A STORY OF LARGE LANDOWNERS AND MATH SKILLS: INEQUALITY AND HUMAN CAPITAL FORMATION IN LONG-RUN DEVELOPMENT, 1820-2000

This is the last working paper version before this study was submitted and accepted. Please cite as

Baten, Jörg and Juif, Dácil (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 (2014), 375- 401. Online: https://doi.org/10.1016/j.jce.2013.11.001

1. Introduction

"How can they understand politics, if they do not even know where they are standing? They might vote for the communists, like the mining workers in the North…" (Isabel Allende, 1982).[1]

In Isabel Allende's novel "La Casa de los Espíritus", the large landowner Esteban Trueba was sure that the day laborers on his estate were not capable of making wise political decisions and should therefore not acquire voting rights. However, he was most likely an exceptional landowner in Chile at the time, as he built a school for his workers during the early 20th century.

Recently, Galor, Moav and Vollrath (2009) (henceforth GMV) argued that two different elite groups strongly influenced political decision-making about educational reforms: large landowners and industrialists. "Latifundia" (large estate) owners were

[1] Own English translation from the Spanish original version of "Ghost House".

typically not interested in tax-financed mass-schooling programs that would remove unskilled day laborers from the agricultural workforce on their estates or make them rebellious. Industrial capitalists, in contrast, typically favored a workforce that had at least some basic skills. In countries that displayed a large amount of land inequality, such as some Latin American countries, the elite landowner group gained political power because their large estates provided financial and social backing to influence the ruling groups. In the process of industrialization, countries with a more equal distribution of land overtook countries with high inequality and performed better in terms of modern income growth. According to Wegenast (2009, 2010), plantation owners also had political reasons to neglect educational expansion, for instance, by monopolizing the decision-making process for literacy requirements for suffrage. GMV provided some evidence regarding regional land inequality and school investment in the U.S. from 1880-1940 as well as some qualitative evidence on land reforms in East Asia and Russia. Hippe and Baten (2012) recently tested and confirmed the link between land inequality and human capital development for European regions.

However, the implications of these theories have not been tested for a large cross-section of countries over the long run, partly because evidence of early inequality for poorer countries was scarce. We create a new dataset to test the influence of land inequality on long-run human capital formation. The methodological idea is to (1) use evidence of 20th century land inequality, (2) assess the influence of land reforms and (3) estimate late 19th century land inequality on the basis of backward projection.

In this study, we assess the influence of both within-country land and income inequality on human capital formation for the period from 1964 to 2003. As a measure of income inequality, we use a recently created global dataset by Van Zanden, Baten, Földvari and Van Leeuwen (2012) and Blum and Baten (2011). This dataset documents within-country income inequality, employing anthropometric inequality measures as well as income distribution (and GDP per capita divided by wages of unskilled workers, so-called “Williamson Ginis”, and other indicators).[2]

Which measure of human capital should be employed as a dependent variable? We argue that a measure should be used that is most conducive to economic growth. For example, school years have often been criticized because the productivity of a school hour differs between countries and cultures and, thus, is not a perfect growth predictor. Recently, the leading human capital economists Hanushek and Woessmann (2012a, 2012b) argued that cognitive skill test results related to math and science abilities are the strongest correlates of economic growth. They extended the famous PISA results from the 2000s into the period 1964-2003 by recalibrating a large number of international math and science tests; they also developed a comprehensive index of those core skills that will be our dependent variable.

A second major contribution of our study is our assessment of the long-run persistence of numerical cognitive skills, which are an important component of the Hanushek-Woessmann measure. We include a new measure of numeracy around 1820 that is constructed on the basis of age-heaping estimates. We argue that countries with early investments in numerical education (and perhaps cultures that promoted numerical skills) entered a path-dependency of human capital-intensive industries, including skill-intensive agriculture and services. Those countries that took the numerical lead (but not necessarily the lead in living standards) in the 1820s were situated in Scandinavia, Central and Western Europe, and East Asia. In the second

[2] For example, in 1890, 11 countries can be documented with income-based Gini estimates, and 68 countries can be documented with anthropometric inequality measures (on the anthropometric measures, see also Baten 1999, and Moradi and Baten 2005).

wave, North American, Southern and East-Central European countries followed. Other world regions lagged behind.

The next section provides a literature overview on the determinants of human capital formation, such as institutional quality, geography, fertility choice and physical capital. In section 3, we introduce the data and sources. Section 4 and 5 describe the OLS and instrumental variable test regression results of our empirical study. Section 6 concludes.

2. Components of Human Capital Formation

A number of theories have been proposed to explain the great divergence in education that took place over the past two centuries and had a strong influence on income divergence. Some influential studies have emphasized the detrimental effect that early inequality had on the emergence of human capital and growth-promoting historical institutions (Engerman and Sokoloff 1997, 2000; Acemoglu et al. 2001, 2002).[3]

Sokoloff and Engerman (2000) hypothesized that in more unequal societies, the elites gained power to influence the choice of legal and economic institutions. In those countries that were unequal, a small number of elites restricted the rights of the majority of people, such as in education and voting, to perpetuate the existing social structures and maintain or even reinforce their elite status.[4]

GMV (2009) set up a theoretical model showing that an unequal land distribution in a country negatively affects per capita income in the long run by

[3] Institutions are normally divided into growth-promoting and growth-retarding types of institutional set-ups. For example, if the risk of expropriating wealth is high due to “bad institutions”, incentives to invest are lacking that otherwise could have promoted economic and technological advances (North 1981, Acemoglu et al. 2001, 2002).

[4] According to Lindert (2004), the burden of educating laborers had more weight than its benefits (preventing crime and sedition) from the landlords’ point of view. The premise was that as educated workers sought better-paid jobs outside the agricultural sector, the political supremacy of the landed elite could be threatened, and their taxes would be raised to subsidize the masses.

delaying the implementation of human capital-promoting institutions. These authors argued that two different elite groups strongly influenced political decision-making about educational reforms: large landowners and industrialists. In contrast to "latifundia" owners – who were typically not interested in tax-financed mass-schooling programs – industrial capitalists preferred a workforce with at least some basic skills because education increased productivity in the industrial sector more than in agricultural production. Thus, concentration of landownership hampered the emergence of institutions such as public schools and child labor regulations, as well as the process of industrialization. This was the case, for example, in some 19th century Latin American countries in which this elite group had particularly strong political power because large estates provided financial and social backing to influence the ruling groups, and the institutions were set up to maintain inequality. In contrast, countries with lower land inequality or countries that had installed land redistribution programs were more often governed by capitalist interests; hence, mass-schooling was frequently extended. Galor et al. (2009) provided some evidence on regional land inequality and school investment in the U.S. from 1880-1940, as well as some qualitative evidence on land reforms in East Asia and Russia. Ramcharan (2010) assessed the impact of educational spending in U.S. counties and states from 1890 to 1930.

According to Wegenast (2009, 2010), plantation owners had economic and *political* reasons to neglect educational expansion, for instance, by impeding the mobilization of rural workers so that a cheap supply of labor was secured and by monopolizing political participation, given the literacy requirements for suffrage.

Another theory exploring this link is the "credit market imperfection approach", which holds that land inequality, usually implying the landlessness of many, results in credit rationing, in which most individuals cannot undertake investments in human capital (Galor and Zeira[5] 1993, Deininger and Squire 1998).[6] Supporting this view and adding a labor market-power component, Deininger (2003) noted that unequal land distribution removes incentives of rural workers to accumulate human capital. When landlords have a monopsony over the labor market and keep all wages at subsistence levels, investing in human capital seems much less rewarding. As an example, the author described the growing gap in the development of literacy rates between 19th century El Salvador and Guatemala on the one hand and Colombia and Costa Rica on the other hand. While in El Salvador and Guatemala a few large landowners held a monopoly over the labor market and people had no incentive to invest in the accumulation of human capital, in Colombia and Costa Rica the coffee boom led to a small landholder coffee economy where literacy rates increased (Deininger 2003, p. 20).

A body of literature that emphasizes the effect of economic or social heterogeneity on the access to public goods, holds that land unequal areas perform worse in terms of public investments.[7] For instance, Banerjee and Iyer (2005) find that in India, areas where proprietary rights in land were historically given to landlords, today invest less in public infrastructure such as primary schooling, higher schools and health centers. They also find accordingly lower human capital outcomes as literacy and health outcomes like infant mortality in this kind of areas.

Hippe and Baten (2012) recently tested and confirmed the link between land inequality and human capital formation by regressing the share of large landholdings

[5] In their study, Galor and Zeira show that in a cross-country comparison, wealth and income inequality has a detrimental effect on subsequent economic growth by hampering investments in human capital.
[6] For an extensive literature review on how inequality affects growth, see also Galor (2011).
[7] One possible reason is that the conflict between stake-holders (here: landlords and the masses) could hamper unanimous policy decisions (on this body of literature, see also Bates 1981 and Easterly and Levine 1997).

(above 100 hectares) on numeracy levels for several hundred regions of ten European countries (see also Cinnirella and Hornung 2011 for Prussia). In contrast to those authors, we use a global sample of Gini coefficients of land inequality during the late 19th century to assess this theory both in industrialized countries and in developing world regions.[8]

Various exogenous factors have been suggested to determine inequality, including initial endowments of geographical conditions, soil quality and surplus labor. Engerman and Sokoloff (1997, 2000, 2002) argued that countries in tropical zones that are suited for the cultivation of cash crops (such as sugar cane, tobacco, coffee, rubber or cocoa) were more likely to develop into unequal societies than countries in temperate zones, which were more suitable for the production of food crops such as wheat, maize or rice, because the production of cash crops was subject to substantial economies of scale (as substantial fixed costs were unavoidable, but fell per unit of production at higher scale). Thus, cash crop production favored large-scale plantations and attracted large inflows of mainly slave labor. Geography has also been alleged to have had a direct productivity-retarding effect through climate and climate-induced diseases in some world regions (Montesquieu 1748, Jones 1981, Sachs and Warner 1997, Diamond 1997, Sachs 2001).[9]

Fertility choice is another potential determinant of human capital accumulation, especially during the 20th century: The famous quantity-quality trade-off might determine the level of educational investment because if a family decides to have more children, the human capital investment per child might be lower (see Becker 1960, Easterlin 1980). However, Clark (2007) recently presented historical

---

[8] Easterly (2007) uses agricultural endowments as an instrument for income inequality as well, but not specifically for land distribution.

[9] However, Easterly and Levine (2003) show in a cross-country study that climatic conditions ("tropics, germs and crops") affect growth only through institutions and not directly.

evidence that this result does not necessarily hold for early societies in which richer individuals had more surviving children (Galor and Moav (2002) proposed a theory of natural selection earlier than Clark).

Human capital clearly has endogenous as well as exogenous components. Exogenous factors (not influenced by economic variables) include, for example, religious educational values (Botticini and Eckstein 2007, Becker and Woessmann 2009). Selective migration (in terms of skills) can affect the human capital development of a country, both positively and negatively (Glaeser et al. 2004). For example, when we speak of a "brain drain" phenomenon, the home country's average human capital is negatively affected by emigration of high skilled workers, while the host country benefits from positively selected immigration (and sometimes vice versa).[10]

Our study contributes to this research stream by providing empirical evidence on the causal link between inequality more than one hundred years ago and recent human capital levels for a large cross-section of countries.

3. Data

The data on income inequality was recently compiled by Van Zanden, Baten, Földvari and Van Leeuwen (2012) and Blum and Baten (2011). This global dataset documents within-country inequality for a large number of countries, employing anthropometric inequality measures as well as income distributions. This effort allows us to cover income inequality in a substantial number of countries.

How can we estimate land inequality? Unfortunately, the number of available estimates of land inequality for the 19th century is much more limited than even the

[10] Stolz and Baten (2012) find the opposite – arithmetic "brain gain" for the home countries which "lost" many their less skilled inhabitants -- for a number of countries during the 19th century era of mass migration.

number of income Ginis. Frankema (2010) presented 199 Gini coefficients of inequality in land holdings[11] for 111 countries in the 20th century. His tables are based on the World Census of Agriculture published by the International Institute of Agriculture (IIA) and its subsequent incarnation, the Food and Agriculture Organization (FAO). Before the 1920s, only 27 land Ginis are available, and for the 1960s, 63 are available. In examining those countries for which two or more land Ginis are available, it becomes clear how little land inequality has changed, especially in developing and threshold economies. For example, Argentina's Gini coefficient of land inequality was 0.803 in the 1910s. In the 1940s, it was 0.806, and in the 1960s and 1980s, it was around 0.81.

