Gabriele Cappelli (U Siena) and Joerg Baten (U Tuebingen, CESifo, CEPR)

Numeracy development in Africa: new evidence from a long-term perspective (1730-1970)

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

Cappelli, Gabriele and Joerg Baten (2021). Numeracy development in Africa: new evidence from a long-term perspective (1730-1970)”, *Journal of Development Economics* (2021). Online: https://doi.org/10.1016/j.jdeveco.2021.102630

Abstract

Historical evidence of numeracy on the African continent since the 18th century is presented for the first time based on a panel dataset of 43 African countries covering the periods before, during and after colonialism (1730–1970). Estimates of numeracy draw on the age-heaping methodology: we carefully discuss the potential biases and sources of measurement error concerning the use of this index for long-term analyses. These new estimates enable us to gain a better understanding of long-term African development. We find that the evolution of numeracy over time correlates with differences in colonial education systems, even when controlling for other variables.

Keywords

Human capital, Africa, Colonial education systems, Numeracy, History

1. Introduction

Historical events are fundamental determinants of economic performance (Allen 2011). However, until recently, research in quantitative economic history has mainly focused on today's high- and middle-income economies. For example, no economic history study on Somalia – a sizeable African country – appears in standard datasets of economic literature, while thousands of contributions on the UK can be found. Although a number of studies on long-term development and growth have started to address developing countries (see Nunn 2020 for a review), evidence on some developing regions is still very limited due to the limited availability of historical data.

This lack of research is particularly evident for the African continent: several populations did not produce a substantial number of written records before European rule, and during and following the colonial period, statistical information was seldom representative of the whole population. However, some aspects of Africa's long-term development can be reconstructed using new proxy indicators. Although these indicators might capture dimensions of social and economic change with a certain degree of measurement error, they are crucial to write an economic history of the continent.

The main contribution of this study is to construct a panel dataset of numeracy that includes several African countries from the beginning of the 18th century until the end of the colonial period (1970) in ten-year intervals. Numeracy, which is an important component of overall human capital, is estimated by studying age heaping, i.e., the tendency of individuals who do not know their exact age to declare it as ending in digits such as zero or five.

This study represents the first time that numeracy – and its increase over time – can be reconstructed for a large number of African countries over such a long time period, covering the precolonial, colonial and early postcolonial phases. We admit that the evidence is characterized by potential issues and gaps, which are addressed in the paper. By exploring numeracy in our panel dataset, we highlight various phases and cross-country trends. In general, the increase in numeracy was steady during the early 18th century, and its pace increased by the eve of the 20th century. This descriptive evidence suggests that the increase in numeracy was more rapid in areas that were influenced by the British education system.

As a second contribution, we use our panel dataset to study whether our new evidence is consistent with a set of hypotheses, based on conditional correlations (without claiming a causal relationship). One hypothesis is that colonial school systems are associated with more rapid increase in numeracy than noncolonial systems: although human capital formation in Africa had long been mainly based on household education and traditional schooling (Mosweunyane 2013; Bolak Funteh 2015), when the European occupation was completed at the turn of the 20th century, colonial administrations sought to strengthen the diffusion of Western schooling as a way to legitimize their rule (White 1996). More specifically, we explore whether the British education system is associated with a more rapid increase in numeracy compared to other colonial school systems in Africa, against the prior that the territories that would be occupied by the British had already developed high numeracy during precolonial times. We control for a number of variables that are known to be relevant for the historical development of Africa, such as the slave trade, the cash-crop trade, the strength of precolonial institutions, suitability to the tsetse fly, and the presence of colonial railways.

Although we focus on colonial education systems, our results are consistent with the view that African agency was a crucial aspect of the numeracy path followed by African countries. The choices and responses of Africans to changing incentives during the colonial period, as well as the role played by African missionaries at the turn of the 20th century, were factors that were strongly associated with the increase in numeracy and, in general, human capital accumulation (Frankema 2012; Meier Zu Selhausen 2019).

While the hypotheses that we explore have been studied in related work (Frankema 2012; Cogneau and Moradi 2014; Dupraz 2019), previous analyses could not focus on precolonial numeracy as an important component of human capital and did not consider such a large sample of African countries. Indeed, doubts have long remained as to whether the British territories with high levels of education in the 20th century might have had a precolonial advantage. Although we do not claim to establish causal relationships (and hence avoid causal language), the age-heaping-based numeracy estimates provide substantial value-added in this respect because we can include evidence from before colonization, hence controlling for precolonial starting conditions.

Our work contributes to the literature on numerical skills, which are among the most crucial determinants of economic growth, today as well as in the past (Hanushek and Woessmann 2012; Baten and Juif 2014). Indeed, if workers are able to understand and process numerical information and if they can work with calendars and plan working hours, they are far more productive than an uneducated labor force. Importantly, numerical skills were also important in early agricultural economies. For example, the treatment of cattle disease and the protection of crops against insects and parasitic plants were more efficient in places where farmers were numerate (A'Hearn et al. 2009; Tollnek and Baten 2017). Concerning Africa, even in contexts that were not yet fully integrated into relevant trade networks, local farmers had to make decisions on producing, selling and buying specific quantities of different products, implying numerical skills and the will to pass them on to their offspring, as was the case in West Africa in the first decades of the 19th century (Fall 2010).

This analysis is also related to a substantial literature on colonial education legacies in Africa. Quantitative evidence obtained by Bolt and Bezemer (2009) has shown that the variation in growth across Sub-Saharan African countries is explained better by cross-country variation in colonial education than by legal origins and geography. The link between colonial-era education and socioeconomic wellbeing has been confirmed by the individual-level analyses of Nigeria (Okoye et al. 2016) and Benin (Wantchekon et al. 2015). Our work is also related to the influential studies about persistent challenges to development in Africa arising from the slave trade and its effects on trust (Nunn 2008; Nunn and Wantchekon 2011), from noninclusive institutions brought by colonial occupiers (Acemoglu et al. 2001), and from precolonial governance institutions (Michalopoulos and Papaioannou 2013). In the following sections, we first describe the data and sources, then discuss potential selectivity bias and other caveats before presenting the numeracy trends and finally the conditional correlations outlined above.

2. Data and sources

We estimate numeracy trends by using the ABCC index, a proxy for basic quantitative skills that relies on the assessment of age heaping within a sample or population. Numeracy, i.e., the capability to count and conduct basic calculations, is negatively correlated with the extent of heaping—defined as people's tendency to round their age to the nearest 0 or 5. This negative correlation has been shown to hold across samples, countries and historical periods. Thus, age heaping provides an innovative tool to explore trends of an important component of human capital in the distant past, typically when (and where) information on literacy and schooling is not available or not reliable (A'Hearn et al. 2009). The extent of age heaping in a sample of individuals within a specific age group can be summarized by the Whipple Index (WI in Equation 1):

$$(1) \qquad WI_{23-32} = \frac{\sum(n_{25}+n_{30})}{1/5\ \sum(n_{23}+n_{24}+\cdots....+n_{32}+)} \times 100$$

$$(2) \qquad WI_{63-72} = \frac{\sum(n_{65}+n_{70})}{1/5\ \sum(n_{63}+n_{64}+\cdots....+n_{72}+)} \times 100$$

The numerator of Equation 1 adds together the number of individuals in an age group (here 23 – 32) who report ages rounded on digits ending by zero and five (n_{25} and n_{30} in formula 1). This is divided by the total number of individuals in the same group (23 – 32) whose age should actually end with zero or five – which is obtained by multiplying the total number of individuals in the concerned group by 1/5 under the assumption that the distribution of cohorts across ages is uniform – and then multiplied by 100. The Whipple Index is calculated for the age groups 23 – 32, 33 – 42, 43 – 52, 53 – 62 and 63 – 72 (Equation 2 shows Whipple Index formula for the age group 63 – 72 to give another example). If there is no heaping, the Whipple Index will be equal to 100. As heaping increases, the index grows from 100 up to a maximum of 500 (only rounded ages). The requirement of an approximately uniform distribution of ages explains why the age brackets 23 – 32 etc. to age bracket 63 – 72 are used in the literature: alternative age brackets such as e.g. 60 – 69 would bias the index, as more individuals are alive at age 60 than at age 69 in high mortality environments (A'Hearn et al. 2009). Individuals below 23 years old are excluded since parents might declare the age of their children, even when these are teenagers. Clearly, 23 is a conservative threshold, which also relates to the requirement of a uniform age bracket. Additionally, individuals older than 73 years are excluded because the numerosity of the sample is greatly reduced due to high mortality rates (particularly within African countries), thus the Whipple Index may be biased.

The ABCC index is a linear transformation of the Whipple Index and ranges between 0 and 100 (from full to no heaping, respectively):

$$(3) \qquad ABCC = \left(1 - \frac{(WI-100)}{400}\right) \times 100$$

If a census or other source displaying the age of people from a specific country contains enough individuals between 23 and 72 years old, the ABCC can be calculated for the five groups shown above. Because we know when a census or other source was produced, each one of these age groups corresponds to a decade of birth. Since basic education is normally acquired during the first decade of a person's life, we can thus aggregate the ABCC by birth decades and countries, gaining the possibility to work with long time series of numeracy. For example, if ages for a sample of people between 33 and 42 years old were recorded via a census in 1900, we can estimate the numeracy of the same sample born in 1858 – 67. This is rounded to represent the birth decade of the 1860s.[1]

One potential objection to age heaping-based estimates for African history might be that the knowledge of the correct age was not important before colonization and that this became more relevant through contact with Europeans. This could potentially bias the results of our study, due to the fact that the relationship between colonial education systems and human capital accumulation would be spurious if it just depended on such institutional changes. For example, A'Hearn et al. (2019) warn that cultural preferences and state capacity might have a separate effect on age-heaping, and that both numeracy and literacy are common outcomes of the general development process that is partly determined by state capability and culture; despite this, they maintain that numeracy based on age heaping and other educational indicators, e.g. literacy, are correlated and can thus be used to shed light on human capital trends. We assess this in the following and argue that, in the context of our study, these objections do not make the numeracy evidence uninformative.

