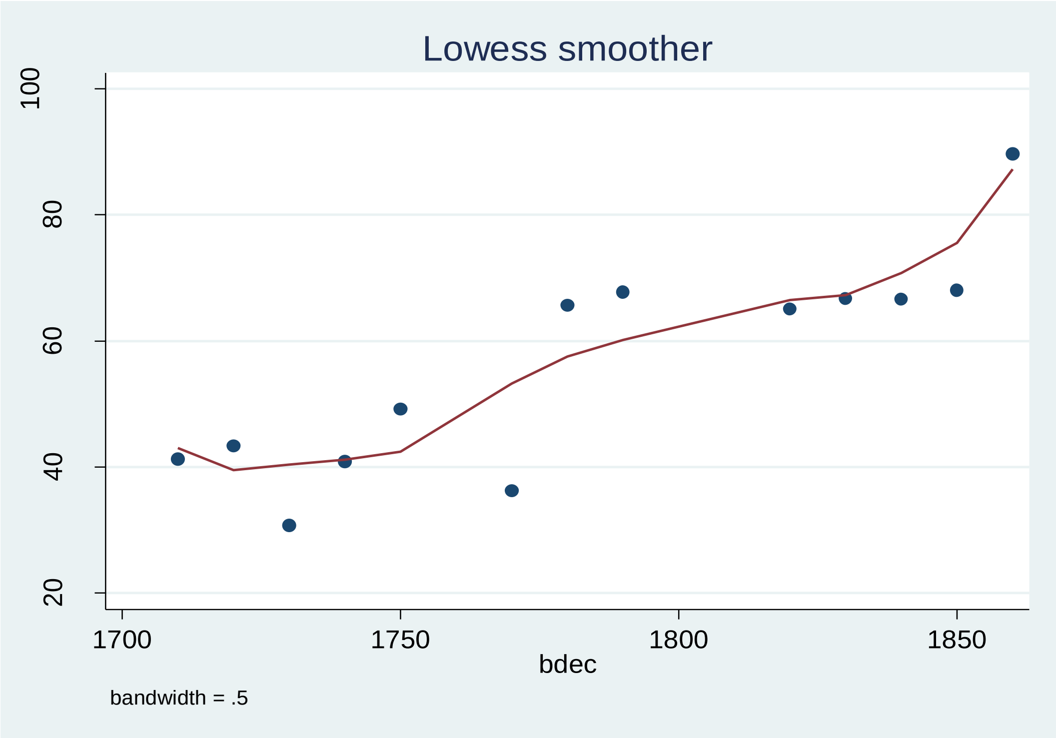
Note: bdec= birth decade

Figure C.6: Russia



Note: bdec= birth decade

Figure C.7: Ukraine



Note: bdec= birth decade

Internet Appendix D: Is the share of never married women at age 50 correlated with poverty?

In Table D.1 we analyzed whether the share of never married women at age 50 correlates with poverty or the opposite, i.e., relatively higher welfare. Dennison (2011, p.71-75) demonstrated that the popular view that marriage was universal in 19th century Russia was not true everywhere. In her sample of one estate of central Russia, females, particularly from poor families, often did not find partners. Can we observe a systematic relationship in provincial data between poverty indicators and the share of never married at age 50?

First of all, we need to keep in mind that never married rates were very low. As welfare/poverty measures we take human height because it has been shown to correlate with health and nutrition quality. Furthermore, nominal daily earnings of the population were included. It would have been better to use real earnings, but there is no harmonized data available for all 49 Russian provinces.

As control variables, the share of Muslims is included. Some of the Russian provinces have a relatively high percentage of Muslims and we would expect this to impact negatively on the share of never married women. Some regions show some degree of voluntary celibacy. This was restricted to the most urban regions like St. Petersburg, Moscow and the Baltic regions, which were the most economically developed in the empire. Thus, a dummy variable for an urban environment, called “metropolis”, and for the Baltic regions, is included. We would expect a positive coefficient.

After the inclusion of all these variables, we find the expected signs. The taller the male recruits in a province, the lower the share of never married women at age 50. The higher the nominal earnings, the lower the share of never married women age.

Multiplying the height coefficient with its standard deviation, we obtain an effect of 0.29. This corresponds to 49 percent of the standard deviation of never-married women. The same calculation for earnings yields 21 percent of the standard deviation of never-married women at age 50. The other control variables have the expected signs, too. Finally, the evidence on wages confirms this relationship.