Over nearly a century, there was barely any change. The same applies to Brazil in the 1920s through the 1960s, despite land reform initiatives during the 1930s that did not have much effect. In Chile, there was a small increase from an already high inequality level of 83.7 in the 1920s to 86.5 in the 1960s, and the results are similar in Uruguay (77.5 to 79.1 in the 1930s and 1960s). Because we know that all inequality measures contain some measurement error, the hypothesis of no change can most likely not be rejected. The stability of land inequality can be observed for a large number of countries. Given the overwhelming stability of land inequality values, we estimate a cross-sectional dataset for the late 19th century by assuming that the earliest value of land inequality is a good proxy for the initial conditions in each country if there was no land reform.[12]

---

[11] The size distribution of the land holdings reflects the access to land as a production factor, rather than only ownership. It is usually more equally distributed than land ownerships. However, this data includes land holdings used for all purposes – arable land, land used for permanent crops, pastures or woods – and do not take into consideration the quality or the location of the land (Frankema 2010).

[12] Our *early land inequality* variable contains the land Gini values of the earliest year available for each country from Frankema's (2010) dataset.

While land inequality typically did not change much over several decades, change in land inequality is likely if a land reform took place in the early 20th century.[13] Because we are concerned that land reforms implemented between 1890 and the earliest year for which we have information on land concentration could have had a substantial effect on land distribution, we need to control for the average effect of land reform on the land Gini coefficient. To test whether land reforms might explain the change in land inequality significantly, we collected evidence on the implementation and success of land reforms (listed in Appendix B). With this information, we assessed the average reduction effect of a land reform on the change of the Gini coefficient of land concentration in a regression (Table II).[14] The coefficient of the explanatory variable 'land reform' (a dummy variable that takes the value of 1 if a land reform was carried out in the previous period) in both specifications implies that land reform will reduce the Gini coefficient in the following period by six Gini points. We estimated a second specification in which we controlled for "Western Europe and offshoots", a variable that includes developed economies in which land concentration declined during the twentieth century, typically without land reforms. The results were robust when including this additional variable. According to the results, we correct our values of early land Ginis – using those that are the closest to 1890 contained in the Frankema data set – by adding six points when land reforms were performed successfully. This tactic allows us to back-project early land inequality estimates to the level around 1890. In Appendix C, we list the sources from which we derive information on land reforms for the various countries. Of course, the ideal variable would have measured the extent of land

[13] One could imagine that also a change in crop production might alter the distribution of land holdings. However, for the time period and the countries observed here, crop production hardly changed. All Latin American countries, for example, had already specialized in cash crop production.

[14] We included only those countries for which Gini coefficients of land holdings for more than one point in time were available.

reforms, which our binary variable does not measure. However, adjusting for the existence of land reforms already improves the measurement of our land inequality variable substantially.

We should consider the question, whether changes in population or in crop production did impact on the change in land inequality, especially between the 1890s and the 1950/60s. Clearly, substantial changes in the crop composition or the share of land used for livestock production could have had an influence on landholding patterns. Especially a switch from small-scale subsistence farming to cash crop commodity production in large plantations could lead to higher land inequality. Expropriation of native small land holders to introduce large scale primary goods production was a typical effect of the colonial 'land-grab'. For example, in Malaysia and Sri Lanka colonialists produced rubber on a large scale, a cause of the high land Gini coefficient in the second half of the 20th century (Frankema 2010). Moreover, the adoption of cattle-raising activities for leather production or beef export could have lead to a higher concentration of land holdings. We will thus have a careful look at the changes in the composition of crops grown and the time of adoption of cattle-raising activities in former colonies included in our data set. We will then draw the link to the evolution of land inequality.

In Latin America already the Spanish colonizers of the 16th to 18th century introduced cash crops and livestock production on large farms (Bulmer-Thomas 1994, p. 9). This was the case for the cultivation of sugar in Cuba, the Caribbean and Brazil, which began in the 16th century. Production and export of tropical crops such as coffee[15], cocoa, tobacco and bananas boomed in Latin America since the middle of the 19th century, to satisfy a rising demand in Europe and North America after the

[15] However, coffee seems to be an exception of a crop that is more effective to be planted in small plantations. There seem to be no scale economies and cultivation is labor intensive, thus more profitable for smallholders (Nugent and Robinson 2010)

Industrial Revolution (Bulmer-Thomas 1994, p. 36). Even rubber production in Brazil, a commodity that was exploited later in Asia, reached its peak already in the middle of the 19th century (Resor 1977). Although small variations in the composition of the cash crop cultivation may have occurred later, mostly landholding patterns were already established by 1890. Hence the Latin American development of the early 20th century does not affect our back-projection of land inequality much.

What about African and Asian colonies? In some more recent colonies, the British rule between the end of the 19th century and the first half of the 20th century might have had an increasing impact on land inequality. Frankema and Waijenburg (2012) described how British colonies in sub-Saharan Africa ("non-settler colonies") mostly adopted the existing land tenure system. However, commodities were cultivated by small native landholders before and after colonization. In French colonies of the same region land expropriations were also rare. In other sub-Saharan colonies, such as Sierra Leone, the native population did not switch to commercial agriculture. Land continued to be held by many smallholding subsistent farmers (Frankema 2010). Yet in South Africa – a "settler colony" – the native population was expropriated in large scale and land was redistributed to the white population at the turn of the century, resulting in a more unequal land distribution (Binswanger and Deininger 1993). The situation was similar in Malaysia, where British colonizers introduced large rubber plantations at the end of the 19th century and land inequality probably intensified (Frankema 2010). In the regression analyses below, however, South Africa, Malaysia and Sri Lanka do not play a role because early numeracy could not be estimated. One missing value in one variable causes the observation to be dropped. In sum, crop change probably influenced our back projection of land inequality mildly but not substantially. The same is probably true for immigration and population change.[16]

Comparing land and income inequality around 1890, we find that a general correlation exists (Figure IV). Chile and Peru (“cl” and “pe”) both had high land and income inequalities, whereas land and income inequalities were low in Finland and Estonia (“fi” and “ee”). On the other hand, some countries had high land inequality but not high income inequality (for example, Egypt, “eg”, and Australia, “au”), whereas other countries deviated to the other side of the regression line. An example of the latter case of low land inequality and elevated income inequality is Sweden in 1890. In this period, the country experienced its industrialization boom, which temporarily shifted income inequality. Inequality was not structural and permanent in Sweden and moved downward later. Looking at modern math and science skill levels, Egypt (high land and low income inequality) had low values, whereas Australia was characterized by high values. Therefore, we need to perform a regression analysis to cope with potential outliers.

Our equation for human capital determinants also includes the ABCC-Index values for 1820 – the earliest year for which we obtained a sufficiently large set of country data – to assess the path-dependence of education reflected in numeracy skills. A considerable number of recent studies have used this innovative proxy of

---

[16] Did changes of population over time, for example due to high immigration, affect land inequality? We find that land inequality was not substantially affected by migration. First, a large share of migrants went to urban centers (for example in Argentina) or to work as employees in already existent crop plantations. If immigrants arriving had no access to land in the destination country, this population increase would not affect our land inequality measure, because it does not account for the landless population (only for land holders). Second, those immigrants who were brought to settle in rural communities as small landowners in Chile and Argentina obtained land pieces of middle size, and overall distribution was not affected. Argentina was one of the most immigration intensive countries between 1870 and 1930. However, the early land Gini measures for Argentina by Frankema (2010) show that land inequality remained stable. Summerhill (2010) provides further evidence for an unchanged land Gini in Sao Paolo state (Brazil) between 1905 and 1995 (0.688 for 1905 and 0.677 for 1995).

basic numeracy, which was introduced by A'Hearn, Crayen and Baten (2009) and is based on the "age-heaping" technique. The age-heaping phenomenon applies to historical populations (as well as people in the poorest countries today) when a substantial share of the people are not able to state their exact age and therefore reported a rounded age, such as "I am 30", when they are in fact 29 or 31. The ABCC-Index reflects the share of people who were able to state an exact age (or the degree to which the distribution of age statements approaches an equal distribution). Crayen and Baten (2010) showed that this proxy of early human capital has a strong correlation with other measures such as literacy and schooling. The calculation of the ABCC Index is shown here as a derivation of the Whipple Index:[17]

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

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

As mentioned in the introduction, we argue that countries that made early investments in numerical education (and perhaps cultures that promoted numerical skills) entered a path-dependency of human capital-intensive industries, including skill-intensive agriculture (such as dairy farming in Denmark) and services. After these skill-intensive industries became important during the second Industrial Revolution of the 19th century, the following generations again invested in skill formation, both through formal schooling and cognitive training within families. For example, they used basic numerical training devices such as calendars and games that required basic numerical skills (Baten et al. 2009).

[17] A'Hearn, Baten and Crayen (2009) found that this index is the only one that fulfils the desired properties of scale independence (a linear response to the degree of heaping), and that it ranks samples with different degrees of heaping reliably. For an easier interpretation, A'Hearn et al. (2009) introduced the ABCC index. It is a simple linear transformation of the Whipple index and yields an estimate of the share of individuals who correctly report their age

Crayen and Baten (2010) and Manzel et al. (2012) -- as well as recently Hippe and Baten (2012) have already carefully discussed the evidence on early numeracy. However, to evaluate its contribution to modern differences of math and science skills, a short summary of the institional sources of the evidence and the discussion of the degree of (sometimes not entirely avoidable) bias can be quite helpful. For Western Europe, Crayen and Baten could rely on census material which reported the age distribution repeatedly in many different censuses. For the UK and the United States, even individual level census data was available. For Eastern Europe, they could rely on the Russian Imperial Census of 1897, which was taken in quite disaggregated regional units, and hence could be re-aggregated in order to make the evidence comparable with modern data and math and science skills which are organized by modern state boundaries. In other words, for the numeracy of Armenia in 1820 only those districts were included which fall into the modern boundaries of Armenia. Similar procedures could be applied to the Austro-Hungarian (Habsburg) Empire evidence.

For Asia and Africa, the basic sources of evidence were more challenging. For example, for China, Baten et al. (2010) used a large variety of different data sets in order to arrive at the most likely estimate for the early 19th century. For Southeast Asia, census records were unavailable before the 1870s and had to be replaced by court records (which required adjustments due to the fact that persons appearing in court mostly for property crimes are not representative for the whole population). In contrast, for India, census material of quite high quality was available thanks to the intensive colonial administration of the British. The evidence on Latin America for the 19th century is quite comprehensive, and Manzel et al (2012) took care that any unrepresentativeness of census collection activity was taken into account (see Appendix G on the sources of the ABCC 1820 estimates in detail).

Figure I illustrates the strong correlation between the 1820s measure of numerical skills and the Hanushek-Woessmann measure for the late 20th century. The dependent variable *cognitive* is our measure for the cognitive skills of human capital. This output measure of human capital reflects the knowledge and abilities that are most favorable for subsequent success. Hanushek and Woessmann (2012a) constructed this variable from test scores in mathematics and science for 77 countries between 1964 and 2003. More specifically, they calculated the simple average of all math and science scores on International Student Achievement Tests conducted by the Organization for Economic Co-Operation and Development (OECD) and the International Association for the Evaluation of Educational Achievement (IEA) during that period. With linear regressions, they found that by adding cognitive skills to a growth model with school attainment as a dependent variable, the model explains three quarters of the variance in growth rates (instead of one quarter if only school years are included). Further, the coefficient for school years turns statistically insignificant in the presence of cognitive skills. When testing the correlation between test score improvements and growth rates by world regions – conditional on the initial level of real GDP per capita in 1960 – they find an R^2 of 0.98 (see Figure II, albeit with N=5). To control for possible endogeneity of cognitive skills, which might be present if the factors leading to growth are also related to high cognitive skills and have been omitted from the estimation, they instrument cognitive skills with institutional measures of schooling and confirm that schooling-induced differences in cognitive skills are significantly related to economic growth.

This measure of human capital is said to best explain economic growth because, unlike other measurements such as schooling, it controls for differences in the quality of education across countries. A second test for the impact of the cognitive skills measure on economic performance was carried out in a case study on US-migrants. The authors looked at the performance in the labor market of US-migrants educated in the US and US-migrants educated at home, holding constant cultural and other country-of-origin fixed effects. The study again confirmed the growth-inducing effect of the *cognitive skills* measurement.[18]

Further variables included in our dataset are as follows: initial GDP in 1910, population density (logarithm), a measure of institutional quality, fertility in 1950 (the earliest estimates available for a sufficient number of countries), the share of the population living in the tropical zone in 1995, ethnic fractionalization, and a measure of physical capital that could proxy industrial development, constructed by Enflo and Baten (2006). Population density was included because visiting schools is less costly in countries with high population density.

4. Base Regression Results

Our model for the estimation of human capital has the following form:

$$Cognitive_i = \alpha + \beta_1\ land\ ineq_i + \beta_2\ income\ ineq_i + \beta_3\ early\ ABCC_i + \gamma X_i + \varepsilon_i$$

The dependent variable *cognitive* is our measure for cognitive skills. This output measure of human capital best reflects the knowledge and abilities that are favorable for subsequent success. Hanushek and Woessmann (2012a) constructed this variable from test scores in mathematics and science for 50 countries between 1964 and 2003.