First, historically, assessing the precision of age statements has not been a crucial aim of public authorities—at least not until the rise of retirement schemes and social security (Soneji and King 2012). Since, within present-day Africa, such systems still cover only 5 to 10 percent of the population (Stewart and Yermo 2009), the age reported in past African censuses is unlikely to be biased due to double-checking by public officials and enumerators trying to limit cheating on actual age. Indeed, Nagi et al. (1973) have shown that heaping is a central feature of many postcolonial African censuses, arguing that ignorance of correct age and age heaping are the most common source of error in age reporting. Furthermore, Lyons-Amos and Stones

[1] See Appendix A for more information on the use of time series produced from birth cohorts. It is worth noting that Crayen and Baten (2010) suggested a downward adjustment of 25 percent for the people within the age group 23 to 32 since, within this range, many individuals also rounded on multiples of two – not only multiples of zero and five (e.g. on 28 rather than 30). The Whipple Index only captures rounding on multiples of zero or five, thus an adjustment is necessary. The one proposed by Crayen and Baten (2010) was widely used in the subsequent literature (for a survey, see Tollnek and Baten 2017).

(2017) find no evidence of increased accuracy in DHS surveys across African countries over the period of 1987–2015, specifically rejecting the hypothesis that age heaping would decline over time due to improving the techniques used to collect the data.

The function of age heaping as an index of numeracy in African surveys and censuses is consistent with the results provided by Crayen and Baten (2010), who explore the correlates of age misreporting in historical sources through regression analysis. Their results highlight that state antiquity and other institutional variables display a zero or marginal impact on age heaping once education and schooling are controlled for (Bockstette et al. 2002). This evidence, too, supports the argument that age heaping can be used as an index of human capital that is not substantially biased from the influence of institutional features.

Second, solid evidence showing that age heaping is a strong informative indicator of numeracy comes from historical periods in which the latter stagnated or declined for a long time. Typically, households did not invest in their children's education during periods of economic or political crisis because adverse conditions like e.g. malnutrition did not allow children to develop cognitive skills, and instability did not allow the attendance of schools. Mid-19th century China is an example of this: one can observe a strong decline of numeracy (based on age heaping) when the Taiping Rebellion brought about crisis and caused severe malnutrition among a broad segment of the population (Baten et al. 2010). Similarly, Latin-American countries in the early 19th century, during and after the conflicts of their independence struggle, saw declining or stagnating numeracy (Manzel et al. 2012).

Our numeracy dataset draws on three main sources.[2] First, concerning the period 1800–1970, we rely on census data included in official national statistics from the 20th century, as well as censuses conducted in the second half of the 19th century. Such data have been assessed by studies in economic history and have been found to be reliable for our aims (Crayen and

[2] See Data Appendix 1 for details on numeracy sources, and Data Appendix 2 for a discussion on how they allowed the reconstruction of numeracy according to present-day country boundaries. Data Appendix 3 reports a short summary of all variables used in this study.

Baten 2010). Most of these data come from censuses taken from the late-1940s to the present day. Given that numeracy is mainly determined in the first decade of life, we can study numeracy since the 1870s thanks to 20th-century censuses. In addition, we obtain long-term series of numeracy in Senegal from French archival colonial census data (Cappelli and Baten 2017). Numeracy figures for South Africa are obtained from the joint analysis of several consecutive censuses, stretching from the mid-19th century to the 1970s. Second, a small number of additional observations for the period of 1850–1900 were obtained from contemporary anthropological reports: Kenya in 1840, Togo and Tanzania in 1860, Sudan 1860–1870, Cameroon in 1870, and Somalia 1880–1890. The anthropologists coded males and females in very similar numbers and assessed many different regions within countries, since they were interested in ethnic diversity. Indeed, these sources are not socially or regionally biased to a substantial extent, as detailed studies found (Stegl and Baten 2009). The number of cases is small in these anthropological data collections, but they help to limit data scarcity in the mid-19th century. Third, we also complement this evidence with slave censuses compiled in the second half of the 18th and early 19th century, as to obtain numeracy series in the period 1730–1800. Midlo-Hall collected numerous sources of this kind, which were written when slaves from Africa were deported to the state of Louisiana, in the US. These sources include statements about their origin, which allows us to locate and assign them to present-day country boundaries. The Midlo-Hall data were merged with a dataset concerning Brazil, the Maranhão Inventories Slave Database (MISD), to reconstruct 18th-century numeracy for Angola, Benin, Congo, Ghana, Guinea, Guinea-Bissau, Mali, Mozambique, Nigeria, Senegal, Sudan and Togo (Midlo-Hall 2000; Hawthorne 2010).

This research effort has culminated in an unbalanced panel of ten-year-interval numeracy figures for 43 African countries from the 1730s to the 1960s. Data could not be provided at any point in time in the case of Comoros, the Democratic Republic of the Congo, Djibouti, Equatorial Guinea, Eritrea, Mauritania, Mauritius, Niger, Sao Tome and Principe, Seychelles and South Sudan (our sources refer to the territory of pre-partition Sudan). Table 1 shows the availability of numeracy data by country and period (non-missing values for each country-decade pair), as well as by main type of source (cells shaded in gray show census-based numeracy estimates), reporting the relative size of the sample (number of individuals) for each country-decade observation. Figure 1 visualizes data availability for each country and century in our study: dark-grey areas show countries for which numeracy data could be retrieved in the 18th (left-hand map), 19th (central map) and 20th century (right-hand side map). It is obvious that we have many observations for the years 1870-1970, but we also have a substantial number of observations for the 19th century (Table 1). The 18th century can be assessed using slave sources, although potential selectivity issues need to be discussed (see next section). The early and mid-19th century is the most poorly documented, although for South Africa, Senegal and Egypt, we have substantial evidence, and Cameroon, Kenya, Somalia, Sudan, Togo and Tanzania can at least be partly covered.

3. Assessing potential selectivity in historical sources

Concerning the data drawing on official censuses and national statistics, the selection bias is not expected to be substantial (Crayen and Baten 2010); this also applies to anthropological data collections, as argued above (Stegl and Baten 2009). The question remains as to whether the slave data that we employ for the 18th century are prone to strong selectivity.

Eltis (1982) argued that the difference between samples of freed slaves and the populations from which they were drawn was, in terms of health and social status, very limited. First, during the 19th century, slave labor was increasingly demanded by the growing economic activities within Africa, such as plantations and physically demanding transport-related tasks (e.g., canoeing and portering) – and these occupations required relatively healthy and strong workers. Second, Eltis shows that no premium can be observed in the data concerning the tallest (hence strongest) slaves, implying no substantial positive selection. Third, if e.g. positive selectivity characterized slavery, the distribution of the heights in the samples of adult slaves would necessarily be skewed towards taller stature ranges – which is not the case in Eltis' data (see Austin et al. 2012 for a similar conclusion).

We also compare numeracy obtained from our dataset on slaves deported to the Americas with the numeracy of the natives in the Cape Colony around the same time (obtained from Baten and Fourie 2015) to check whether non-self-reported ages bias our estimates to a large extent. Indeed, one issue could be that slaves did not self-report their age and that this was thus estimated by e.g. their owners. Since the court-register records used to reconstruct numeracy in the Cape Colony explicitly state that people were declaring their age themselves in court, this is an effective way to test whether our sources suffer from substantial non-self-reporting bias. The average numeracy in our slave dataset for the Southern part of Africa (Mozambique, Angola) is 35 in the period of 1750–1780. The numeracy of natives in South Africa ranged from approximately 25 (Khoesan) to approximately 30 (slaves) and up to approximately 35 (South African blacks, mostly Xhosa, see Figure 4 in Baten and Fourie 2015, p. 648). Although explorative, this check suggests that our numeracy estimates from deported slaves are probably not severely affected by self-reporting bias.

To consider the possibility that numeracy estimates from slaves are biased because of a preponderance of males in the sample, we also weight our estimates by the gender share. Women represent, on average, approximately 40 percent of the slave dataset. Therefore, the female (male) observations are assigned a correspondingly higher (lower) weight: the results are virtually unchanged when unweighted estimates are used.

Similarly, to assess the geography-related selectivity of our slave samples, we compare the regional distribution of numeracy in Western Sudan (today's Senegal) across two distinct datasets to test whether our estimates of numeracy suffer from specific bias. We find that numeracy calculated from local census lists is highly consistent with figures obtained from slave records (Figure 2, estimates drawing on census lists are from Cappelli and Baten 2017).

Cultural differences might also invalidate the use of age heaping for calculating numeracy: this technique assumes that basic ways of counting are broadly similar across different countries as well as across different cultures. We do not dismiss the argument that different populations and societies may value counting differently or count in different ways, with biases towards certain digits that are different from multiples of 5. For example, the Yoruba in Nigeria partly use a vigesimal counting system (based on multiples of 20, see Babarinde 2014). However, in earlier studies, cultural preferences for certain numbers were found to play only a very modest role in rounding behavior. The fact that most human beings start counting with their hands implies a strong tendency to round on multiples of 5, which has been shown by studies on the Inca-Indio culture and East Asia (Juif and Baten 2013; Baten et al. 2010). It is also worth noting that the strong correlations we found between numeracy and gross enrollment ratios in 1940 and between numeracy and average years of schooling in the same year (Figure 3) suggest that the ABCC index is a reasonable measure of educational outcomes, even within Africa.

4. Descriptive results

We first take a look at the overall numeracy trends (Figure 4). Numeracy improved since the late 18th century from a relatively low level, converging with low speed towards the higher levels in East Asia, the Americas and Europe (Crayen and Baten 2010, Cappelli and Baten 2017). The increase in numeracy accelerated in the 1830s, when the slave trade gave way to legitimate trades, and kept increasing at roughly the same pace until the 1880s, when numeracy started to improve at an increased speed. In the first two decades of the 20th century, despite having increased already, numeracy grew at the most rapid pace.[3] After the 1920s, the increase in numeracy slowed but remained steady.

[3] This cannot be due to numeracy figures drawing on different sources: starting with the 1870s, only census data (or very similar sources) were used.