Table D.1: Regression of the share of never married women at age 50 on welfare/poverty indicators (height, earnings) and other determinants

	(1)	(2)	(3)	(4)
Height	-0.03***	-0.02**	-0.03***	
	(0.000)	(0.027)	(0.000)	
Muslshare	-0.02***		-0.02***	-0.01***
	(0.000)		(0.000)	(0.000)
Baltic	1.69***	1.26***	1.61***	1.09***
	(0.000)	(0.002)	(0.000)	(0.000)
Metropolis	1.25***	1.32***		1.31***
	(0.000)	(0.000)		(0.000)
Earnings				-0.56**
				(0.029)
Constant	45.76***	28.86**	42.94***	-0.88
	(0.000)	(0.046)	(0.000)	(0.386)
Observations	49	49	49	49
Adjusted R-squared	0.53	0.36	0.35	0.41

Note: Robust p-values are indicated in brackets

Table D.2 Descriptives:

Variable	Obs	Mean	Std. Dev.	Min	Max
Never married at age 50 (log)	49	-3.104	0.591	-3.990	-1.775
Height_mm	49	1623.245	9.658	1607	1648
Muslshare	49	6.082	14.745	0	72
Baltic	49	0.082	0.277	0	1

Earnings	49	4.064	0.222	3.784	4.771

Sources: Earnings and Height: Mironov, B. (1998). “Height and Welfare Growth: the Standard of Living in Imperial Russia, 1855-1989, Conference Paper Munich; See also: Mironov, B. (2012). The Standard of Living and Revolutions in Russia, 1700-1917, Routledge:Abingdon; Muslim share: Census 1897, recorded on the internet page of “Demoscope Weekly”.

InternetAppendix E: Review of literacy and school systems in ECE

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. He cited the estimate by the Russian historian Sapunov that a mere one to 1.5 percent of the Russian population in the mid-thirteenth 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 seventeenth century, literacy rates (based on the signatures of witnesses in legal sources) increased to between two and 2.5 percent by the late seventeenth century. Finally, after organizing the 1897 census by birth cohort, 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).

Throughout the territories under consideration, educational institutions were in short supply, and educational inequality was high (See Krumbholz 1982). Especially in Poland-Lithuania, where institutionalized schooling was at the low-moderate level in the early modern period, the nobility and aristocracy could benefit from high-level schools, merchants and urban upper classes from district schools. For example, in the seventeenth–eighteenth century, literacy rates in the Krakowskie voivodship amounted to 98.6 for the high nobility (magnates) and 84.3 for lower nobility (szlachta), but only 9.5 percent for peasants (males only, see Zarys (*2012*), pp. 247). Peasants only had access to primary schools, and hence, the bulk of low-level learning took place as intra-household education and self-learning of children. In the Russian Empire, even primary schooling was not accessible for many peasants until the Orthodox Church created a number of additional schools in the early nineteenth century (with a religious bias). Only in the second half of the nineteenth century did schooling institutions develop more rapidly. When Poland was partitioned in the late eighteenth century, the political leaders of the Russian, Prussian and Austria-Hungarian empires did not favor education for the Polish population, especially not in their native language (Baten 2003).

References to Appendix E (not in main reference list)

BATEN, J. (2003). 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, 383-422.

KRUMBHOLZ, J. (1982). *Die Elementarbildung in Russland bis zum Jahre 1864.*Wiesbaden: Steiner.

REIS, JAIME (2005). “Economic Growth, Human Capital Formation, and Consumption in Western Europe Before 1800.” In *Living Standards in the Past*, Robert C. Allen, Tommy Bengtsson, Martin Dribe (ed.), Oxford: OUP, 195–225.

ZARYS historii Polski w liczbach. (GUS, Warszawa, *2012*).

InternetAppendix F: Labor during child ages

Labor is negative at child age as children do not develop well under the usual pressure of early modern labor market situations. Given that children have not yet reached sufficient maturity, their productivity is usually established by harsh punishments and other disciplinary tools. Examples are contemporary descriptions of child labor in mining and textile factories but also child labor in agriculture, in which the punishment of children is described intensively. In contrast, women in their late teens or early twentieth have in most cases reached sufficient maturity to be able to gain work experience.