[18] The finding that numerical cognitive skills are most conductive to economic growth further corroborates the argument of Schumpeter for the use of “numeracy” as an important proxy of human capital (see also Baten 2010).

We test both income inequality around 1890 and early land inequality as determinants of cognitive skills. We add numerical abilities ("early ABCC") to assess the persistence of numerical skills. This measure also proxies to a certain extent the GMV factor of industrialists promoting basic education because countries that were numerate in 1820 also tended to be industrial.

In the econometric specification above, the additional vector of explanatory variables *X* captures other factors that could have an effect on human capital for all countries *i*: initial GDP in 1910, population density (logarithm), institutional quality, fertility in 1950, the share of the population living in the tropical zone in 1995, ethnic fractionalization and physical capital in 1925. Table I provides an overview of the summary statistics for the variables included in the model. The Skewness/Kurtosis tests for normality indicate that physical capital in 1925 and population density were right skewed; hence, we chose to take the natural logarithm of those variables.

Based on the basic model of human capital introduced in the last section, we perform cross-country analyses of the effect of inequality and early human capital on cognitive skills (Table III). A large part of the variation in cognitive skills in the period 1964-2003 – as much as 54 percent – can be explained by early numeracy from the ABCC-Index in 1820 (see specification (1)).

In specifications (2) to (7), we assess the additional long-run impact of the Gini coefficient of early land inequality in OLS regressions for all the countries for which we had data. Land inequality always has a significantly negative impact on the dependent variable. When examining the coefficients, we can observe that a one standard deviation increase in land concentration leads to a 0.35 standard deviation decrease in human capital. We can best illustrate the impact of the land Gini coefficient with an example: a one standard deviation increase in land inequality reduces the level of cognitive skills from Serbia and Montenegro to the level of India. The same is true if we instead include income inequality (columns 5 and 7). In the case of income inequality, a one standard deviation increase of the Gini leads to a 0.26 to 0.31 standard deviation decrease of cognitive skills.

We also include a large set of other variables that could exert an influence besides inequality and early human capital. First, we control for early income (GDP in 1910) because the differences in early human capital could arise from the fact that richer economies spend more on education. However, all except one of the income coefficients are insignificant, and they do not lead to insignificance of early numeracy or inequality. We further insert controls for institutions, geography, population density, fertility, ethnic fractionalization and physical capital, but these variables are often insignificant (except tropic share).

Moreover, if we run a horse race between the two concepts of inequality, land inequality has the greater explanatory power (columns 12-14). Land inequality remains significant, whereas income inequality is no longer statistically significant. This result is consistent with Easterly's (2007) argument that land inequality is an important component of "structural inequality" because income inequality can temporarily rise, especially in growth and transformation processes. Structural inequality is mostly caused by the long-run effect of colonial land grabbing. Structural inequality has a much more depressing effect on human capital and development because educational efforts may not be rewarded in a situation of permanent, structural inequality.

The negative effect of land inequality, with other factors held equal, ranges from minus 0.81 to minus 1.60 and remains strongly significant through all specifications. These results confirm the GMV theory of a persistent negative effect of land inequality and early numeracy on the average quality of education during the second half of the 20th century.

In all specifications of Table III we include robust standard errors in order to cope with heteroskedasticity, because, as is common in this kind of dataset, the variance of the residuals is not identically distributed across all fitted values. We decided to cluster the standard errors by region (running the `stata`-option, `vce(cluster)` after running the regressions with `reg`), given that standard errors could be correlated within world regions.

We additionally use alternative techniques like regressions with weighted least squares by population numbers in the countries assessed (columns 10 and 11) and robust against outliers together with population weights (columns 8 and 9). [19] Assigning a greater weight to more populated countries does not alter the significance of the coefficients for land inequality and their sizes stay almost unchanged. However, when running a robust regression (which gives less weight to outliers), the estimated effect of land inequality remains statistically significant, but the coefficient of income inequality turns insignificant. As a further robustness check, we run all specifications shown in Table III, this time always using the same set of countries (number of observations equal to 30). The results are very similar and are shown in the Appendix Table I.

5. Instrumental Variable Models

Next, we need to consider endogeneity. The results of the ordinary least squares regressions could be affected by reverse causality. For example, apart from the

[19] For running regressions with weighted least squares by population, we use the Stata option `[aweight=popsqrt]`. We used the R-statistics- package command `lmRob` to compute robust regressions and the option `weights=popsqrt` to include weights by countries' population.

direction of causation running from the inequality of land via the political economy of landlords who oppose primary schooling (and the tax burden that comes with it), one can also imagine that in the long run, regions with relatively good education even for small landholders could reach a lower level of inequality of land distribution as those peasants are able to buy more and more land. The peasants might also influence political activity in favor of land reforms, as Cinnirella and Hornung (2011) have noted for the historical German kingdom of Prussia. On the other hand, educated small landholders might decide to sell their plots to obtain the return on their human capital investment in other industries – in nearby cities, for example. Instrumental variable estimation allows us to circumvent these issues of endogeneity.

We base our first stage of the two stage-least-square estimate on the following equation:

$$Landineq_i = \beta_1 + \beta_2\, Sugar/wheat\text{-}rice_i + \beta_3\, lowpop1500\text{-}Southern_i + \beta_4\, X_i + \varepsilon_i$$

where *sugar/wheat-rice* is an Easterly-type instrumental variable of relative soil suitability, *lowpop1500 –Southern* is the interaction of inverse population density in 1500 with "Southerness" of latitude and *X* is a vector of other exogenous variables.

Easterly (2007) and GMV recently advocated for the use of climatic, geological and similar variables that allow types of agriculture that correlate either with higher or lower efficient sizes of scale. Sugar plantation and cattle-raising for hide and beef exports are clear examples of agricultural production types in large-scale economies. On the other hand, wheat and rice production is already highly productive on much smaller farm units, as has been amply demonstrated in the agricultural economics literature. The specialization of a country on the cultivation of large scale cash crops is negatively associated to the distribution of landholdings, whereas food crops such as wheat or rice are not scale intensive and were historically planted in smallholdings. Sugar, wheat and rice production requires relatively clear-cut climatic and soil characteristics. Based on this premise, the UN- Food and Agriculture Organization quantified the share of a country's area that is suitable for the production of each of those crops. In the spirit of Easterly's (2007) we use the ratio of the share of the land suitable for the cultivation of the "inequality crop" (sugar) to the share of the land suitable for the cultivation of the "equality crops" (wheat and rice).

The cattle-raising activity associated with inequality is unfortunately less concentrated on specific soils (Figure III). However, there is a clear correlation with southern locations and initially low population density. Argentina became the prototype of this type of land use because the indigenous Indian population on its great plains was always very thin, and in contrast to the United States, European immigration was relatively limited in numbers until the late 19^{th} century. Therefore, the U.S. type of grain farming associated with lower inequality developed on a smaller scale. Thus, we instrument this type of agricultural inequality in southern latitudes in interaction with population density in 1500. In Figure III, we show the level of land inequality around 1890. It is clear that the sparsely populated countries in the most southern locations showed the largest land inequality. Apart from Argentina and Australia, South Africa, New Zealand and Uruguay also had very unequal distributions following the colonial land grab (Eastwood et al. 2010). In contrast, the equality crops of wheat and rice led to low land inequality in East Asia and some of the European countries.

The advantage of the ratio between the climatic and geological suitability ratio of sugar and wheat/rice is its intrinsically exogenous nature, whereas actual crop use could be influenced by educational levels. Similarly, population density around 1500 is a very popular instrumental variable because it captures human development in a very early time period (Acemoglu et al. 2002). The autocorrelation of early population densities and those of the 20th century is quite limited due to the unequal population increase in some world regions and massive migration movements. Similarly, southern latitude is most likely very exogenous.

The results of the two-stage-least-squares regressions confirm that both variables fulfill the necessary requirements to be good measures of land inequality: First, they mostly correlate with land inequality, as is documented by the "first stage" section of Table IV. The F-Test is mostly below 10, requiring a Limited Information Maximum Likelihood (LIML) estimation, but it is still in the range of 5, which means that the instruments have some strength (see Stock and Yogo 2005). In the case of Model 1, in which we use both instrumental variables, the F-Statistic is larger than 10. Hence, this model should be preferred from the criterion "strength of instruments". However, because the instrumental variable component of "Southerness" (southern latitude) might be correlated with other factors, we find it important to show that the suitability variables alone yield similar results in Models (2) to (6). Second, we argue that the instruments influence the dependent variable only through the potentially endogenous variable, land inequality (see Figure V and the discussion on the exclusion restriction in Appendix F). The p-value of the Durbin-Wu-Hausman test indicates only once (column 2) that the IV approach is necessary, as the results are significantly different from OLS estimates. As a result, the significant impact of land inequality and early numeracy remains a consistent determinant of today's human capital. Most of the other controls are insignificant, except for early numeracy, which is always positive and significant, and in one specification the coefficient of income per capita in 1910 becomes significant. In Appendix A, we show the results of income inequality as a determinant of human capital in the long run. When using the same instruments for income inequality, TSLS regression results become far less robust and usually insignificant.

6. Conclusion

The combined long-run effects of land inequality and human capital path-dependence were assessed in this article for the first time. We found that initial income inequality in a cross-section of countries might play a major role in determining the development of human capital formation in the long run on the first view. However, a more important predictor of recent human capital formation was early land concentration. This finding supports the GMV (2009) theory that the behavior of landowners is crucial to development relative to the influence of industrialists. The industrialist component of the model was to some extent proxied by early numeracy because the top performers in early numeracy were also quite industrial. We tested the robustness of our results by including various control variables and running a series of alternative regression models. We find that this influence is actually causal, using an Instrumental Variable (IV) approach with geological, climatic and other variables that are intrinsically exogenous.

These results are consistent with Easterly's (2007) argument that land inequality is an important component of "structural inequality" because income inequality can temporarily rise, especially in growth and transformation processes. Structural inequality is mostly caused by the long-run effects of colonial land grabbing. Structural inequality has a much more depressing effect on human capital and development because it seems that educational efforts might not be rewarded in a situation of permanent, structural inequality.

A second major contribution of our study was to assess the persistence of numerical cognitive skills, which are an important component of the Hanushek-Woessmann measurement. Our results confirm that countries with early investments in numerical education (and perhaps cultures that promote numerical skills) entered a path-dependency of human capital-intensive industries, including skill-intensive agriculture and industries. Early numeracy explained a considerable portion of recent cognitive skills.

Joerg Baten, University of Tübingen, and Dácil Juif, University of Tübingen

Appendix A. Regressions of Cognitive Abilities on Income Inequality and other Factors: Instrumental Variable Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation	TSLS	TSLS	LIML	TSLS	TSLS	TSLS
	Sugar/ (WheatRice) Def. 1 + LowPop* Southern	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3
Second Stage						
Early numeracy	1.75***	1.80***	1.40*	1.25**	1.30***	1.35**
	(0.00)	(0.00)	(0.06)	(0.03)	(0.00)	(0.01)
Early income inequality (around 1890)	-5.04	-7.06**	-11.13	-7.97	-9.37	-10.11
	(0.17)	(0.04)	(0.46)	(0.45)	(0.13)	(0.38)
GDP/c 1910			0.47	0.36	0.43	0.41
			(0.60)	(0.61)	(0.14)	(0.57)
Tropic share			0.23	0.19	0.06	0.21
			(0.60)	(0.56)	(0.86)	(0.57)
Institutional quality			0.01		0.01	0.01
			(0.43)		(0.60)	(0.35)
Fertility			-0.02	-0.06		-0.03
			(0.94)	(0.73)		(0.87)
Constant	5.24***	6.00***	4.37	4.29	4.01***	4.48
	(0.00)	(0.00)	(0.17)	(0.11)	(0.01)	(0.12)
Observations	37	39	33	34	35	33
Second Stage R-squared	0.61	0.38			0.11	0.02
First Stage						
Sugar/(WheatRice) Def. 1	0.21					
	(0.13)					
LowPop*Southern	0.00					
	(0.94)					
Sugar/(WheatRice) Def. 2		0.20*	0.07			
		(0.07)	(0.57)			
Sugar/(WheatRice) Def. 3				0.72	0.13	0.08
				(0.48)	(0.18)	(0.49)
F-Statistic	1.31	5.02	0.32	0.52	1.91	0.48
Durbin-Wu-Hausman p-val.	0.40	0.07	0.33	0.50	0.14	0.33
Sargan p-value	0.88					