Within this overall pattern, we can focus on country-specific trends, particularly those depicted by long time series during the 19th century: Senegal, Egypt, South Africa and Tanzania (also depicted in Figure 4). In Senegal, the stagnation of numeracy in the early 19th century gives way to a steady increase in the second half of the century (albeit not particularly fast compared to the overall trend). By contrast, in South Africa – where the colonial education and political system switched from Dutch to British in 1806 – the increase in numeracy was rapid since the early-19th century. In North Africa (and Egypt in particular), we observe a stagnation of numeracy at a very low level until the end of the 19th century. This stagnation around the mid-19th century and a later steady growth can be confirmed by evidence on individual birth decades relating to Tanzania. These country-specific trends are suggestive of a potential link between the increase in numeracy and the rise of colonial education, even though other factors were at play – for example, cash-crop trade (Cappelli and Baten 2017).

In the next section, we compare the increase in numeracy with a number of potential correlates, including colonization, cash-crop trade, the density of railways, and others, to isolate systematic patterns. By looking at the numeracy levels (not variations), one may note that African countries with high numeracy in 1900 tended to be characterized by a rather centralized political organization, better access to transport infrastructure (railways), a more limited incidence of slavery and a higher density of main mission stations, while basically no clear correlation emerges between numeracy and the tsetse suitability index or cash-crop trade (Figure 5). Due to this, we need to control for these factors in our explorative regressions, since increases in numeracy might be conditional on initial numeracy levels and on these variables.

In Table 2 we compare the average numeracy (not variations) in free and colonized territories, by African macro-region and period – i.e. dividing pre- from post-1900 years. Before the Scramble for Africa, average numeracy was higher in the few colonized territories than in noncolonized ones (Panel A in Table 2).[4] Looking at the post-1900 figures – when only Ethiopia[5] and Liberia were formally free from European occupation – numeracy levels were still higher in colonial education systems (Panel B in Table 2).

Descriptive evidence of the importance of colonial education systems for numeracy can also be discerned through border analysis – that is, a comparison of two bordering regions that are quite similar in terms of most economic, social and cultural features but differ concerning the colonial education system. We compare the region of Bench-Maji in Ethiopia (not colonized, except during the late 1930s) with Turkana County in Kenya, colonized by the British (Figure 6). The introduction of the British border in 1895 can even be regarded as a natural experiment because there was no obvious economic reason to draw the border in that particular place—in fact, even on the Ethiopian side, there was no ethnic or economic homogeneity, as Ethiopia had conquered these territories inhabited by various ethnic groups at the end of the 19th century.[6] Although numeracy estimates for the Ethiopian region of Bench-Maji are available only starting in 1920, the series shows a declining or, at best, stagnant numeracy up to the 1940s. By contrast, numeracy in Turkana County (south of the border, in Kenya) shows a clear positive linear trend since 1900 (see Appendix B for further evidence using the border analysis).

5. Conditional Correlations

We study whether the new evidence presented in this study is compatible with the hypothesis that African countries experienced more rapid increases in numeracy under education systems during the colonial period (hypothesis 1). Clearly, colonialism exacted a heavy toll in Africa in terms of casualties (Rummel 1997), forced labour (Waijenburg 2018)

[4] Since the link between the colonial education system and higher numeracy may be spurious due to the positive trend in numeracy and the (late) timing of colonization, we need to control for common trends affecting the increase in numeracy in the econometric analysis.
[5] Ethiopia was a free country until Italy's occupation at the end of the 1930s, and thereafter independent again.
[6] We cannot obtain long-term estimates of numeracy at the subnational level covering longer time frames.

and colonial extraction through trade (Tadei 2018), among others. However, one may put forward that, compared to precolonial times and non-colonized territories, a positive relationship might have existed between colonial education systems and numeracy, based on two main mechanisms. First, the effort to spread European schooling among the native population to consolidate and legitimate colonial power in the eyes of Africans, which pushed the metropoles to organize colonial schooling; and second, the fact that the systematic occupation of the African territory allowed missions (and thus mission schools) to expand their operations in the continent, from coastal areas to inner region (Jedwab, Selhausen, and Moradi 2019). Furthermore, the new export opportunities brought about by the colonial economy prompted an increasing demand for European-provided education: African farmers responded to changing incentives, thus African agency also prompted more rapid increase in numeracy under a new context (Frankema 2012).

We use a weighted least squares (WLS)[7] panel data model where the dependent variable is the average annual increase in numeracy ngr_{it} within each ten-year interval t concerning country i (1730–1970):

$$(4)\ ngr_{it} = \beta_0 + \beta_1 n_{it} + \beta_2 colony_{it} + \alpha X'_{it} + \delta_i + \gamma_t + \varepsilon_{it}$$

The main independent variable is a time-varying dummy ($colony_{it}$) capturing whether a country was subject to a colonial education system in period t (for the definition, we follow Henderson and Whatley 2014, see Appendix C for details). We are interested in the beginning of systematic colonization, not “outside contact” or “commercial contact”. Therefore, a country is defined as being colonized when the last territory within its present-day borders was occupied. Generally, this implied a substantial reduction of conflicts and political instability, allowing missions to operate in the territory with a lower risk of being compromised. Moreover, the ending of colonial-expansion conflicts also made it possible to organize colonial education

[7] Weights are the country-year population figures.

systems. Once it became clear that Europeans had come to stay in the medium or long term, the benefit of European education grew compared to its initial costs, raising the demand for education by Africans (Frankema 2012, p. 349).

n_{it} is the level of numeracy at the beginning of each period to control for convergence effects.[8] We always include country fixed effects and time-dummy variables to capture bias from omitted time-invariant factors and country-specific unobservables. In the most complete specifications, we add a vector X of control variables, including the average annual rate of population growth in the ten-year interval concerned, the urbanization rate, the intensity of international and domestic conflict in each period, cattle per capita and cropland per capita, railways per sq. km and cash-crop trade per capita (Appendix D discusses the motivation for including the controls as well as the underlying sources – see the complete list in Data Appendix 3). We also control for state capacity following the definition of Borcan et al. (2018), as this might have had a partial effect on age-heaping (A'Hearn et al. 2019).[9] Although there might be reverse causality from numeracy levels to colonial status, this potential issue does not necessarily relate to increases in numeracy, particularly when we condition on its initial level. Table 3 shows summary statistics.

Results are reported in Table 4, columns 1 to 3. In column 1, the estimated coefficient of colonial education systems is positive and statistically significant, being equal to a 1.79 percentage-point increase in the average annual increase in numeracy. We obtain similar results if we use the same equation and restrict the analysis to the period 1800–1970 (column 2). The estimated coefficient is quite large, as the mean annual increase in numeracy in our sample from 1730 to 1970 is equal to 1.19 percent (0.70 percent if post-1800). Column 3 includes the full set of control variables. The coefficients of urbanization, conflicts and agricultural production

[8] Included on the right-hand side of the equation to model the fact that education normally spreads following an S-shaped growth curve, i.e., increasing more slowly when it is already high.
[9] This aspect relates to the fact that more organized states could better manage and process information on their subjects.

exhibit the expected signs, although they are mostly statistically insignificant. The rise of cash crops and trade networks have a negative sign, which might not be the expected one. Bauer argued that cash crop trade was largely beneficial for West Africa (Bauer 1965, p. 104). However, other studies highlight that the relationship between cash crops and development depends on local conditions and can be negative (Kennedy, Bouis, and von Braun 1992; Cappelli and Baten 2017). The same can be posited for railways. A substantial literature argues that cash-crop exports replace high-quality food available to the poor local population in the first phase of transport-infrastructure expansion (Baten and Fertig 2010). Only later are the benefits of the railway likely to be better distributed.

We also consider potential heterogeneity between different colonial education systems, focusing on the British and the French against the others (hypothesis 2). The openness of the British system towards missionary schools and the use of native languages, particularly in rural areas, was possibly linked to a more rapid increase in numeracy (Frankema 2012; Cogneau and Moradi 2014; Dupraz 2019). Indeed, Carvalho and Koyama (2016) argue that the adoption of local languages by missionaries and authorities in British Africa reduced the cultural distance between the metropole and the native populations, and that native Africans responded to new incentives and returns to European-style education in the context of the colonial economy. By contrast, the French applied a policy of strong assimilation in areas where returns to education were not clear for the natives – apart from a few coastal regions such as those in today's Senegal. Last but not least, the local-taxation system to fund education and other public investments was also identified as one peculiar element that might differentiate the British territories from areas occupied by other metropoles: local taxes were partly used for education, and some elements of local monitoring existed (Bolt and Gardner 2019).

In columns 1 and 2 in Table 5, the positive within-country correlation between the increase in numeracy and colonial status appears to be related to the performance of the British education system. The British coefficient is positive with a magnitude similar to that of the mean effect estimated in Table 4. The coefficient for French colonies is not statistically significant, and if anything, the effect appears to have been negative throughout the period of 1800–1970. The results are robust if we include only the census evidence (dropping slave based and anthropological evidence, see Appendix E) and if we adjust for the boundary effect that numeracy cannot be higher than 100 percent (see Appendix F). All in all, this evidence is consistent with the view that the choices made by the British administration – not only laissez-faire towards missions but also active cooperation with the latter and subsidized missionary schooling – and missionary activity largely run by Africans – i.e. the Africanization of the mission and the reliance on local resources and languages – was associated with a more rapid expansion of schooling, and hence numeracy (Meier zu Selhausen 2019). Indeed, qualitative and micro evidence on the relationship between colonial administrators from various metropoles and missions suggests that the British policy brought about a sustained development of colonial schooling (see Appendix H).

6. Conclusions

We provide new estimates of numeracy for an unbalanced panel of African countries from 1730 to 1970. The age-heaping technique and the reliance on new sources—such as slave records and untapped census data—allow a reconstruction of trends in numeracy on the continent before, during and following the Scramble for Africa.

The main contribution of this study is to show that numeracy evidence can fill important gaps in our knowledge of the long-term development of the African continent. Naturally, data spanning three centuries of development cannot be constructed without raising potential issues about selectivity and measurement errors. We analyzed potential biases and measurement issues in great detail, concluding that our index can be used as a first step towards tracing the history of increases in African numeracy in since the early 18th century, which has important implications for development economics.