InternetAppendix G: Nutrition and Numeracy

In the early-modern disease environment, human beings needed much more protein than today to develop antibodies in their fight against infectious disease. This is also reflected in higher human stature values of persons with better net nutrition. Paxson and Schady (2007) studied whether height can be used as an indicator of cognitive abilities. They found for a large sample from modern Ecuador that the test scores for children of lower height were significantly lower relative to the taller children. Paxson and Schady argued that while parental genes might play a certain role in individual human stature, the nutritional impact on cognitive ability was the decisive factor. Magnusson et al. (2006) confirmed this with a twin study: keeping genetic influences constant, there is still a substantial correlation between height and cognitive ability. Malnutrition is a forecasting indicator of limited cognitive abilities, especially during the first one to two years of life (Lloyd-Still 1976). Many studies have confirmed the influence of protein availability on numeracy, such as Baten and Fourie for Africa (2014). Milk could not be transported over more than a few kilometers before refrigeration technology was introduced. Hence, poor regions sometimes benefitted from a "protein proximity" factor, and this introduced a strong exogenous component to the variable "cattle per capita". Baten et al. (2015) assessed the "natural experiment" of the blockade Napoleon introduced against the British in the early 1800s. This blockade caused severe malnutrition among the poorest. The cognitive ability and the numeracy of the British "blockade cohorts" was substantially lower. The British poor and vulnerable groups who suffered the most during their first decade of life subsequently were only able to find work in poorly paid occupations when competing with cohorts born before or after them.

References to Appendix G (not in the main reference list)

BATEN, JOERG AND JOHAN FOURIE (2014), “Numeracy of Africans, Asians, and Europeans during the Early Modern Period: New Evidence from Cape Colony Court Registers”, *Economic History Review* 68, 2 (2015), 632–656.

LLOYD-STILL, JOHN. 1976. “Clinical Studies on the Effects of Malnutrition During Infancy and Subsequent Physical and Intellectual Development.” *Malnutrition and Intellectual Development*. Lancaster: MTP.

MAGNUSSON; P., F. RASMUSSEN AND U. GYLLENSTEN (2006). “Height at Age 18 Years is a Strong Predictor of Attained Education Later in Life.” *International Journal of Epidemiology* 35:1: 658-63.

PAXSON, CHRISTINA AND NORBERT SCHADY. 2007. “Cognitive Development among Young Children in Ecuador.”*The Journal of Human Resources* 42 (1): 49-84.

InternetAppendix H: Girl Power and literacy

As a robustness check, we also performed regressions of literacy and obtained results of similar sign (see table H.1) but less statistical significance.

Did the determinant age at marriage and “milder serfdom” have a different effect on literacy as a dependent variable? An answer to this question can potentially be obtained by calculating the standard deviation effect on numeracy and literacy. We obtain a numeracy increase of 3.74 by multiplying the standard deviation (0.64) with the largest coefficient of marriage age (5.85; note that because measurement error biases coefficients downward, we use the largest coefficient.). This represents 33 percent of the standard deviation of numeracy (11.46), whereas if we perform the same calculation for literacy, we obtain a product of coefficient (0.66) and standard deviation (0.64) of 0.42 (Appendix A, Table A.6). This represents 77 percent of the standard deviation of literacy (0.55). Hence, the effect of girl power on literacy between the Russian provinces is even stronger than on numeracy. In contrast, in terms of serfdom, the effect is between 70 and 100 percent of the standard deviation of numeracy, whereas for literacy it is mostly insignificant. The largest coefficient of 0.40 multiplied with the standard deviation is 0.19. It represents 34 percent of the standard deviation of literacy. Hence, literacy might, in fact, be slightly less affected by serfdom compared to numeracy. It might be that the incentives for numeracy investment in the family were even more strongly reduced by tough serfdom compared to the effects on literacy. As Hanushek and Woessmann (2012) argued for a particularly strong role for math skills and numeracy, the effect of serfdom on overall development was probably even stronger.

Table H.1: Regression of literacy: cross-section of provinces of the Russian Empire (nineteenth century)

	(1)	(2)	(3)	(4)	(5)
Region omitted	None	None	None	None	Russia
Girl power	0.51***	0.50***	0.54***	0.66***	0.43***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Milder Serfdom	0.16	0.21	0.27	0.40***	0.07
	(0.386)	(0.208)	(0.109)	(0.003)	(0.674)
Cereal suit.	-0.00		-0.01***	-0.00	0.00
	(0.509)		(0.002)	(0.333)	(0.561)
Protein Proximity	0.12	0.12	0.20***	0.13**	0.11
	(0.236)	(0.233)	(0.006)	(0.017)	(0.318)
Pop. Density		0.34***			
		(0.001)			
Orthodox				-0.34**	
				(0.021)	
Protestant				0.38***	
				(0.004)	
Altitude					-0.38**
					(0.019)
Ruggedness					0.31
					(0.316)
Constant	4.27***	4.08***	4.41***	4.80***	4.20***
Observations	48	48	48	18	48
Adjusted R-squared	0.33	0.34	0.48	0.94	0.39

Note: Robust p-values are indicated in brackets.