Dependent variable (of the second stage): Cognitive Skills. Control variables: ABCC 1820, tropic share, population density (ln), institutional quality, GDP per capita in 1910 (ln) and physical capital (ln). Instrument "Sugar/(WheatRice) Def. 1" includes the share of the land that is very suitable and suitable for the plantation of the three crops (sugar, wheat and rice), Instrument "Sugar/(WheatRice) Def. 2" the share being very suitable, suitable and moderately suitable and, last, "Sugar/(WheatRice) Def. 3" includes the share of the land very suitable, suitable, moderately suitable *and* marginally suitable for crop plantation. Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Appendix B: Land reforms with some potential impact on land inequality.
Abbreviations: See Appendix C

co	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990
al							1	1				1
am												
ar												
at												
au												1
be												
bg	1				1							
bh												
bo								1				1
br												1
bw												
ca	1											
ch												
cl									1	1		
cn							1	1		1		
co						1			1			
cs												
cu									1			
cy												
cz					1		1					
de					1							
dk					1							
dz										1		
ec									1	1		
ee				1								
eg								1				
es						1						
et										1		
fi				1								
fr												
gh												
gr												
gt								1				
hk												
hn										1		
hr					1		1					
hu							1	1				1
id									1			
ie	1											
il												

co	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990
in										1		
iq										1		
ir									1	1		
is												
it												
jo												
jp							1					
ke									1			
kr							1					
kw												
la										1		
lb												
li					1							
lk										1		
lt												
lu												
lv				1								
ma												
md												
mk												
mo												
mx				1		1	1					
my												
na												1
ng												
ni											1	
nl												
no					1							
nz												
pe								1	1		1	
ph									1		1	
pl					1		1					
pr							1	1				
ps												
pt												
ro					1		1					
ru				1								1
sa												

Appendix C: Sources for land reforms

Co. abbr.	Country	Source
AL	Albania 1946, 1950, 1991	Statistical Yearbook of the Popular Republic of Albania 1963. Tirana: Department of Statistics, 1964 Lerman, Z., Csaki, C., and Feder, G. (2004). Agriculture in Transition: Land Policies and Evolving Farm Structures in Post-Soviet Countries. Landham, MD: Lexington Books.
DZ	Algeria 1971	http://www.economywatch.com/agrarian/land/algeria.html
AU	Australia 1993	United Nations Economic and Social Council (2007). Indigenous Land Reform in Australia.
BO	Bolivia 1953, 1996	Tai, H. C. (1974). Land Reform and Politics: A comparative Analysis. Berkeley: University of California Press. Baranyi, S., Deere, C. D., and Morales, M. (2004). Land and Development in Latin America: Openings for Policy Research. National Library of Canada Cataloguing in Publication.
BG	Bulgaria, end of 19th c., 1920-1923	Cochrane, N. J. (1993): Central European Agrarian Reforms in a Historical Perspective. *American Journal of Agricultural Economics*, 75(3) : 851-856
BR	Brazil 1990s	Deininger, K. (1999). Making Negotiated Land Reform Work: Initial Experience from Colombia, Brazil and South Africa. *World Development*, 27: 651-672.
CA	Canada 1873	Smith, P. J. (1987). The Ideological Origins of Canadian Confederation. *Canadian Journal of Political Science*: 3-30
CL	Chile 1964-1973	Collier, S., and Sater, W. F. (1996). A History of Chile: 1808-1994. Cambridge: Cambridge University Press.
CN	China 1948, 1950s, 1970s	Twitchett, D., Fairbank, J. K., and MacFarquhar, R. (1992). The Cambridge history of China. Cambridge: Cambridge University Press.
CO	Colombia 1936, 1958-1970	Deininger, K. (1999). Making Negotiated Land Reform Work: Initial Experience from Colombia, Brazil and South Africa. *World Development*, 27: 651-672.
CS	Serbia 1919, 1945	Z. Lerman, C. Csaki, and G. Feder (2004). Agriculture in Transition: Land Policies and Evolving Farm Structures in Post-Soviet Countries. Lanham, MD: Lexington Books.
CU	Cuba 1959-1963	Kellner, D. (1988). Ernesto “Che” Guevara - World Leaders Past and Present. Chelsea House Publishers.
CZ	Czech Rep. 1919, 1945, 1948	Cornwall, M. (1997). 'National Reparation'?: The Czech Land Reform and the Sudeten Germans 1918-38. *The Slavonic and East European Review*, 75(2): 259-280.
DE	Germany	Brockhaus Encyclopedia (1929), 15th edition, Article “Bodenreform”, Leipzig: Brockhaus.
DK	Denmark 1919	Encyclopedia Britannica [http://www.britannica.com/]
EC	Ecuador 1964, 1973	Blankstein, C.S., and Zuvekas, C. (1973). Agrarian Reform in Ecuador: An Evaluation of past Efforts and the Development of a New Approach. *Economic Development and Cultural Change*, 22(1): 73-94
EG	Egypt 1952	Abdel Malek, A.(1968). Egypt: Military Society. New York: Random House.
ES	Spain 1936-1939	Bernecker, W. L. (1991). Krieg in Spanien 1936-1939. Darmstadt: Wissenschaftliche Buchgesellschaft.
EE	Estonia 1919	Meyers, W. H., and Kazlauskiene, N. (2005). Land reform in Estonia, Latvia, and Lithuania: A comparative analysis. In: Land Reform in the Former Soviet Union and Eastern Europe. Edited by Wegren, S. K. London: Routledge.
FI	Finland 1922	Encyclopedia Britannica [http://www.britannica.com/]
GT	Guatemala 1944-1954	Gleijeses, P. (1989). The Agrarian Reform of Jacobo Arbenz. *Journal of Latin American Studies*, 21(3): 453-480.
HN	Honduras 1973-1977	http://www.fao.org/righttofood/KC/downloads/vl/docs/AH265.pdf
HU	Hungary 1945	Wegren, S. K. (2005). Land Reform in the Former Soviet Union and Eastern Europe. London: Routledge.
ID	Indonesia 1960	A rather ineffective landreform: http://countrystudies.us/indonesia/65.htm

IN	India 1977	Deininger, K. (2003). Land Policies for Growth and Poverty Reduction. Oxford: World Bank and Oxford University Press.
IR	Iran 1962-1971	Amid, M. (1990). Agriculture, Poverty and Reform in Iran. London: Routledge.
IQ	Iraq 1970	Rodinson, M. (1979). Marxism and the Muslim world. Zed Press.
JP	Japan 1947	Kawagoe, T. (1999). Agricultural Land Reform in Postwar Japan: Experiences and Issues. World Bank Policy Research Working Paper.
KE	Kenya 1963	Kanyinga, K. (2009): Land Redistribution in Kenya, In: *Agricultural Land Redistribution: Toward Greater Concesus* (Binswanger-Mkhize, H. et al., ed.). Washington: The World Bank Press.
KR	Korea, Rep. 1945-1950	Jeon, Y.D., and Kim, Y.Y. (2000):. Land Reform, Income Redistribution, and Agricultural Production in Korea. *Economic Development and Cultural Change*, 48(2): 253-68.
LA	Laos 1975	Stuart-Fox, M: (1997). A History of Laos. Cambridge: Cambridge University Press.
LK	Sri Lanka 1972-1975	Abt Associates (1999). *The Land Tenure System in Sri Lanka*. Bethesda, Md.
LV	Latvia 1918-1919	Meyers, W. H. and Kazlauskiene, N. (2005). Land reform in Estonia, Latvia, and Lithuania: A comparative analysis. In: *Land Reform in the Former Soviet Union and Eastern Europe* (Wegren, S. K., ed.). London: Routledge.
MX	Mexico 1910, 1934	Dunn, M. (2000). Privatization, Land Reform, and Property Rights: The Mexican Experience. *Constitutional Political Economy*, 11
NA	Namibia 1990	Weidlich, B. (2010). Land ministry tests new farm acquisition model. *The Namibian*. Tapscott, C. (1994). Land reform in Namibia: Why not?. *Southern Africa Report*, 9(3): 12
NI	Nicaragua 1979-1990	Deininger, K., Zegarra, E., and Lavadenz, I. (2003). Determinants and Impacts of Rural Land Market Activity: Evidence from Nicaragua. *World Development, Elsevier*, 31(8): 1385-1404.
NO	Norway	Norway (1979). Lov om jordskifte o.a. (jordskifteloven). Oslo.
PE	Peru 1950s, 1968, 1985-1988	Deininger, K. (2003). Land Policies for Growth and Poverty Reduction. World Bank and Oxford University Press. Zegarra Méndez, E. (1999). El Mercado de Tierras Rurales en el Perú. Productive Development Series no. 63. Santiago de Chile: Economic Commission for Latin America and Caribbean.
PH	Philippines 1960, 1985	Borras, Saturnino M. Jr. (2006): The Philippine Land Reform in Comparative Perspective: Some conceptual and Methodological Implications. Journal of Agrarian Change. 6,(1): 69-101
PL	Poland 1919-1928, 1944	Swinnen, J. F. M. (1999). The political economy of land reform choices in Central and Eastern Europe. *Economics of Transition*, 7: 637-664. Wolf, N. (2005). Path dependent border effects: the case of Poland's reunification (1918-1939). *Explorations in Economic History*, 42: 414-438
RO	Romania 1921, 1945, 1991	Sabates, R. (2005). Cooperation in the Romanian Countryside: An Insight into Post-Soviet Agriculture. New York: Lexington Books.
RU	Russian Fed. 1906-1917,	Service, R. (1998): *A history of twentieth-century Russia*. Cambridge, Massachusetts: Harvard University Press. http://www.nytimes.com/2002/07/09/opinion/russian-land-reform.html
SN	Senegal 1964	Ubink, J. M., Hoekema, A.J., and Assies, W.J. (2009). Legalising Land Rights: Local Practices, State Respones and Tenure Security in Africa, Asia and Latin America. Leiden University Press.
SV	El Salvador 1980s	Arnson, C. J. (2003). El Salvador's Democratic Transition Ten Years After the Peace Accord. Washington, D.C: Woodrow Wilson International Center for Scholars.
SY	Syria 1958-1961	Heydemann , S. (1999). Authoritarianism in Syria: Institutions and Social Conflict 1946-1970. Ithaca: Cornell University Press.
TN	Tunisia 1963	http://perspective.usherbrooke.ca/bilan/servlet/BMEve?codeEve=916
TW	Taiwan (Chinese Taipei) 1950s	Yager, J.A. (1988). Transforming Agriculture in Taiwan: The Experience of the Joint Commission on Rural Reconstruction. Ithaca: Cornell University Press.
VN	Viet Nam 1953-1956	Fall, B. (1975). The Viet Minh Regime, Government and Administration in the Democratic Republic of Vietnam. Connecticut: Greenwood Press.

ZA	South Africa 1994	Deininger, K. (1999). Making Negotiated Land Reform Work: Initial Experience from Colombia, Brazil and South Africa. *World Development*, 27: 651-672
ZW	Zimbabwe	Moyo, S. (2000). Land Reform Under Structural Adjustment in Zimbabwe: Land Use Change in the Mashonaland Provinces. Uppsala: Nordiska Afrikainstitutet.

Appendix D: Data Sources

Variable name	Source
Cognitive skills (average scores of international tests in math and science 1964-2003)	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.
Income inequality (Gini 1890)	- Baten, J., Földvari, P., van Leeuwen, B., and van Zanden, J.L. (2012). World Income Inequality 1820-2000. Working Paper. - Blum, M., and Baten, J. (2011). Anthropometric within-country Inequality and the Estimation of Skill Premia with Anthropometric Indicators. *Review of Economics - Jahrbuch fuer Wirtschaftswissenschaften*, 62(2): 107-138.
Early land inequality (Gini coefficient)	[adapted from] Frankema, E. (2010). The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions?. *Economic History Review*, 63(2): 418-451.
Early numeracy (abcc index 1820)	Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History*, 47 (1): 82–99. [Complemented with new evidence based on Manzel, K., Baten, J., Stolz, Y. (2012). Convergence and divergence of numeracy: the development of age heaping in Latin America, 17th to 20th centuries. *Economic History Review*, 65 (3): 932–960.]. For details see Appendix G.
Tropic share (the share of a country's population living in a tropical zone, 1995)	Center for international development at Harvard University, General Measures of Geography [http://www.cid.harvard.edu/ciddata/geographydata.htm]
Population density (ln) (total area population / land area in square kilometers)	Total population numbers: Maddison, A. (2001). The World Economy**:** A Millennial Perspective. Paris: OECD Publishing. Country area: CIA World Fact Book [https://www.cia.gov/library/publications/the-world-factbook/]
Institutional quality (Polity II)	Gurr, T. R. (1990). Polity II: Political Structures and Regime Change, 1800-1986. Boulder, CO: Center for Comparative Politics. [http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/9263]
Fertility 1950 (Births per woman)	World Population Prospects: The 2010 Revision [Downloaded at http://data.un.org/Default.aspx]
Ethnic fractionalization	Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2): 155-194. [http://www.anderson.ucla.edu/]
Student enrolment 1930s (Primary Enrollment*100 / Children of School Age)	Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History*, 47 (1): 82–99. [Based on: Benavot, A., and Riddle, P. (1988). The expansion of primary education, 1870-1940: Trends and Issues. *Sociology of Education*, 66 (3): 191-120 and Lindert, P. (2004). Growing Public: Social Spending and Economic Growth Since the Eighteenth Century, Vol. 2: Further Evidence. Cambridge: Cambridge University Press]
Natural resource exports around 1980 (percentage value of total exports)	World Bank Data 1999 (CD-Rom)
Physical capital (ln) about 1925	Enflo, K. and Baten, J. (2006). Estimates of Early Capital Stock Series. Working Paper Univ. Tuebingen/Lund.