Secondly, we rely on conditional correlations to explore whether the British colonial education system – which relied on African teachers and local languages to a substantial extent – was correlated with a more rapid numeracy development compared to other education systems in Africa. Colonial rule was clearly a catastrophe for many dimensions of African development. Concerning numeracy, though, our evidence suggests that the choices made by the British administration – not only laissez-faire towards missions but also active cooperation with the latter and subsidized missionary schooling – and missionary activity largely run by Africans – i.e. the Africanization of the mission and the reliance on local resources and languages – was correlated with the expansion of numeracy. This finding is consistent with the recent literature on African long-term growth and economic history, which stresses that African agency mattered for the continent's development more than has been commonly acknowledged, even in the recent past.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

A'Hearn, Brian, Jörg Baten, and Dorothee Crayen. 2009. "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital." *The Journal of Economic History* 69 (3): 783–808.

A'Hearn, Brian, Alexia Delfino, and Alessandro Nuvolari. 2019. "Cognition, Culture, and State Capacity: Age-Heaping in XIX Century Italy." 14261. *CEPR Discussion Papers*. CEPR Discussion Papers. C.E.P.R. Discussion Papers.

Allen, Robert C. 2011. *Global Economic History: A Very Short Introduction*. Very Short Introductions. Oxford, New York: Oxford University Press.

Alsan, Marcella. 2015. "The Effect of the TseTse Fly on African Development." *American Economic Review* 105 (1): 382–410.

Austin, Gareth, Joerg Baten, and Bas Van Leeuwen. 2012. "The Biological Standard of Living in Early Nineteenth-Century West Africa: New Anthropometric Evidence for Northern Ghana and Burkina Faso." *The Economic History Review* 65 (4): 1280–1302.

Babarinde, Olusanmi. 2014. "Linguistic Analysis of the Structure of Yoruba Numerals." *Language Matters* 45 (1): 127–47.

Baten, Joerg, and Georg Fertig. 2010. "Did the Railway Increase Inequality? A Micro-Regional Analysis of Heights in the Hinterland of the Booming Ruhr Area during the Late 19 Th Century." *Journal of European Economic History* 38: 263–99.

Baten, Joerg, and Johan Fourie. 2015. "Numeracy of Africans, Asians, and Europeans during the Early Modern Period: New Evidence from Cape Colony Court Registers." *The Economic History Review* 68 (2): 632–56.

Baten, Joerg, and Dácil Juif. 2014. "A Story of Large Landowners and Math Skills: Inequality and Human Capital Formation in Long-Run Development, 1820–2000." *Journal of Comparative Economics*, Economic Systems in the Pacific Rim Region Symposium, 42 (2): 375–401.

Baten, Joerg, 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*, Asian Economic Growth and Development, 47 (3): 347–59.

Bauer, P. T. 1965. *Economic Analysis and Policy in Underdeveloped Countries.* Routledge Library Editions. Durham, NC.

Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth* 7 (4): 347–69. https://doi.org/10.1023/A:1020827801137.

Bolak Funteh, Mark. 2015. "Dimensioning Indigenous African Educational System: A Critical Theory Divide Discourse." *International Journal of Humanities and Social Science* 5 (4): 139–50.

Bolt, Jutta, and Dirk Bezemer. 2009. "Understanding Long-Run African Growth: Colonial Institutions or Colonial Education?" *The Journal of Development Studies* 45 (1): 24–54.

Bolt, Jutta, and Leigh Gardner. 2019. "African Institutions under Colonial Rule." 14198. *CEPR Discussion Papers*. CEPR Discussion Papers. C.E.P.R. Discussion Papers.

Borcan, Oana, Ola Olsson, and Louis Putterman. 2018. "State History and Economic Development: Evidence from Six Millennia." *Journal of Economic Growth* 23 (1): 1–40.

Cameron, A. Colin, and Douglas l. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *The Journal of Human Resources 50 (2): 317-371.*

Cappelli, Gabriele, and Joerg Baten. 2017. "European Trade, Colonialism, and Human Capital Accumulation in Senegal, Gambia and Western Mali, 1770–1900." *The Journal of Economic History* 77 (3): 920–51.

Carvalho, Jean-Paul, and Mark Koyama. 2016. "Resisting Education." Working Paper. MPRA Paper No. 48048.

Cogneau, Denis, and Alexander Moradi. 2014. "Borders That Divide: Education and Religion in Ghana and Togo Since Colonial Times." *The Journal of Economic History* 74 (3): 694–729.

Crayen, Dorothee, and Joerg Baten. 2010. "Global Trends in Numeracy 1820-1949 and Its Implications for Long-Term Growth." *Explorations in Economic History* 47 (1): 82–99.

Dupraz, Yannick. 2019. "French and British Colonial Legacies in Education: Evidence from the Partition of Cameroon." *Journal of Economic History* 79 (3): 628–68.

Dwight, Henry Otis. 1905. *The Blue Book of Missions for 1905*. New York, London, Funk & Wagnalls company.

Eltis, David. 1982. "Nutritional Trends in Africa and the Americas: Heights of Africans, 1819-1839." *The Journal of Interdisciplinary History* 12 (3): 453–75.

Fall, B. 2010. "Sénégal: Le Travail Au XXè Siècle."

Frankema, Ewout. 2012. “The Origins of Formal Education in Sub-Saharan Africa: Was British Rule More Benign?” *European Review of Economic History* 16 (4): 335–55.

Frankema, Ewout, and Morten Jerven. 2014. “Writing History Backwards or Sideways: Towards a Consensus on African Population, 1850–2010.” *The Economic History Review* 67 (4): 907–31.

Frankema, Ewout, Jeffrey Williamson, and Pieter Woltjer. 2018. “An Economic Rationale for the West African Scramble? The Commercial Transition and the Commodity Price Boom of 1835–1885.” *The Journal of Economic History* 78 (1): 231–67.

Hanushek, Eric A., and Ludger Woessmann. 2012. “Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation.” *Journal of Economic Growth* 17 (4): 267–321.

Hawthorne, Walter. 2010. *From Africa to Brazil: Culture, Identity, and an Atlantic Slave Trade, 1600–1830*. Cambridge University Press.

Henderson, Morgan, and Warren Whatley. 2014. “Pacification and Gender in Colonial Africa: Evidence from the Ethnographic Atlas.” 61203. MPRA Paper. University Library of Munich, Germany.

Jedwab, Remi, and Alexander Moradi. 2016. “The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa.” *The Review of Economics and Statistics* 98 (2): 268–84.

Jedwab, Remi, Felix Meier zu Selhausen, and Alexander Moradi. 2019. “The Economics of Missionary Expansion: Evidence from Africa and Implications for Development.” Working Paper. IIEP-WP-2019-10 Elliott School of International Affairs, The George Washington University.

Juif, Dácil-Tania, and Joerg Baten. 2013. “On the Human Capital of Inca Indios before and after the Spanish Conquest. Was There a ‘Pre-Colonial Legacy’?” *Explorations in Economic History* 50 (2): 227–41.

Kennedy, Eileen, Howarth Bouis, and Joachim von Braun. 1992. “Health and Nutrition Effects of Cash Crop Production in Developing Countries: A Comparative Analysis.” *Social Science & Medicine* 35 (5): 689–97.

Lyons-Amos, Mark, and Tara Stones. 2017. “Trends in Demographic and Health Survey Data Quality: An Analysis of Age Heaping over Time in 34 Countries in Sub Saharan Africa between 1987 and 2015.” *BMC Research Notes* 10 (December).

Manzel, Kerstin, Joerg Baten, and Yvonne Stolz. 2012. “Convergence and Divergence of Numeracy: The Development of Age Heaping in Latin America from the Seventeenth to the Twentieth Century1.” *The Economic History Review* 65 (3): 932–60.

Meier Zu Selhausen, Felix. 2019. “Missions, Education and Conversion in Colonial Africa.” In *Globalization and The Rise of Mass Education*, edited by David Mitch and Gabriele Cappelli. London: Palgrave Macmillan.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. “Pre-Colonial Ethnic Institutions and Contemporary African Development.” *Econometrica* 81 (1): 113–52.

Midlo-Hall, Gwendolyn. 2000. *The African Diaspora in the Americas: Regions, Ethnicities, and Cultures*. Chapel Hill: University of North Carolina Press.

Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.1 [dataset]. Minneapolis, MN: IPUMS, 2018.

Mitchell, B. R. (Brian R.). 2007. *International Historical Statistics: Europe, 1750-2005*. Sixth edition. Palgrave Macmillan.

Mosweunyane, Dama. 2013. “The African Educational Evolution: From Traditional Training to Formal Education.” *Higher Education Studies* 3 (4): p50.

Nagi, M. H., E. G. Stockwell, and L. M. Snavley. 1973. “Digit Preference and Avoidance in the Age Statistics of Some Recent African Censuses: Some Patterns and Correlates.” *International Statistical Review / Revue Internationale de Statistique* 41 (2): 165–74.

Nunn, Nathan. 2008. “The Long-Term Effects of Africa’s Slave Trades.” *The Quarterly Journal of Economics* 123 (1): 139–76.

———. 2020. “The Historical Roots of Economic Development.” *Science* 367 (6485).

Nunn, Nathan, and Leonard Wantchekon. 2011. “The Slave Trade and the Origins of Mistrust in Africa.” *American Economic Review* 101 (7): 3221–52.

Okoye, Dozie, Roland Pongou, and Tite Yokossi. 2016. “On the Dispensability of New Transportation Technologies: Evidence from Colonial Railroads in Nigeria.” Working Paper 1620E. University of Ottawa, Department of Economics.

Rummel, Rudolph J. 1997. *Statistics of Democide: Genocide and Mass Murder Since 1900*. Rutgers, NJ: Transaction Publishers.

Soneji, Samir, and Gary King. 2012. “Statistical Security for Social Security.” *Demography* 49 (3): 1037–60.

Stegl, Mojgan, and Joerg Baten. 2009. “Tall and Shrinking Muslims, Short and Growing Europeans: The Long-Run Welfare Development of the Middle East, 1850–1980.” *Explorations in Economic History*, Special Issue on Heights and Human Welfare, 46 (1): 132–48.