Table H.2: Descriptives for literacy analysis

Variable	Obs	Mean	Std. Dev.	Min	Max
Literacy (log)	48	2.8	0.55	2.20	4.51
Girl power	49	0.18	0.12	0.05	0.47
ln(Girl power)	49	-1.91	0.64	-3.09	-0.75
Milder Serfdom	49	0.31	0.47	0.00	1.00
Ln(Cattle p. c.)	49	-3.04	0.34	-4.30	-1.09
Religion (orth.)	49	0.90	0.31	0.00	1.00
Cereal suitability	49	62.73	19.97	10.60	94.90
Altitude	49	1.50	0.60	-0.13	3.22
Ruggedness	49	0.46	0.29	0.14	1.89

Appendix I: OLS- regressions, omitting the Milk and serfdom variables

	(1)	(2)	(3)	(4)	(5)	(6)
Poverty adj.	Yes	Yes	No	Yes	No	Yes
Region omitted	None	None	None	None	None	None
Estimation techn.	OLS	Robust	OLS	OLS	OLS	OLS
Girl power	5.57*	6.03***	7.52**	5.57*	7.17**	5.51*
	(0.084)	(0.007)	(0.041)	(0.078)	(0.021)	(0.071)
Cereal suit.	-0.24*	-0.31***		-0.35***	-0.36**	0.53
	(0.091)	(0.000)		(0.008)	(0.019)	(0.719)
Pop. Density				0.10		0.53
				(0.953)		(0.719)
Altitude				-0.24	-0.77	
				(0.000)	(0.000)	
Ruggedness				0.46	-3.24	
				(0.937)	(0.650)	
Constant	87.82***	94.26***	89.86***	95.34***	96.25***	97.23***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	52	52	52	49	49	48
Adjusted R-squared	0.19	0.32	0.23	0.37	0.40	0.36

Appendix J: OLS- regressions, restricting all regressions to to 48 cases, Russia

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Poverty adj.	Yes	Yes	Yes	No	Yes	No	Yes
Region omitted	None	None	None	None	None	None	None
Estimation techn.	OLS	Robust	OLS	OLS	OLS	OLS	OLS
Girl power	4.81**	5.47***	3.84	5.53**	5.06**	6.79**	5.41**
	(0.039)	(0.007)	(0.109)	(0.045)	(0.030)	(0.020)	(0.047)
Milder Serfdom	9.00**	7.79**	14.06***	14.03***	9.48**	9.36**	9.89**
	(0.011)	(0.017)	(0.003)	(0.002)	(0.011)	(0.010)	(0.015)
Cereal suit.	-0.25***	-0.24***			-0.30***	-0.27***	-0.34**
	(0.000)	(0.001)			(0.003)	(0.009)	(0.025)
Protein Proximity	5.29**		5.52	4.62	5.66**	4.31*	5.73**
	(0.034)		(0.112)	(0.174)	(0.023)	(0.079)	(0.030)
Pop. Density					1.75		2.07
					(0.296)		(0.259)
Altitude						1.33	1.56
						(0.687)	(0.643)
Ruggedness						-0.95	-1.85
						(0.868)	(0.765)
Constant	101.14***	102.95***	82.92***	81.74***	101.80***	99.47***	102.53***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	48	48	48	48	48	48	48
Adjusted R-squared	0.48	0.44	0.33	0.37	0.48	0.50	0.45

Appendix K: Political institutional differences

Part of the underlying reasons for educational underinvestment might have been political institutional differences. In a working paper version, we tentatively included Acemoglu et al.'s (2002) institution variable, being well aware that the measurement quality within ECE is low (not included here, results are available from the authors; the variable had expected signs, but was only sometimes significant).[15] It mostly reflects the slightly higher checks exerted by the Polish nobility. It is doubtful that the additional rights garnered by the nobility always had positive influences on economic growth. However, a lower system of checks might have led to more detrimental wars (Kochanowicz 1991).

[15] We thank an anonymous referee for this insight.

Appendix L: Correlation of Numeracy and Literacy in Russia?

When comparing the log literacy in the Russian Imperial census of 1897 for individuals born between 1825 and 1884 on the vertical axis and their numeracy levels (see Appendix A, Figure A.1) on the horizontal axis, we can see that there is a correlation. The values found for the Baltic regions of Estonia and Livonia, as well as in the capital region of St. Petersburg, were very high, whereas the regions around Belarus had quite low values of both literacy and numeracy. The correlation coefficient between log (literacy) and numeracy is.

Interestingly, the northeastern districts of European Russia such as Archangelsk, Wologda, and Perm had a quite high numeracy level (Figure 3). Literacy rates were lower in comparison. 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 rather than 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