Population 1500 (in millions)	Maddison, A. (2003). The World Economy Historical Statistics**.** Paris: OECD Publishing.
Southern Latitude	Comin, D., Easterly, W., Gong, E., 2010. Was wealth of nations determined in 1000 BC? American Economic Journal: Macroeconomics 2 (3), 65–97.

Appendix E: Educational performance in 1820 and in 1964-2003

Rank	country	Abcc 1820	Rank	country	cognitive
1	Finland	1	1	Taiwan (Chinese Taipei)	5.452
1	Japan	1	2	Korea, Rep.	5.338
1	Sweden	1	3	Singapore	5.330
1	Denmark	1	4	Japan	5.310
2	Switzerland	0.99	5	Macao-China	5.260
2	Slovenia	0.99	6	Hong Kong	5.195
2	Belgium	0.99	7	Estonia	5.192
2	Netherlands	0.99	8	Switzerland	5.142
3	Norway	0.98	9	Liechtenstein	5.128
3	Germany	0.98	10	Finland	5.126
4	France	0.97	11	Netherlands	5.115
5	Czech Rep.	0.96	12	Czech Rep.	5.108
5	Austria	0.96	13	Australia	5.094
5	Estonia	0.96	14	Austria	5.089
6	Canada	0.95	15	Slovak Rep.	5.052
6	United Kingdom	0.95	16	Hungary	5.045
7	Korea, Rep.	0.91	17	Belgium	5.041
8	Bulgaria	0.89	18	France	5.040
8	Hungary	0.89	19	Canada	5.038
9	Australia	0.88	20	Sweden	5.013
10	New Zealand	0.87	21	Ireland	4.995
11	Romania	0.86	22	Slovenia	4.993
11	China	0.86	23	New Zealand	4.978
12	United States	0.85	24	Denmark	4.962
12	Spain	0.85	25	Germany	4.956
12	Hong Kong	0.85	26	United Kingdom	4.950
12	Italy	0.85	27	China	4.939
12	Portugal	0.85	28	Iceland	4.936
13	Ireland	0.84	29	Russian Fed.	4.922
14	Latvia	0.83	30	United States	4.903
15	Slovak Rep.	0.81	31	Poland	4.846
16	**Uruguay**	**0.79**	32	Malaysia	4.838
17	Poland	0.77	33	Norway	4.830
18	**Brazil**	**0.72**	34	Spain	4.829
19	**Argentina**	**0.66**	35	Latvia	4.803
20	Russian Fed.	0.64	36	Bulgaria	4.789
21	Israel	0.62	37	Lithuania	4.779
22	**Colombia**	**0.6**	38	Italy	4.758
22	**Mexico**	**0.6**	39	Israel	4.686
23	Lithuania	0.59	40	Luxembourg	4.641
24	**Chile**	**0.57**	41	Greece	4.608
24	Serbia	0.57	42	Thailand	4.565
25	Greece	0.55	43	Portugal	4.564

25	Moldova	0.55	44	Romania	4.562
26	**Peru**	**0.51**	45	Cyprus	4.542
27	India	0.47	46	Moldova	4.530
28	Cyprus	0.39	47	Serbia	4.447
29	Thailand	0.37	48	Armenia	4.429
29	Indonesia	0.37	49	Swaziland	4.398
29	Philippines	0.37	**50**	**Uruguay**	**4.300**
30	Armenia	0.25	51	India	4.281
31	Turkey	0.21	52	Jordan	4.264
32	Albania	0.18	53	Iran	4.219
33	Egypt	0.08	54	Nigeria	4.154
			55	**Colombia**	**4.152**
			56	Macedonia	4.151
			57	Turkey	4.128
			58	Bahrain	4.114
			59	Zimbabwe	4.107
			60	Palestine	4.062
			61	**Chile**	**4.049**
			62	Kuwait	4.046
			63	Egypt	4.030
			64	**Mexico**	**3.998**
			65	Lebanon	3.950
			66	**Argentina**	**3.920**
			67	Indonesia	3.880
			68	Tunisia	3.795
			69	Albania	3.785
			70	Saudi Arabia	3.663
			71	Philippines	3.647
			72	**Brazil**	**3.638**
			73	Ghana	3.603
			74	Botswana	3.575
			75	Morocco	3.327
			76	**Peru**	**3.125**
			77	South Africa	3.089

Appendix F: The exclusion restriction for instrumental variables

One of the biggest challenges in any instrumental variable approach is the requirement of the exclusion restriction, which implies that the instrumental variables do not have a direct influence on the ultimate dependent variable except via the instrumented variable. In his seminal paper, Easterly (2007) studied the applicability of the exclusion restriction of relative soil and climatic suitability by using both theoretical reasoning and econometric tests. One possibility for such a direct causal channel is the possibility that wheat/rice and sugar have different effects on the wealth of the local population. This wealth difference could be a potential direct causal influence on cognitive abilities because those might depend on different investment possibilities. On the other hand, Easterly argues convincingly that the difference in the wealth effects of those agricultural goods are quite limited compared to all of the other goods that countries are producing.

Another potential violation of the exclusion restriction could stem from the widely discussed concept of the “natural resource curse” (Easterly 2007). Exceptionally high incomes from raw material exports might generate rents that in turn could lead to political economy problems (Sachs und Warner 1995, Auty 1993, see also the recent review by Frankel 2010). Sugar cane is a primary product that might produce such high windfall profits, for example. Isham et al. (2005) have developed a theory of “point-source” agricultural exports. Typical cases are exports such as sugar cane. The idea is that the “point-source” export revenues can more easily be captured by ruling elites than “diffuse” exports such as wheat and rice. Easterly (2007) argues that if these “resource curse” effects operate via inequality, the exclusion restriction is of course not violated. Most of the studies discussing these issues emphasize that the behavior of rich elites and their interactions with the institutional environment is the main issue, which is consistent with the inequality story (Engerman and Sokoloff 1997, Isham et al. 2005).

Nevertheless, one can still imagine that the resource curse works through other channels. One strategy to address these issues is to directly include additional controls for a resource-oriented export structure and determine whether inequality, measured with the sugar/wheat-rice suitability variable, turns insignificant. We again follow the literature in applying this strategy (Easterly 2007). We construct a variable of the share of raw material and mining exports relative to the country's total exports. The "resource curse" variable is insignificant for those cases that could be included and does not affect the significance of land inequality measured by the suitability variable (Table V).

What about the theoretical properties of the instrumental variable "low population density in 1500 interacted with southern location"? Low population density could, for example, have a negative impact on schooling, because commuting costs are much higher in sparsely populated countries. However, the population that experienced this situation around 1500 does not have much in common with the population in the 20th century. When large-scale immigration to the southern cone countries and Australia started in the 19th century, the composition of the population changed completely. Immigration countries such as Argentina, South Africa and Australia have a reputation for much higher schooling independent of high or low population density.

We also consider a potentially similar concern that might arise regarding our second measurement (even if it turns out to be less crucial to the overall analysis above). For example, we could imagine that the "low population density in 1500-southern latitude" countries such as Argentina and Australia were not only unequal in their land distribution but also invested differently in public schooling. However, we expect those countries to have actually invested more in schooling, which should have affected modern math/science skills positively (whereas their land inequality was high, which should have had a negative effect). Nevertheless, using the same strategy and explicitly including school enrollment in the 1930s as an additional control, we again find that measured land inequality is still significant (Table V, Model 2 and 3).

Appendix G: Sources of the ABCC 1820 estimates in detail (for reference abbreviations, see below)

Western Europe

Austria 1880 (Rothenbacher); Belgium 1856, 1866, 1880, 1890 (Rothenbacher); Switzerland 1860, 1870, 1880, 1888 (Rothenbacher); Denmark 1870, 1880, 1890 (Rothenbacher); Finland 1880, 1900 (Rothenbacher); France 1851, 1856, 1861, 1866 (Rothenbacher); Germany 1880 (Hippe and Baten); Ireland 1851, 1861 (Hippe and Baten); Italy 1871, 1931, 1936 (Rothenbacher); Netherlands 1849, 1859, 1869, 1879, 1889, 1899 (Rothenbacher); Norway (Crayen and Baten), using: Norway. Census of Norway 1865 and 1900. Statistics Norway, Oslo; Portugal: Stolz, Yvonne, Baten, J. and Jaime Reis, "Portuguese Living Standards 1720-1980 in European Comparison – Heights, Income and Human Capital", *Economic History Review 66-2* (2013), pp. 545-578; Spain 1900 (Hippe and Baten); Sweden 1880, 1900 (Rothenbacher); Switzerland 1860, 1870, 1888, 1900 (Rothenbacher); United Kingdom 1851 and 1881 (Crayen and Baten), using: Anderson, M. et al., 1979. National sample from the 1851 census of Great Britain [computer file]. Supplied by History Data Service, UK Data Archive (SN: 1316). Colchester, Essex; Schuerer, K., Woollard, M., 2002. National sample from the 1881 census of Great Britain [computer file]. Supplied by History Data Service, UK Data Archive (SN: 4375). Colchester, Essex;

Eastern and Southeastern Europe

Albania 1918 (Hippe and Baten), using: Eberhart, Helmut et al. (2010), Preliminary dataset "Albanische Volkszaehlung von 1918", entstanden an der Karl-Franzens-Universita¨t Graz unter Mitarbeit von Helmut Eberhart, Karl Kaser, Siegfried Gruber, Gentiana Kera, Enriketa Papa-Pandelejmoni und finanziert durch Mittel des Oesterreichischen Fonds zur Foerderung der wissenschaftlichen Forschung; (FWF). On early 19th C numeracy trends in the Balkans, see (Habsburg 1880); ; Armenia (Russian Empire 1897); Azerbeidshan (Russian Empire 1897); Belarus (Russian Empire 1897); Bulgaria 1893 (Hippe and Baten); Croatia 1880 (Rothenbacher); Czech lands 1880 (Rothenbacher); Cyprus, first benchmark decade is the 1870s (Crayen and Baten), on the change of numeracy in the region 1820-70, see Turkey; Estonia (Russian Empire 1897); Georgia (Russian Empire 1897); Greece 1903 (Hippe and Baten). Values are for the 1830s; Hungary 1880 (Rothenbacher); Kazakhstan (Russian Empire 1897); Kyrgystan (Russian Empire 1897); Lithuania (Russian Empire 1897); Latvia (Russian Empire 1897); Moldova (Russian Empire 1897); Poland 1880 (Habsburg part: Rothenbacher), 1880 (Prussian part: Hippe and Baten) and 1897 (Russian part: Russian Empire 1897)); Romania 1880 (Habsburg part: Rothenbacher; Romanian part assumed equal); Russia (Russian Empire 1897); Serbia 1867 (Crayen and Baten), based on friendly communication by Siegfried Gruber, who collected visitation data on a number of Serbian villages. Siegfried Gruber, Karl-Franzens-Universität Graz, Centre for Southeast European History, Project ‘‘Kinship and Social; Security"; Slovenia 1880 (Habsburg 1880); Slovakia 1880 (Habsburg 1880); Tajikistan (Russian Empire 1897); Turkmenistan (Russian Empire 1897); Ukraine (Russian Empire 1897); Uzbekistan (Russian Empire 1897)

Asia/ Oceania and Africa:

Australia: Meinzer, Nicholas (2013) “The selectivity of migrants to Australia: a new methodological approach”. Unpubl. Master thesis Univ. Tuebingen; China: Baten, J., Debin Ma, Stephen Morgan and Qing Wang (2010) “Evolution of Living Standards and Human Capital in China in the 18-20th Centuries: Evidences from Real Wages, Age-heaping, and Anthropometrics”, *Explorations in Economic History* 47-3: *347-359*; Egypt: 1848 (Census of Cairo), 1907 (Census of Egypt: The Statistical Department of the Ministry of Finance Egypt, 1907. Statistical yearbook of Egypt. 3rd census of Egypt 1905. Cairo, The Government Press; Hong Kong Baten, J., Debin Ma, Stephen Morgan and Qing Wang “Evolution of Living Standards and Human Capital in China in the 18-20th Centuries: Evidences from Real Wages, Age-heaping, and Anthropometrics”, *Explorations in Economic History* 47-3 (2010): *347-359*; India: 1891 (Census of India, 1891 (Bombay, Madras, North-Western Provinces) Indian Empire Census of 1891, 1901, 1911 and 1921. The Superintendent of Government Printing India, Calcutta; Indonesia: Southeast Asia estimate of Baten, J, and Johan Fourie “Numeracy in the 18th Century Indian Ocean Region”): ERSA Working Paper No. 270, complemented with evidence used also in the study Baten, J., Mojgan Stegl and Pierre van der Eng (2013). The Biological Standard of Living and Body Height in Colonial and Post-colonial Indonesia, 1770–2000”, *Journal of Bioeconomics 15*: 103-122; Japan: Ministry of Internal Affairs and Communications, 1882. First Statistical Yearbook of the Japan Empire. Population statistics of the Province of Kai 1879 (today’s Yamamashu Prefecture). Government Publications, Tokyo; Korea: Baten, J and Kitae Sohn: “Back to the ‘Normal’ Level of Human-Capital Driven Growth? A Note on Early Numeracy in Korea, China and Japan, 1550–1800”, University of Tübingen Working Papers in economics and finance, No. 52; New Zealand: Meinzer, Nicholas (2013) “The selectivity of migrants to Australia: a new methodological approach”. Unpubl. Master thesis Univ. Tuebingen; Philippines: Southeast Asia estimates of (Crayen and Baten) and Baten, J, and Johan Fourie “Numeracy in the 18th Century Indian Ocean Region,” ERSA Working Paper No. 270, complemented with evidence by Kathrin Grether (2012), Langfristige Humankapitalentwicklung auf den Philippinen im international Vergleich. Unpubl. BA Thesis Univ. Tuebingen; Thailand: Southeast Asia estimates of (Crayen and Baten), and Baten, J, and Johan Fourie “Numeracy in the 18th Century Indian Ocean Region,” ERSA Working Paper No. 270; Turkey (Russian Empire 1897): evidence on the province of Kars; see the discussion in Crayen and Baten (2010) about the representativeness of the province. See also on the Ottoman census of 1831 Starbatty, Peter (2011). Humankapitalentwicklung im Osmanischen Reich 1760-1810. Regionale und ethnische Unterschiede. Unpubl. BA Thesis Univ. Tuebingen.

The Americas:
Argentina (Manzel, Baten, Stolz 2012), based on Argentina: National census data of 1869 and 1895, published in Somoza, J., Lattes, A., 1967. Muestras de los dos primeros censos nacionales de población, 1869 y 1895. Documento de Trabajo No 46, Instituto T. Di Tella, CIS, Buenos Aires; Brazil (Manzel, Baten, Stolz). ; Canada: (Crayen and Baten), using the 1852 and 1881 Historical Censuses of Canada (Canada East, Canada West, New Brunswick and Nova Scotia). Université de Montréal, Montréal; Chile : Robert Pertschy (2012), Regionale Unterschiede der langfristigen Humankapitalentwicklung in Chile im 19. Jahrhundert. Unpubl. BA Thesis Univ. Tuebingen; Colombia (Manzel, Baten, Stolz); Ecuador (Manzel, Baten, Stolz); complemented with new evidence by Christian Schneider (2011), Das Humankapital in den Regionen Ecuadors, Unpubl. Diploma Thesis Univ. Tuebingen; Mexico

(Manzel, Baten, Stolz); Peru (Manzel, Baten, Stolz); Complemented with evidence by Sabin Guettler (2011), Verbreitung der Bildungsinnovationen in Peru und Ecuador im 18. und 19. Jahrhundert, Unpubl. Diploma Thesis Univ. Tuebingen; United States: 1850, 1860, 1870, 1880, 1900: Ruggles, S., Alexander, J.T., Genadek, K., Goeken, R., Schroeder, M.B., and Sobek, M. (2010). Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota; Uruguay (Manzel, Baten, Stolz).

Reference abbreviations above:

(Crayen and Baten): Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. Explorations in Economic History, 47(1): 82-99.

(Habsburg 1880): Austro-Hungarian census of 1880, published as Österreichische Statistik, Band 1, Heft 1–3, Band 2, Heft 1–2 and Band 5, Heft 3, 1882–1884. The evidence covers Austria, Bosnia and Herzegovina, Croatia, Czech Republic, Hungary, Slovakia and Slovenia. We merged Austrian, Russian, and German regional statistics to obtain weighed averages for the modern territories of Ukraine and Poland.

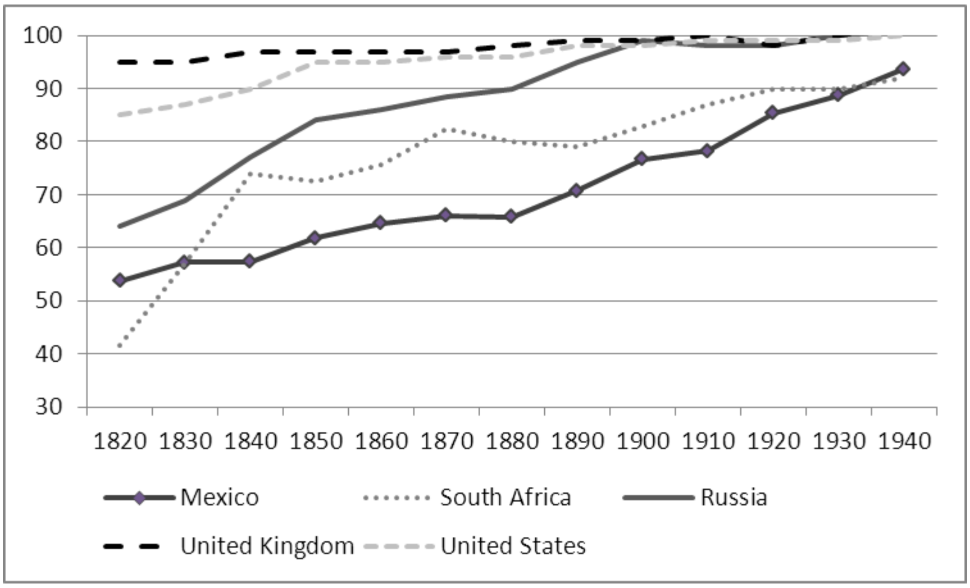
(Hippe and Baten): Hippe, R. and Baten, J. (2012) "The Early Regional Development of Human Capital in Europe, 1790 – 1880, *Scandinavian Economic History Review*, 60, Number 3, 1 November 2012 , pp. 254-289

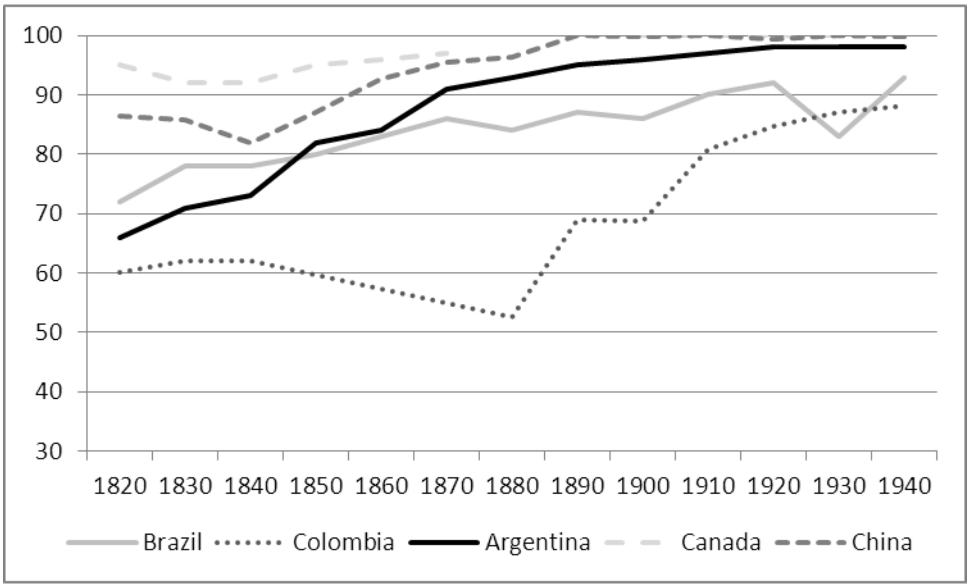
(Manzel, Baten and Stolz) Manzel, K., Baten, J. and Stolz, Y. (2012) "Convergence and Divergence of Numeracy: The Development of Age Heaping in Latin America, 17th to 20th Century", *Economic History Review* 65, 3 (2012), pp. 932–960. Detailed sources are listed in their online appendix p.4/5

(Rothenbacher): Rothenbacher, F. (2002). The European Population 1850-1945. Basingstoke: Palgrave Macmillan.

(Russian Empire 1897): First General Russian Empire Census of 1897. издание центрального статистического комитета министерства внутренних (1899).

Appendix H: How path-dependent is numeracy over time?





How path-dependent is numeracy over time? (Pairwise correlations for dataset used in the Figures above. All correlation coefficients are statistically significant at the 10% level)

	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940
1820	1.00												
1830	0.98	1.00											
1840	0.83	0.91	1.00										
1850	0.83	0.90	0.96	1.00									
1860	0.79	0.87	0.93	0.98	1.00								
1870	0.68	0.78	0.87	0.95	0.98	1.00							
1880	0.65	0.75	0.84	0.94	0.98	1.00	1.00						
1890	0.76	0.83	0.85	0.95	0.97	0.96	0.96	1.00					
1900	0.70	0.76	0.82	0.93	0.96	0.96	0.97	0.98	1.00				
1910	0.76	0.82	0.87	0.95	0.95	0.93	0.94	0.99	0.97	1.00			
1920	0.74	0.79	0.83	0.94	0.95	0.94	0.95	0.99	0.98	0.99	1.00		
1930	0.61	0.61	0.65	0.76	0.75	0.73	0.74	0.80	0.86	0.84	0.86	1.00	
1940	0.73	0.75	0.77	0.88	0.90	0.87	0.89	0.91	0.95	0.89	0.91	0.91	1.00

Appendix I: Inequality and Human Capital in OLS Regressions with constant Observation Numbers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Robust regression	No	No	No	No	No	Yes	Yes	No	No	No	No	No
WLS	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Early numeracy	1.56***	1.51***	1.21***	0.87***	0.84***	0.87***	0.96***	0.78***	0.88***	1.26***	1.15***	0.81***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Early land inequality (around 1890)		-1.52***	-1.70***	-1.52***		-1.03***		-1.31***		-1.48*	-1.28**	-1.10***
		(0.00)	(0.00)	(0.00)		(0.00)		(0.00)		(0.05)	(0.05)	(0.00)
Income inequality (1890)					-2.07**		-1.00*		-1.93**	-0.87	-0.26	-0.83
					(0.03)		(0.08)		(0.03)	(0.18)	(0.69)	(0.41)
GDP/c 1910			0.19	0.13	0.07	-0.07	-0.14	-0.02	-0.09			-0.02
			(0.14)	(0.39)	(0.66)	(0.38)	(0.20)	(0.84)	(0.36)			(0.68)
Tropic share				-0.30	0.05	-0.81***	-0.67***	-0.51**	-0.17	-0.47***	-0.53***	-0.44**
				(0.36)	(0.91)	(0.00)	(0.00)	(0.03)	(0.56)	(0.00)	(0.00)	(0.02)
Population density				0.04	0.05	0.08***	0.08***	0.06*	0.05			0.05
				(0.39)	(0.20)	(0.00)	(0.00)	(0.09)	(0.15)			(0.24)
Institutional quality				-0.01	-0.00	0.00	0.01	-0.01	-0.00			-0.00
				(0.39)	(0.94)	(0.83)	(0.14)	(0.36)	(0.84)			(0.43)
Fertility				-0.05	-0.10	-0.05	-0.05	-0.08*	-0.12**			-0.09**
				(0.41)	(0.11)	(0.18)	(0.29)	(0.09)	(0.03)			(0.04)
Ethnic Fractionalization				-0.02	-0.33	0.24	0.13	0.17	-0.19			0.13
				(0.95)	(0.46)	(0.19)	(0.56)	(0.56)	(0.52)			(0.65)
Constant	3.39***	4.41***	3.32***	4.08***	4.58***	5.12***	5.32***	5.17***	5.81***	4.99***	4.74***	5.43***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	30	30	30	30	30	30	30	30	30	30	30	30
Adj. R-Squared	0.54	0.69	0.70	0.73	0.73	0.93	0.86	0.85	0.81	0.73	0.81	0.85

Notes: Dependent variable: Cognitive Skills, average test scores 1964 to 2003. Early numeracy: ABCC 1820; Early land inequality: Gini coefficient; Institutional quality: Polity 2; Fertility: refers to 1950; Fractionalization: Index from Alesina et al. (2003); Population density and GDP/c are in logs; Specifications (6) to (12) are weighted by the square root of population numbers of each country. All specifications include heteroskedasticity robust standard errors and clustered standard errors by world region. Models (6) and (7) were estimated with R-command mlrobust. Robust p-values in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.10$

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review,* 91 (5): 1369-1401.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution *The Quarterly Journal of Economics*, 117(4): 1231-1294.