Stewart, Fiona, and Juan Yermo. 2009. "Pensions in Africa." OECD Working Papers on Insurance and Private Pensions, No. 30, OECD Publishing, Paris.
Tadei, Federico. 2018. "The Long-Term Effects of Extractive Institutions: Evidence from Trade Policies in Colonial French Africa." *Economic History of Developing Regions* 33 (3): 183–208.
Tollnek, Franziska, and Joerg Baten. 2017. "Farmers at the Heart of the 'Human Capital Revolution'? Decomposing the Numeracy Increase in Early Modern Europe†." *The Economic History Review* 70 (3): 779–809.
Waijenburg, Marlous van. 2018. "Financing the African Colonial State: The Revenue Imperative and Forced Labor." *The Journal of Economic History* 78 (1): 40–80.
Wantchekon, Leonard, Marko Klašnja, and Natalija Novta. 2015. "Education and Human Capital Externalities: Evidence from Colonial Benin." *The Quarterly Journal of Economics* 130 (2): 703–57.
White, Bob W. 1996. "Talk about School: Education and the Colonial Project in French and British Africa, (1860-1960)." *Comparative Education* 32 (1): 9–25.

Tables

Table 1 – (continues on the next page): size of each country-period sample used to calculate numeracy across African countries throughout 1730 – 1970 (empty cells are missing values).

Country	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960
Algeria															77822	261662	835386	1338465	1793366	2358220	1590000			
Angola			68	175	277	192	112																	
Benin		39	76	94	42													24387	93116	133157	201415	289841	474180	
Botswana																6875	15882	25758	57311	91986	125582	77375	122142	191885
Burkina Faso																		58596	227264	376072	520953	697642	983266	
Burundi																			68218	126524	208342	289673	506879	826666
Cameroon			46	83	109	79									54		78107	172630	352113	559962	783479	983233		
Cape Verde																			7655	10842	18993	16842	21619	48305
CAR																		30162	118531	228413	336322	410366		
Chad																							773	2854
Congo	64	126	182	295	345	483	167																	
Cote d'Ivoire																			59074	186140	394245	641410	1034246	1762942
Egypt								112	166	2484	5245	9395	15537	160307	448732	911142	1731326	2483161	2963330					
Ethiopia																			99779	158314	231459	354786	421567	
Gabon																							205	1225
Gambia																	5953	11337	20500	33150	53751	89202		
Ghana																		434724	726314	1260501	2017223	1291483		
Guinea	31	108	173	219	138	40													15007	23237	34050	47726	69679	
Guinea-Bissau			55	137	223	125									4470	10974	22402	44077	71832	96884				
Kenya												53				74726	262272	570668	1413207	2347536	3777164	3797298	3838217	3126275
Lesotho																			199	572	1069	1747	2957	
Liberia																	31630	87230	155948	236617	353792	240683		
Libya																	29109	56297	82665	138840	185543	235327		
Madagascar																	83545	220589	409787	595266	728035	948105		
Malawi																			22118	33693	48705	77588	115117	
Mali		35	92	116	47												71417	167233	308729	434918	644528	922529		
Morocco																134260	541317	1077923	1507302	2418564	3540039			
Mozambique				37	43	27														36023	63159	100813	151900	228593

(*Cont'd*) Country	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960
Namibia																6412	5813	2777	2499	1700				
Nigeria		55	130	193	127	142	36									329022	645469	1279842	2506681	5195384	10900000			
Rwanda																				18638	29548	11954	21821	47890
Senegal	39	154	340	373	219	199	48	365	585	283	56	49	59	84	105	117	149				11031	24805	37887	63645
Sierra Leone																					12134	18303	32064	53329
Somalia																106	72							
South Africa								179	283	14284	34783	59900	98041	155082	252423	251803	405009	746416	1206860	1603278	81028	111245	329863	457385
Sudan			40	48	56									62	67							102094	180524	290099
Swaziland																3436	12300	21650	38973	60561	93703	59904	92178	
Togo	31	83	188	304	169	97								42		19245	53590	101604	170030	270869	432563	279586		
Tunisia																45102	122006	248254	342205	499451	595009			
Uganda																3171	1738	965	424	227	99	84273	125449	237057
Tanzania														32						60935	95353	138276	204199	341074
Zambia																	43493	151917	331380	15868	30825	45657	66373	115261
Zimbabwe																						16985	27932	37257

Note: grey cells represent figures obtained from census data. Sources: see Data Appendix 1.

Table 2 – Mean numeracy rates across African regions, pre and post 1900

PANEL A: Pre 1900

African Regions	Free Regions	Colonized Regions
East	52	
Central	32	47
North	21	52
South	69	89
West	31	57
Total	37	57

PANEL B: Post 1900

African Regions	Free Regions	Colonized Regions
East	62	80
Central	64	76
North	54	71
South	90	91
West	71	72
Total	69	77

Notes to Panel A. The number of countries/birth decade observations pre-1900 is, (a) for the free countries: in East Africa 13, in Central Africa 17, in North Africa 18, in Southern Africa 15 and in West Africa 53 for a total of 116 observations. Meanwhile, the number of countries/birth decade observations pre-1900 is (b) for colonized countries: in East Africa not available, in Central Africa 1, in North Africa 5, in Southern Africa 1 and in West Africa 5, thus a total of 12 observations.

Notes to Panel B. The number of countries/birth decade observations post 1900 is, (a) for the free countries: in East Africa 5, in Central Africa 1, in North Africa 5, in Southern Africa 5, in West Africa 8. A total of 24 observations. Meanwhile, the number of countries/birth decade observations post 1900 is (b) for colonized countries: in East Africa 55, in Central Africa 13, in North Africa 17, in Southern Africa 23, in West Africa 62. A total of 170 observations.

Table 3 – Summary statistics for the panel data, 1730 – 1970.

Variable	**N**	**Mean**	**SD**	**Min**	**Max**
Annual increase, numeracy (%)	263	1.19	3.01	-5.6	26.67
Initial numeracy (0-100)	321	61.31	26.54	5.00	100
Colony (HW code)	983	0.30	0.46	0.00	1.00
Annual growth population	703	0.90	0.87	-0.98	4.20
Urbanization rate	669	0.07	0.07	0.00	0.51
Intensity of int. conflicts	686	0.09	0.21	0.00	1.00
Intensity of dom. conflicts	686	0.03	0.14	0.00	1.00
Cattle p.c.	669	0.37	0.50	0.00	3.24
Cropland (sq. km) p.c.	669	0.61	0.34	0.09	1.58
Cash-crop-trade value per capita (1,000 2018 GBP)	983	5.80	23.14	1.00	323.45
Km railways per sq. km (Mitchell)	982	0.00	0.01	0.00	0.41
State capacity	983	28.01	13.05	0.00	50.00

Note: The data are characterized by little variation in terms of colonial rule before 1900: Cape Verde is the only country in our sample to be coded as colonized between 1730 and 1870; Algeria and Senegal are coded as completely occupied by European countries starting in 1870, while other countries were colonized from c. 1890 on.

Table 4 – FE panel-data model.

Dep. Var.: Av. annual increase in numeracy	(1) 1730 – 1970	(2) Post 1800	(3) Post 1800 w/controls
Initial Numeracy	0.0244	0.0054	0.0145
	(0.044)	(0.059)	(0.068)
Colony	1.7861***	1.7178***	1.9345**
	(0.579)	(0.529)	(0.930)
Urbanization rate			5.2249
			(5.326)
Intensity of int. conflicts			-1.9115*
			(1.103)
Intensity of dom. conflicts			-1.7517
			(1.825)
Cattle p.c.			0.3123
			(0.652)
Cropland (sq. km) p.c.			1.0411
			(1.926)
Annual growth population			0.8445***
			(0.285)
Cash-crop-trade value per capita (log)			-0.6835**
			(0.264)
State capacity			-0.0039
			(0.031)
Km railways per sq. km (Mitchell)			-6.0616**
			(2.630)
Observations	263	216	209
Country & Time FE	Y	Y	Y
N. of countries	43	41	39
Adjusted R-squared	0.376	0.413	0.476

Notes: Robust standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Although Cameron and Miller (2015) suggest that it is not necessary to compute bootstrapped standard errors with a limited number of clusters, we have calculated them. The results remain virtually the same. Data sources and their calculation: Data on cropland per capita, cattle per capita and domestic as well as international conflicts are drawn from the Clio-Infra dataset. To have population estimates going back to 1800, we expand the Frankema-Jerven dataset backwards by means of linear extrapolation (Frankema and Jerven 2014). This dataset is state-of-the-art concerning the population in precolonial and colonial Africa. In any case, the overall index of correlation between these figures and the population figures from the Clio-Infra data in the pooled sample is a reassuring 0.95. Cash-crop trade is coded from the dataset provided by Frankema et al. (2015), while the density of the railways has been reconstructed by relying on data from Mitchell (2007) and Jedwab and Moradi (2016). The intensity of international and domestic conflict in each period is measured as share of years characterized by a major conflict within the relative decade. Cattle per capita and cropland per capita, railways per sq. km: see Data Appendix. State capacity is based on Borcan, et al. (2018), see Data Appendix 3. The log of the monetary value of cash crop trade per capita is measured in thousands of GBP in real prices for 2018 (see also Appendix D for details about motivation for including the controls). N declines from 216 to 209 due to missing values for the controls. Estimates for the control variables are only available for the period of 1800–1970, making a regression with controls possible only for the 19th and 20th centuries.

Table 5 – FE panel-data model with colonial-education-system specific effects.