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

Allende, Isabel (1982). La Casa de los Espiritus. Barcelona: Plaza and Janés.

Auty, R. M. (1993). *Sustaining Development in Mineral Economies: The Resource Curse Thesis*. London: Routledge.

Banerjee, A., and Iyer, L. (2005): History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India, *American Economic Review* 95(4): 1190-1213.

Barro, R. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5(1): 5-32.

Baten, J. (1999). Ernaehrung und wirtschaftliche Entwicklung in Bayern 1730-1880. Stuttgart: Steiner.

Baten, J., Földvari, P., van Leeuwen, B., and van Zanden, J.L. (2012). World Income Inequality 1820-2000. Working Paper.

Bates, R. (1981). States and Markets in Tropical Africa: The Political Basis of Agricultural Policy. Berkeley: University of California Press, Series on Social Choice and Political Economy.

Becker, G. S. (1960). An economic analysis of fertility. In: Demographic and economic change in developed countries, edited by G.S. Becker. Princeton, NJ: Princeton University Press (pp. 209–240).

Becker, S., and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, 124(2): 531-596.

Binswanger, H. P., and Deininger, K. (1993). South African land policy: The legacy of history and current options. *World Development,* 21(9): 1451-1475.

Blum, M., and Baten, J. (2011). Anthropometric within-country Inequality and the Estimation of Skill Premia with Anthropometric Indicators. *Review of Economics - Jahrbuch fuer Wirtschaftswissenschaften,* 62(2): 107-138.

Botticini, M., and Eckstein, Z. (2007). From Farmers to Merchants, Conversions and Diaspora: Human Capital and Jewish History. *Journal of the European Economic Association*, 5(5): 885-926.

Bulmer-Thomas, V. (1994). The Economic History of Latin America since Independence. Cambridge: Cambridge University Press.

Cinnirella, F., and Hornung, E. (2011). Landownership Concentration and the Expansion of Education (September 30, 2011). CESifo Working Paper Series No. 3603.

Clark, G. (2007). A Farewell to Alms. A Brief Economic History of the World. Princeton: Princeton University Press.

Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History,* 47 (1): 82–99.

Deininger, K. (2003). Land Policies for Growth and Poverty Reduction. Oxford: Oxford University Press.

Deininger , K., and Squire, L. (1998). New Ways of Looking at Old Issues: Inequality and Growth. *Journal of Development Economics,* 57(2): 259–287.

Diamond, J. (1997). Guns, Germs and Steel: The Fates of Human Societies. New York: W.W. Norton.

Easterly, W., and Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Division. *Quarterly Journal of Economics*, 152(4): 1203-1250.

Eastwood, R., Lipton, M., and Newell, A. (2010). Farm Size, in Handbook of Agricultural Economics, edited by R. Evenson and P. Pingali, 3323-3397. Amsterdam: Elsevier.

Easterlin, R. A. (1980). Towards a more General Economic Model of Fertility Determination: Endogenous Preferences and Natural Fertility. In: Population and economic change in developing countries, edited by R. A. Easterlin. Chicago: University of Chicago Press for National Bureau of Educational Research.

Easterly, W., and Levine, R. (2003). Tropics, Germs, and Crops: How Endowments influence Economic Development. *Journal of Monetary Economics*, 50: 3-39

Easterly, W. (2007). Inequality Does Cause Underdevelopment. *Journal of Development Economics*, 84(2): 755-776.

Enflo, K., and Baten, J. (2006). Estimates of Early Capital Stock Series. Working Paper Univ. Tuebingen/Lund.

Engerman, S., and Sokoloff, K. (1997). Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic

Historians of the United States. In: How Latin America Fell Behind, edited by S. Haber. Stanford, CA: Stanford University Press.

Engerman, S., and Sokoloff, K. (2002). Factor Endowments, Inequality and Paths of Development among New World Economies. *Economia*, 3(1): 41-109.

Frankel, J. (2010). The Natural Resource Curse: A Survey. Harvard Kennedy School Working Paper 10-005.

Frankema, E. (2010). The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions?. *Economic History Review*, 63(2): 418-451.

Frankema, E., and Van Waijenburg, M. (2012). Structural impediments to African Growth? New evidence from real wages in British Africa, 1880–1965. *The Journal of Economic History*, 72(4): 895-926.

Galor, O. (2011). Inequality, Human Capital Formation and the Process of Development. NBER Working Paper Series No. 17058.

Galor, O., Moav, O., and Vollrath, D. (2009). Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. *Review of Economic Studies*, 76(1): 143-179.

Galor, O., and Moav, O. (2002). Natural Selection And The Origin Of Economic Growth. *The Quarterly Journal of Economics*, 117(4): 1133-1191.

Galor, O., and Weil, D. (2000). Population, Technology and Growth: from the Malthusian Regime to the Demographic Transition and beyond. *American Economic Review*, 90: 806–828.

Galor, O., and Zeira, J. (1993). Income Distribution and Macroeconomics. *Review of Economic Studies*, 60: 35-52.

Gelman, J., and Santilli, D. (2010). Rising Inequality. Land property in Buenos Aires between 1839 and 1855. *Investigaciones de Historia Económica (IHE) Journal of the Spanish Economic History Association*, 18: 11-33.

Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004). Do Institutions Cause Growth?. *Journal of Economic Growth*, 9(3): 271-303.

Hanushek, E.A., and Woessmann, L. (2012a). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *Journal of Economic Growth* , 17(4): 267-321.

Hanushek, E.A., and Woessmann, L. (2012b). Schooling. Educational Achievement, and the Latin American Growth Puzzle. *Journal of Development Economics*, 99 (2): 497–512.

Hippe, R., and Baten, J. (2012). ‘Keep them ignorant.’ Did inequality in land distribution delay regional numeracy development?. University of Tuebingen Working Papers.

Isham, J., Woolcock, M., Pritchett, L., and Busby, G. (2005). The varieties of resource experience: natural resource export structures and the political economy of economic growth. *The World Bank Economic Review*, 19:141–174.

Jones, E. L. (1981). The European miracle. Cambridge: Cambridge University Press.

Lindert, P. (2004): Growing Public. Social Spending and Economic Growth Since the Eighteenth Centruy. Cambridge: Cambridge University Press.

Lindert, P. H., and Williamson, J. G. (2003). Does Globalization Make the World More Unequal? In: Globalization in Historical Perspective, edited by M.D. Bordo, A. M. Taylor and J.G. Williamson. Chicago: University of Chicago Press.

Lipton, M. (2009). Land Reform in Developing Countries: Property Rights and Property Wrongs. London: Routledge.

Montesquieu, Charles de Secondat (1748): The Spirit of the Laws. New York: Cambridge University Press, 1989.

Moradi, A., and Baten, J. (2005). Inequality in Sub-Saharan Africa 1950-80: New Estimates and New Results. *World Development*, 33(8): 1233-1265.

North, D. (1981). Growth and Structural Change. New York: W. W. Norton.

Nugent, J. B., and Robinson, J. A. (2010). Are factor endowments fate?. *Revista de Historia Economica/Journal of Iberian and Latin American Economic History*, 28(01): 45-82.

Ramcharan, R. (2010). Inequality and Redistribution: Evidence from U.S. counties and states, 1890-1930. *Review of Economics and Statistics*, 92(4): 729–744.

Resor, R. R. (1977). Rubber in Brazil: Dominance and Collapse, 1876-1945. *The Business History Review*, 341-366.

Sachs, J.D., and Warner, A.M. (1997). Sources of slow growth in African economies. *Journal of African Economies*, 6 (3): 335-376.

Sachs, J. D. (2001). Tropical Underdevelopment, NBER Working Paper No. 8119.

Sachs, J. D., and Warner, A.M. (1995). Natural resource abundance and economic growth. NBER Working Paper No. 5398.

Solimano, A. (2001). *International migration and the global economic order: An overview*. Vol. 2720. World Bank Publications.

Summerhill, W. R. (2010). *Colonial Institutions, Slavery, Inequality, and Development: Evidence from São Paulo, Brazil*. MPRA Paper, No. 22162.

Stock, J.H., and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In: Identification and Infernce for Econometric Models: Essays in Honor of Thomas Rothenberg, edited by D.W.K. Andrews and J.H. Stock, 30-108. Cambridge: Cambridge University Press.

Stolz, Y., Baten, J.“Brain Drain, Numeracy and Skill Premia during the Era of Mass Migration: Reassessing the Roy-Borjas Model”, with Yvonne Stolz, *Explorations in Economic History 49* (2012), pp. 205-20.

Sokoloff, K. L., and Engerman, S. L. (2000). History Lessons: Institutions, Factors Endowments, and Paths of development in the new world. *Journal of Economic Perspectives*, 14(3): 217-232.

Wegenast, T. (2009). The Legacy of Landlords: Educational Distribution and Development in a Comparative Perspective. *Comparative Governance and Politics*, 3(1): 81-107.

Wegenast, T. (2010). Cana, Café, Cacau: Agrarian Structure and Educational Inequalities in Brazil. *Revista de Historia Económica – Journal of Iberian and Latin American Economic History*, 28(1): 103-137.

Table I: Summary Statistics of Data included in Regressions

Variable name	Obs. No.	Mean value	Standard deviation	Minimum value	Maximum value
Cognitive skills	69	4.62	0.55	3.09	5.45
Income inequality (Gini 1890)	42	0.38	0.07	0.28	0.62
Early land inequality (Gini coefficient)	52	0.61	0.14	0.29	0.91
Early numeracy (abcc 1820)	54	0.74	0.25	0.08	1.00
Tropic share	69	0.11	0.28	0.00	1.00
Population density (ln)	61	2.22	1.93	-3.14	4.65
Institutional quality (Polity 2)	50	3.25	7.46	-9.00	10.00
Fertility 1950	57	4.14	1.80	1.98	7.34
Ethnic fractionalization	74	0.33	0.23	0.00	0.85
Student enrolment 1930s	58	0.45	0.27	0.03	0.97
Natural resource exports	51	0.33	0.29	0.02	0.99
Physical capital (ln)	38	7.76	1.18	4.78	9.99

Notes: Cognitive Skills is a measure of human capital introduced by Hanushek and Woessmann (2012a) that is constructed out of international test scores in math and sciences.
The measure of land inequality refers to the size distribution of the land holdings. Landholders were those who produced on their own land or those who or rented land. It reflects the control over land as a production factor (at least in the medium run). However, this data includes land holdings used for all purposes – arable land, land used for permanent crops, or pastures. The "quality" of land has not been reflected in most land inequality estimations.
'Tropic share' refers to the share of a country's population living in a tropical zone (1995). Source: Center for international development, Geography datasets (download at http://www.cid.harvard.edu/ciddata/geographydata.htm, General Measures of Geography).
Fertility data are constructed as an average of United Nations country data from 1950-1955 and from 1955-1960. See also the Data Appendix.