Dep. Var.: Av. annual increase in numeracy	(1) Post 1800 w/controls + interactions	(2) Post 1900 w/controls + interactions
Initial Numeracy	0.0274	0.1317***
	(0.081)	(0.032)
Colony	0.6738	-1.4518***
	(1.051)	(0.337)
Colony x UK	1.4921*	2.6671***
	(0.826)	(0.692)
Colony x France	0.2033	-1.2719
	(1.081)	(0.754)
France	-0.4094	0.0430
	(1.319)	(0.854)
Urbanization rate	4.9982	3.2412*
	(5.525)	(1.885)
Intensity of int. conflicts	-1.6827	1.2249
	(1.256)	(0.828)
Intensity of dom. conflicts	-1.9950	-0.3845
	(1.570)	(1.175)
Cattle p.c.	0.1846	-0.8946*
	(0.631)	(0.529)
Cropland (sq. km) p.c.	1.2840	0.0898
	(2.198)	(0.948)
Annual growth population	0.8152***	0.1920
	(0.222)	(0.151)
Cash-crop-trade value per capita (log)	-0.6706**	-0.0197
	(0.269)	(0.150)
State capacity	-0.0099	0.0091
	(0.031)	(0.034)
Km railways per sq. km (Mitchell)	-4.4371	2.5405
	(3.125)	(1.714)
Observations	209	167
Country & Time FE	Y	Y
N. of countries	39	38
Adjusted R-squared	0.473	0.540

Notes: Robust standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Although Cameron and Miller (2015) suggest that it is not necessary to compute bootstrapped standard errors with a limited number of clusters, we have calculated them. The results remain virtually the same. Data sources and their calculation: Data on cropland per capita, cattle per capita and domestic as well as international conflicts are drawn from the Clio-Infra dataset. To have population estimates going back to 1800, we expand the Frankema-Jerven dataset backwards by means of linear extrapolation (Frankema and Jerven 2014). This dataset is state-of-the-art concerning the population in precolonial and colonial Africa. In any case, the overall index of correlation between these figures and the population figures from the Clio-Infra data in the pooled sample is a reassuring 0.95. Cash-crop trade is coded from the dataset provided by Frankema et al. (2015), while the density of the railways has been reconstructed by relying on data from Mitchell (2007) and Jedwab and Moradi (2016). The intensity of international and domestic conflict in each period is measured as share of years characterized by a major conflict within the relative decade. Cattle per capita and cropland per capita, railways per sq. km: see Data Appendix. State capacity is based on Borcan et al. (2018), see Data Appendix 3. The log of the monetary value of cash crop trade per capita is measured in thousands of GBP in real prices for 2018 (see also Appendix D for details about motivation for including the controls). N declines from 216 to 209 due to missing values for the controls. Estimates for the control variables are only available for the period of 1800–1970, making a regression with controls possible only for the 19th and 20th centuries. The coefficient for the dummy relating to British territories is dropped in this model due to high collinearity with the fixed effects and the other Colony and Colony x UK interaction terms.

Figures

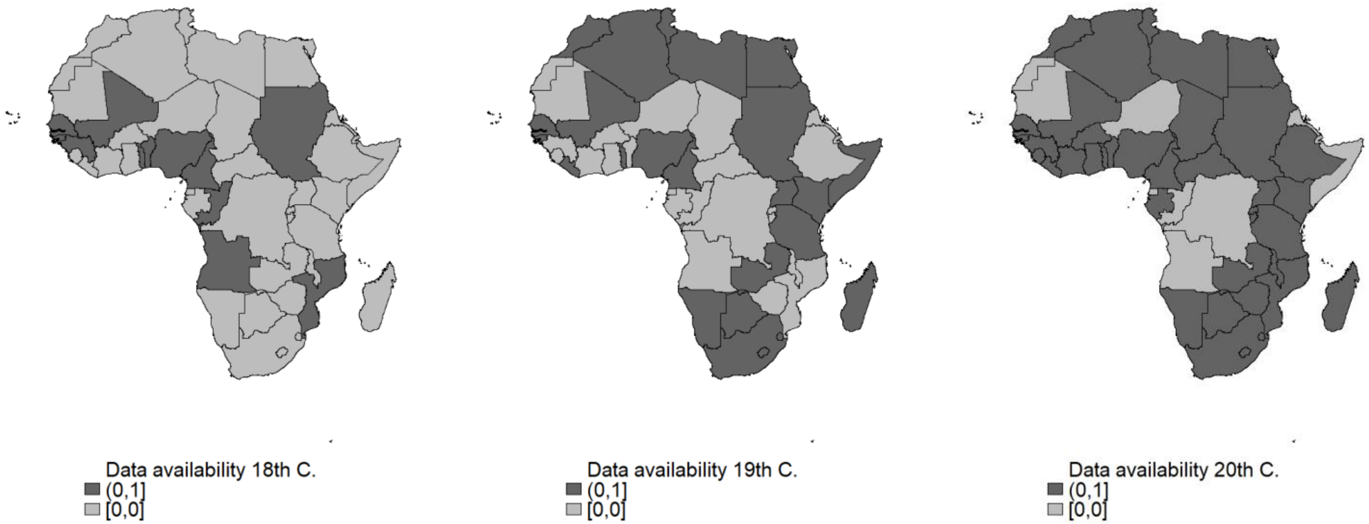


Figure 1: Data availability by country and period (18th century (left-hand side), 19th century (centre) and 20th century (right-hand side). Sources: see Data Appendix 1.

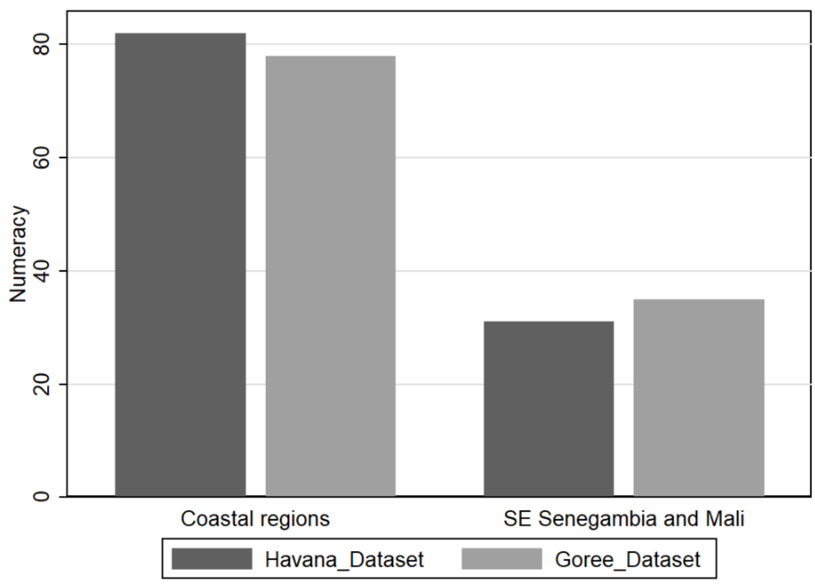
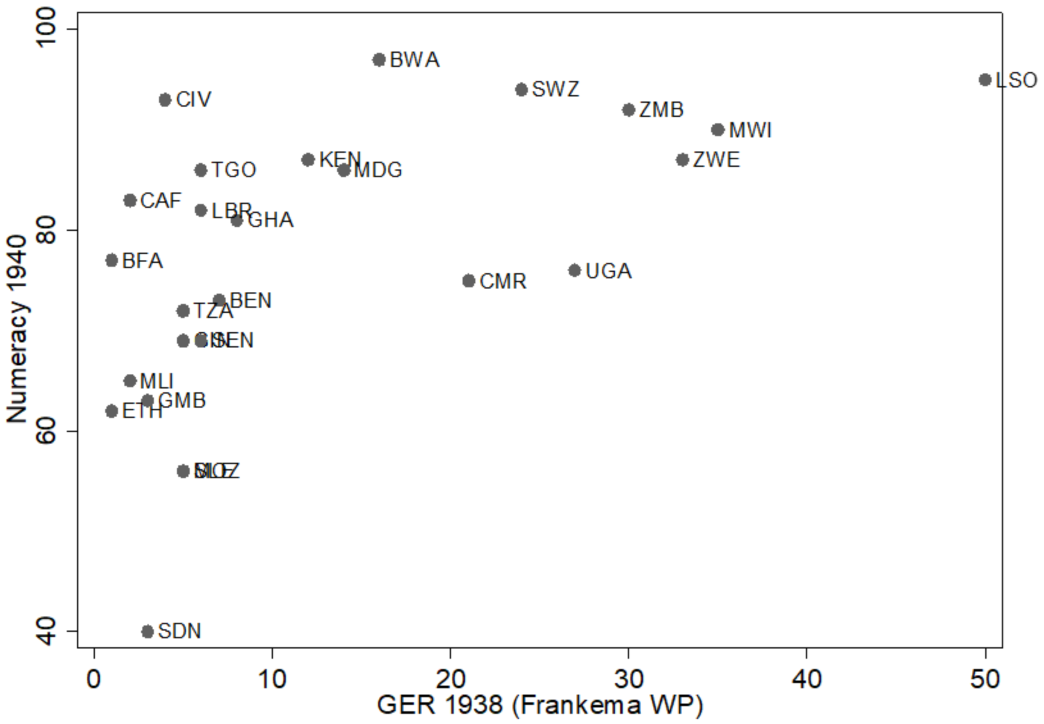


Figure 2: An assessment of potential selectivity through estimates of the numeracy of Africans (in %) from two regions and two different sources: Havana (a slave census at the port) and Gorée (a population census).

Sources: "Havana_Dataset" refers to a dataset on the shipment of African slaves to the port of Havana in the early nineteenth century (1825 – 1840), elaborated from the Transatlantic Slave Trade Database. Two main categories of individuals are considered: "coast" refers to slaves shipped from the port of Bissau (Portuguese Guinea) while "interior" represents slaves belonging to Mandingo tribes, who inhabited an area that stretched from today's southwest Mali to southeast Senegal. The "Goree_Dataset" refers to a different dataset on numeracy across regions of Western Sudan (see Cappelli and Baten 2017). Here, "coast" refers to Muslim individuals born in Senegal, in the coastal cities called the Four Communes, while "interior" refers to a sample of Muslims born in French Sudan, today's Kayes, in Mali.

Panel A



Panel B

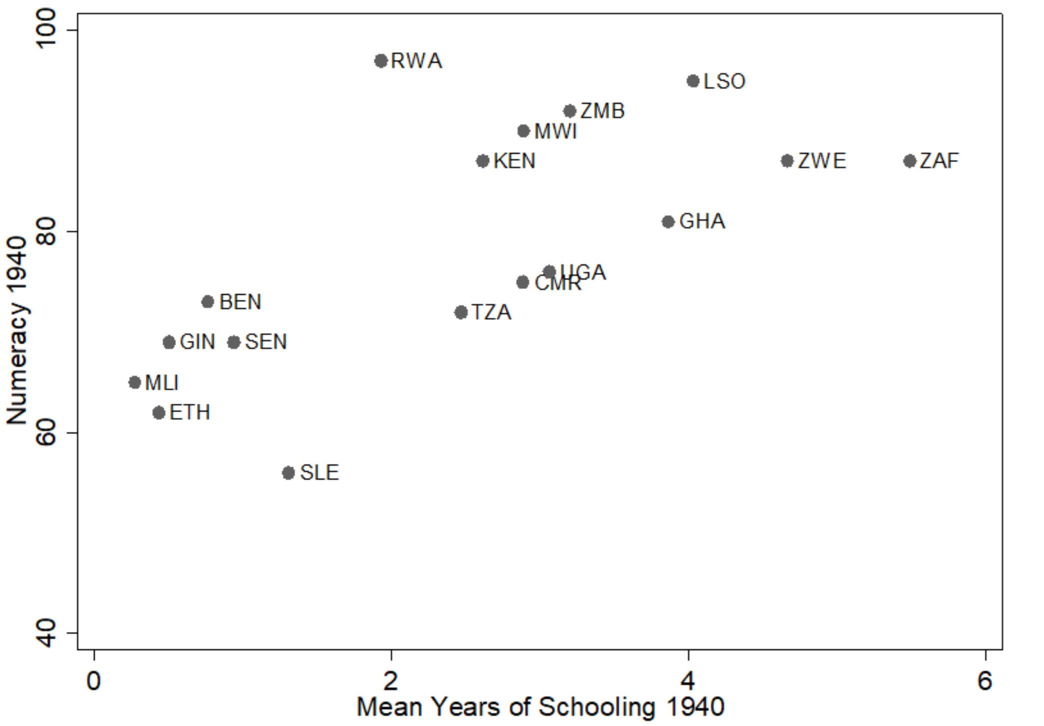


Figure 3: Numeracy compared with other educational indicators

Notes: Panel A shows average numeracy (%) in 1940 plotted against the Gross Enrolment Ratios (in %) in 1938 across African countries. Panel B shows the same correlation between numeracy (%) in 1940 and Average Years of Schooling. Source: own elaboration from primary sources and from Frankema (2012). Data on Average Years of Schooling are from IPUMS international data (Minnesota Population Center 2018).

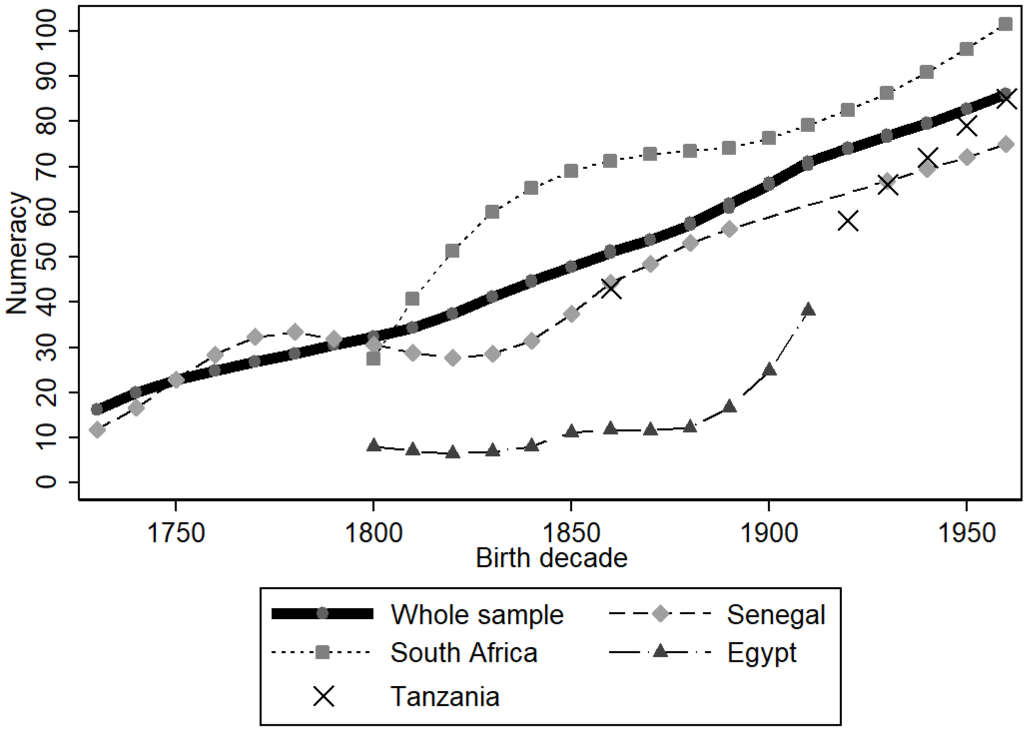


Figure 4 – Increase in numeracy over time, 1730 – 1970 (lowess smoother). Sources: see text.

Notes: The trends for Senegal, South Africa, Egypt and Tanzania are examples. The whole data covering 43 countries is included in the "whole sample" lowess-smoothed line. Source: Data Appendix 1.

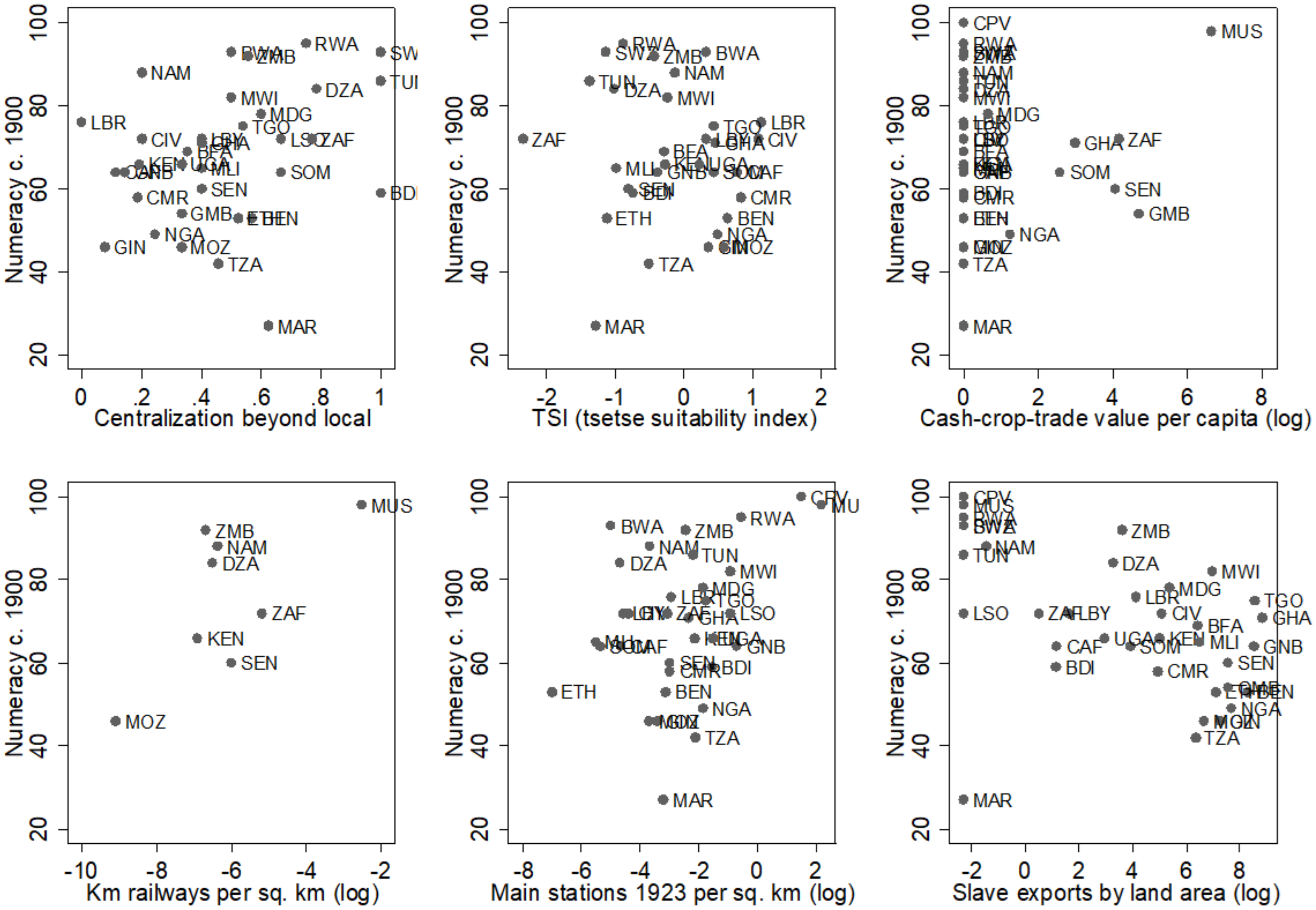


Figure 5 – Correlations between numeracy in 1900 and selected variables. Sources: see Data Appendix 1 and 3.