Table II: Average Influence of a Land Reform on the Change in Land Inequality

	(1)	(2)
Method	LSDV	LSDV
Land reform	-5.57*	-5.57*
	(0.05)	(0.06)
Western Europe and Offshoots		-11.31***
		(0.00)
Country fixed effects	Yes	Yes
Constant	0.37	0.37
	(0.24)	(0.25)
Observations	60	60
R-squared	0.65	0.80

Notes: Dependent variable: Change in land inequality. Only those countries were included for which the Gini coefficient was available in more than one year. "Change in land inequality" is constructed by the difference in the Gini from one point in time to another. "Land reform" is a dummy variable that takes the value of 1 if a land reform took place between two Gini estimates. "Western Europe and Offshoots" refers to countries with high GDP per capita (over 25000 U.S. Dollar, Maddison 2001), namely Austria, Australia, Finland, France, Italy, Netherlands, New Zealand and Sweden. The estimation technique is Least Squares estimation with country dummy variables.
Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table III: Inequality and Human Capital in OLS Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Robust regression	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No
WLS	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes
Early numeracy	1.48***	1.47***	1.30***	0.91**	0.93**	1.16**	1.05***	1.34***	1.03***	0.91***	0.95***	1.27***	1.16***	0.81***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Early land inequality (around 1890)		-1.37*	-1.50*	-1.39**		-1.44**		-0.81***		-1.19**		-1.43*	-1.22*	-1.10***
		(0.07)	(0.07)	(0.02)		(0.04)		(0.00)		(0.02)		(0.05)	(0.05)	(0.00)
Income inequality (1890)					-1.93**		-2.28*		-0.88		-1.85**	-0.93	-0.37	-0.83
					(0.04)		(0.06)		(0.27)		(0.04)	(0.16)	(0.59)	(0.41)
GDP/c 1910			0.16	0.21	0.08	0.46	0.28	-0.07	-0.10	0.01	-0.08			-0.02
			(0.16)	(0.14)	(0.60)	(0.13)	(0.39)	(0.35)	(0.407)	(0.80)	(0.40)			(0.68)
Tropic share				-0.27	0.10	-0.42*	0.00	-0.85***	-0.59***	-0.49***	-0.15	-0.46***	-0.52***	-0.44**
				(0.21)	(0.68)	(0.06)	(0.98)	(0.00)	(0.007)	(0.00)	(0.38)	(0.00)	(0.00)	(0.02)
Population density				0.03	0.05	0.04	0.04	0.03	0.07*	0.05	0.06			0.05
				(0.38)	(0.30)	(0.48)	(0.51)	(0.15)	(0.078)	(0.21)	(0.22)			(0.24)
Institutional quality				-0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00			-0.00
				(0.70)	(0.81)	(0.68)	(0.86)	(0.95)	(0.626)	(0.81)	(0.98)			(0.43)
Fertility				-0.02	-0.09	0.03	-0.04	0.03	-0.05	-0.04	-0.10*			-0.09**
				(0.75)	(0.10)	(0.62)	(0.60)	(0.53)	(0.372)	(0.27)	(0.09)			(0.04)
Ethnic Fractionalization				-0.07	-0.34	-0.14	-0.78	0.30*	0.03	0.09	-0.22			0.13
				(0.81)	(0.52)	(0.75)	(0.28)	(0.07)	(0.919)	(0.75)	(0.54)			(0.65)
Physical capital						-0.17	-0.16							
						(0.30)	(0.29)							
Constant	3.57***	4.39***	3.36***	3.24***	4.30***	2.36*	4.10**	4.14**	4.49***	4.68***	5.57***	4.97***	4.74***	5.43***
	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.089)	(0.015)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Observations	54	44	37	35	32	31	27	35	32	35	32	34	34	30
Adj. R-squared	0.54	0.68	0.71	0.71	0.71	0.69	0.71	0.61	0.82	0.81	0.81	0.76	0.82	0.85

Notes: Dependent variable: Cognitive Skills, average test scores 1964 to 2003. Early numeracy: ABCC 1820; Early land inequality: Gini coefficient; Institutional quality: Polity 2; Fertility: refers to 1950; Fractionalization: Index from Alesina et al. (2003); Population density and GDP/c are in logs; Specifications (8) to (14) are weighted by the square root of population numbers of each country. Models 8 and 9 were estimated with R-command `mlrobust` and the option `weights=popsqrt` to include weights by population. Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table IV: Regressions of Cognitive Abilities on Land Inequality and other Factors: Instrumental Variable Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation	TSLS	TSLS	LIML	TSLS	TSLS	TSLS
Instruments	Sugar/ (WheatRice) Def. 1 + LowPop* Southern	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3
Early numeracy	1.57***	1.41***	1.41***	1.00***	1.09***	1.01***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Early land inequality (around 1890)	-1.28*	-2.75**	-2.75**	-2.24*	-2.41**	-2.24*
	(0.07)	(0.01)	(0.01)	(0.08)	(0.02)	(0.08)
GDP/c 1910				0.23	0.25*	0.24
				(0.14)	(0.08)	(0.19)
Tropic share				-0.20	-0.29	-0.18
				(0.43)	(0.22)	(0.45)
Institutional quality					-0.01	-0.00
					(0.60)	(0.77)
Fertility				-0.04		-0.03
				(0.59)		(0.61)
Constant	4.25***	5.28***	5.28***	3.71***	3.44***	3.55***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	44	47	47	37	40	37
Second Stage R-squared	0.71	0.58	0.58	0.74	0.72	0.74
First Stage						
Sugar/(WheatRice) Def. 1	0.73**					
	(0.01)					
LowPop*Southern	0.01***					
	(0.00)					
Sugar/(WheatRice) Def. 2		0.54**	0.54**			
		(0.01)	(0.01)			
Sugar/(WheatRice) Def. 3				0.56**	0.66**	0.56**
				(0.02)	(0.03)	(0.03)
F-Statistic	11.24	6.78	LIML	5.63	5.51	5.29
Durbin-Wu-Hausman p-value	0.99	0.79	LIML	0.41	0.31	0.37
Sargan p-value	0.17					

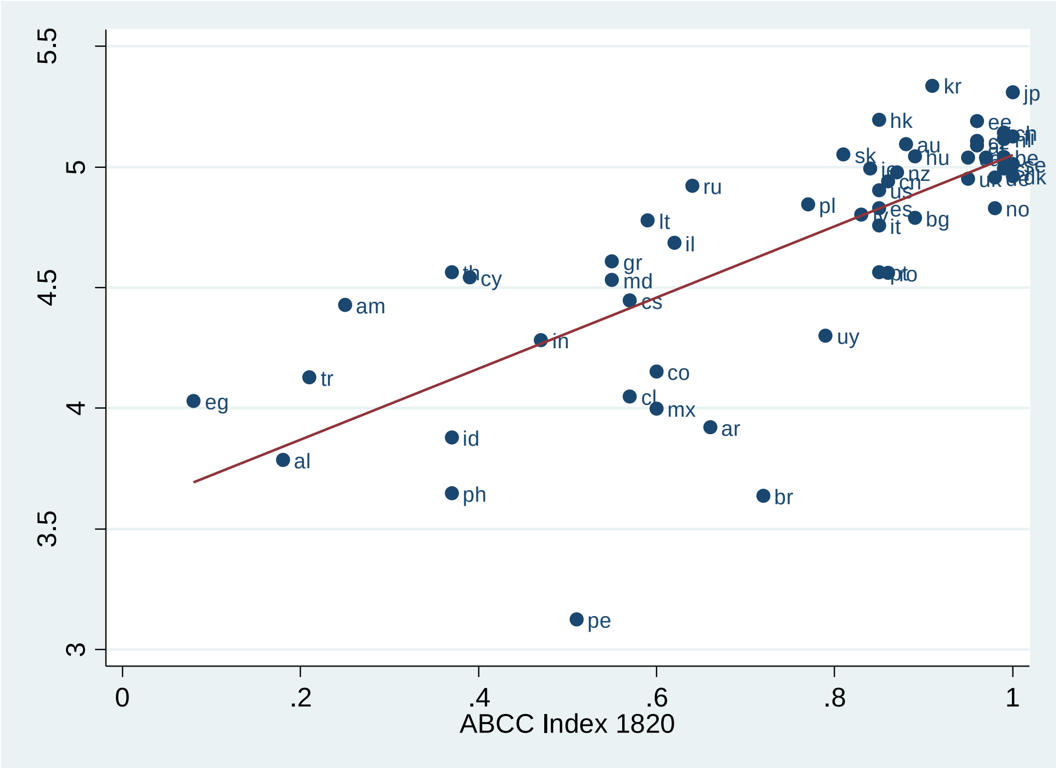
Notes: Dependent variable (of the second stage): Cognitive Skills. Control variables: ABCC 1820, tropic share, population density (ln), institutional quality, GDP per capita in 1910 (ln) and physical capital (ln). Instrument "Sugar/ (wheat-rice) Def. 1" includes the share of the land that is very suitable and suitable for the plantation of the three crops (sugar, wheat and rice), Instrument "Sugar/ (wheat-rice) Def. 2" the share being very suitable, suitable and moderately suitable and, last, "Sugar/ (wheat-rice) Def. 3" includes the share of the land very suitable, suitable, moderately suitable *and* marginally suitable for crop plantation.
Robust p-values in parentheses:
***p<0.01, **p<0.05, *p<0.10

Table V: Instrumental Variable Regression controlling for Primary Goods Exports

	(1)	(2)	(3)
Method	TSLS	TSLS	TSLS
Instruments	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 1 + LowPop* Southern	Sugar/ (WheatRice) Def. 2 + LowPop* Southern
Second Stage			
Land inequality (Gini 1890)	-2.67**	-1.10*	-1.42*
	(0.04)	(0.09)	(0.07)
Early numeracy (abcc 1820)	1.21***	1.21***	1.18***
	(0.00)	(0.00)	(0.00)
Natural resource exports 1980	-0.05		
	(0.85)		
Tropic share	-0.55**	-0.35	-0.36
	(0.03)	(0.16)	(0.13)
School enrollment		0.34	0.36*
		(0.12)	(0.08)
Constant	5.47***	4.26***	4.47***
	(0.00)	(0.00)	(0.00)
Observations	40	41	41
Second Stage R-squared	0.65	0.75	0.76
First Stage			
Sugar/(WheatRice) Def. 1		0.71**	
		(0.03)	
Sugar/(WheatRice) Def. 2	0.66**		0.63**
	(0.04)		(0.01)
LowPop*Southern		0.01***	0.01***
		(0.00)	(0.00)
F-Statistic	4.63	8.68	6.69

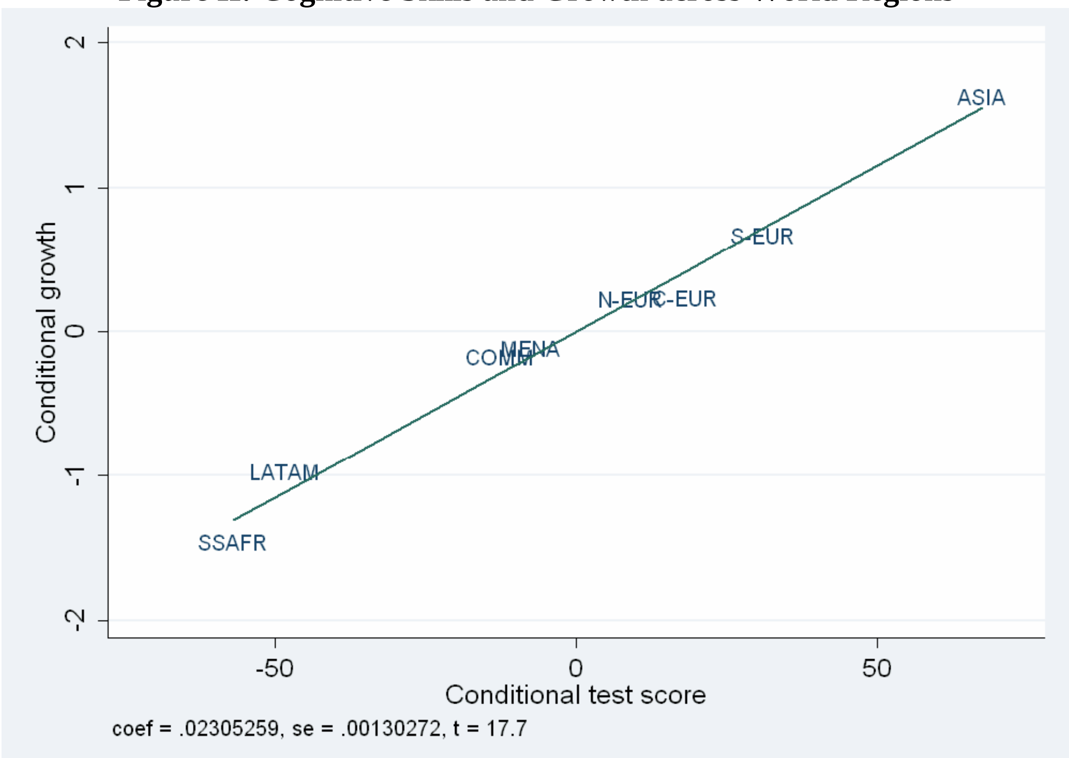
Notes: Dependent variable (of the second stage): Cognitive Skills. Control variables: ABCC 1820, tropic share, primary goods exports and school enrollment in the 1930s. Instrument "Sugar/ (wheat-rice) Def. 3" includes the share of the land that is very suitable, suitable and moderately for crop plantation. Natural resource exports: raw material and mining exports per GDP. Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Figure I: Math and Science Cognitive Skills 1964-2003 and Numeracy 1820 (ABCC Index)



Note: Abbreviatios follow ISO codes. See also Appendix C

Figure II: Cognitive Skills and Growth across World Regions



Source: Hanushek and Woessmann (2012a)

Notes: Region codes are East Asia and India (ASIA), Central Europe (C-EUR), Commonwealth OECD members (COMM), Latin America (LATAM), Middle East and North Africa (MENA), Northern Europe (N-EUR), Southern Europe (S-EUR), Sub-Saharan Africa (SSAFR).

Figure III: Early Land Inequality (around 1890, Gini coefficients)

Figure IV: Land Inequality versus Income Inequality

Note: Abbreviatios follow ISO codes. See also Appendix C

Figure V: Cognitive Skills, 1964-2003