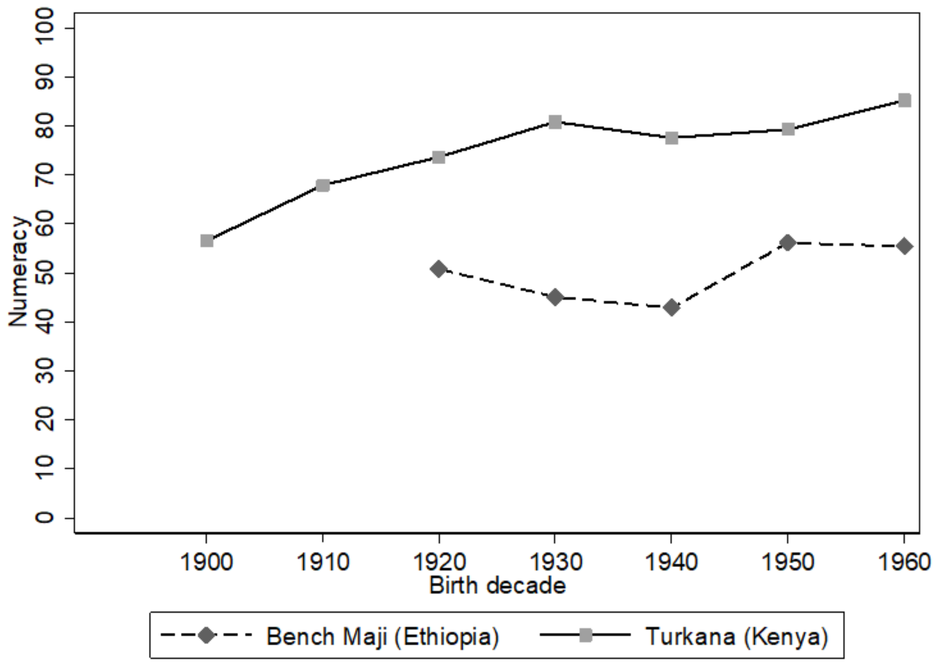


Figure 6 – Border analysis, c. 1900 – 1960. Notes: the graph shows the border analysis between the Bench-Maji zone of Ethiopia (free apart from the short Italian occupation at the end of the 1930s) and the Turkana county in Kenya (colonized by the British). Numeracy, as previously discussed, is estimated by organizing the data by birth decades; the sources are censuses as reported by IPUMS international data (Minnesota Population Center 2018).

Note: the result does not change if one chooses a different zone of Ethiopia that is more in the southeast (South Omo) or a different county in Kenya that is more in the northeast (Marsabit) (Appendix B). In sum, the border analysis of Kenya—suffering occupation from British colonialism but benefitting from its education system—and mostly independent Ethiopia suggest that the former realized more rapid increases in numeracy. The same holds true when changing the two bordering countries (e.g., Liberia vs Guinea).

(NOT FOR PUBLICATION – ONLINE SUPPLEMENTARY MATERIAL)

Data Appendix 1

Numeracy estimates for African countries by birth decade, 1730 – 1970

The age-heaping technique: the quality of the estimates for Africa

Age data frequently display excess frequencies on attractive numbers, such as multiples of zero and five. We exploit this phenomenon to measure cognitive ability in quantitative reasoning, i.e. numeracy. We construct a database of age-heaping-based estimates of basic numeracy with exceptional geographic and temporal coverage. The index that we use, the ABCC index, is a linear transformation of the Whipple index that is commonly used to capture age heaping. The reliability of our estimates is solid. By using the ClioInfra quality code – also employed by the OECD in the publication "How was life?" –, none of the ABCC estimates equals 1 (i.e. from official governmental statistics) or 4 (conjecture or guesstimate). However, all the ABCC values that are based on UN Demographic Yearbooks (from the late-19th century to the post-WWII period) or that show the word "census" in the title referenced should obtain a score equal to 2 – reflecting highly reliable estimates calculated through official statistical data; since this calculation is performed by the authors, not the government, the quality code attached to such values cannot be equal to one. All other estimates (e.g. slave records, very early censuses and anthropological reports) may obtain a value equal to 3 on the ClioInfra quality scale.

1. **Data sources for African countries, 1880 to 1970**

Official publications: a large share of our numeracy estimates is based on census data and official publications (e.g. UN Demographic Yearbooks) for different countries, as described in the www.clio-infra.eu database. We further integrated the information available in ClioInfra with new census figures that have recently become available: we added information from the following censuses: Ethiopia (1984), Guinea (1983), Malawi (1987), Rwanda (2012), Senegal (2002), Sierra Leone (2004), Sudan (2008), and early censuses for Uganda (1959; 1991) and Tanzania (1988). These sources allow us to estimate numeracy for most African countries between 1880 and 1970. Estimates based on this kind of source constitute roughly 70 percent of our whole panel-data sample.

2. **Data sources for African countries, 19th century**

Census data for Senegal: national archives of Senegal, Dakar: 3G2120 to 3G2140 and in folders 22G29 and 22G34, and include enumerations collected between 1832 and 1921; Complementary sources: 3G262 and 3G265 folders (1851, 1887 and 1913 3G2120 to 3G2140 and in folders 22G29 and 22G34. See G. Cappelli and J. Baten (2017). "European Trade, Colonialism and Human Capital Accumulation in Senegal, Gambia and Western Mali, 1770 – 1900." *The Journal of Economic History* 77 (3): 920-951.

Census Data for Algeria: Archives Nationales (France, Outre Mer: Aix-en-Provence): 107 MIOM 70, 71, 72, 73, 74, and 79.

Census Data Egypt: National Archives of Egypt, Census 1848 and 1868.

Census Data for Mauritius: Archives Nationales (France, Outre Mer: Aix-en-Provence): G1 473, G1 161.

Census data for South Africa: see Juif, Dacil, Joerg Baten and Ewout Frankema. 2018. "Ambivalent Colonial Encounters: The Rise of South African Indigenous Numeracy, 1800-1920". Manuscript Univ. C3M and Univ. Tuebingen.

Court registers for South Africa (NB: for data consistency checks, not in the main dataset): TEPC (Transcription of Estate Papers at the Cape of Good Hope) Transcription Project, *Documents of convicts and exiles of the Court of Justice of the Cape of Good Hope: CJ3186–CJ3192A and CJ2562–CJ2568* (Cape Town, 2005–7).

3. Slave records, 18th century

Slave registers from the Midlo-Hall and the Maranhão Inventories Slave Database (MISD) datasets: the data includes tens of thousands of slaves deported from Africa to the Americas (US and Brazil) in the late-18th and early-19th centuries. See: http://slavebiographies.org/databases.php

4. Anthropological reports

Anthropological reports: reports by anthropologists have been used to fill missing age-heaping values, when the sources did not present bias-related issues (e.g. selectivity).

We obtained data for Cameroon (1870s), Kenya (1840s), Somalia (1880s, 1890s), Sudan (1860s, 1870s), Togo (1860s) and Tanzania (1860s): the detailed list of primary sources is available upon request.

Data Appendix 2

Organization by modern countries, and potential regional selectivity

We organize the numeracy by birth cohort. For example, if ages for a sample of people between 30 and 40 years old are recorded via a census in 1900, we can proxy the numeracy of the same sample in 1860 – 70. Given that we can project modern data backwards, our methodology automatically reflects modern-country boundaries for the post-1880 period. However, we can frame earlier numeracy data within modern-country borders, too. Indeed, since most of our numeracy figures are constructed starting from individual data, the earlier individual age information (anthropological surveys, slave data and early local censuses) is assigned to the territories of present-day countries using the name of the birthplace (country, city etc.) of each individual. In a few instances, the African ethnic group is reported, which can also be assigned to a country with a minimal error margin. If a modern-country numeracy observation could only be estimated from just a very numerate or very innumerate region of the country, we dropped that figure. Whenever possible, we compared the regional distribution of numeracy in the past with later regional cross-sections (Baten et al. 2017). This procedure might be prone to a certain amount of measurement error; yet, this strategy is the only possibility to make informative comparisons over time, hence it was used in a large amount of studies (see van Zanden et al. 2014).

References

Baten, Joerg, Mikolaj Szołtysek, and Monica Campestrini. 2017. "'Girl Power' in Eastern Europe? The Human Capital Development of Central-Eastern and Eastern Europe in the Seventeenth to Nineteenth Centuries and Its Determinants." *European Review of Economic History* 21 (1): 29–63.

Zanden, Jan Luiten van, Joerg Baten, Marco Mira D'Ercole, Auke Rijpma, and Marcel Timmer, eds. 2014. *How Was Life? Global Well-Being Since 1820*. Paris: OECD.

Data Appendix 3

Table 2 – Variable definitions and sources

Variable	Definition	Source
Initial numeracy & average annual increases in numeracy	Numeracy is an index of basic cognitive and quantitative skills based on age-heaping. We calculate the increase in numeracy (%) in a decade and divide it by ten (years) to obtain the average annual increase.	The primary sources used to calculate numeracy and its growth rate are discussed in the text, see also the Data Appendix 1.
Colony dummy	A dummy variable that is equal to one if a country was a colony in each period. The date of colonization for each present-day country is based on the date representing the moment when the last portion of territory was occupied.	Own elaboration from Henderson and Whatley (2014).
Annual growth population	The population growth rate calculated from Frankema and Jerven (2014)	Own elaborations from Frankema and Jerven (2014).
Urbanization rate	Ratio between urban population and the population figures as provided by Frankema and Jerven (2014)	Own elaborations from the Clio-infra dataset and Frankema and Jerven (2014).
Intensity of international conflicts	For each country, the index measures the number of years within a ten-year period that were characterized by international conflicts. The index ranges between 0 and 1.	Clio-infra dataset.
Intensity of domestic conflicts	For each country, the index measures the number of years within a ten-year period that were characterized by internal conflicts. The index ranges between 0 and 1.	Clio-infra dataset.
Cattle p.c.	Number of cattle per head.	Clio-infra dataset.
Cropland (sq. km) p.c.	Cropland (measured in squared kilometres) per head.	Clio-infra dataset.
Cash-crop value per capita	To calculate cash crop value, we included the following most-important exported goods based on the literature on African cash-crop trade: kola nuts, coffee, oleaginous fruits, groundnuts, pal kernels, gum (and resins), wax, sugar, cocoa, fat and oil from cocoa, rubber, hides, wool, and raw cotton. We converted all figures into thousands of 2018 GBP	The African Commodity Trade Database. The data are discussed by Frankema et al. (2018).

(*Cont'd*) Variable	Definition	Source
Railway (km) per squared km	The length on railway lines (km) per squared kilometre.	Mitchell (2007) cross-checked geo-coded data by Jedwab and Moradi (2016).
Tsetse suitability index	An index showing the suitability to the tsetse fly.	Alsan (2015).
Intensity of slave deportations	The logarithm of the number of slaves deported from each present-day country due to slave trades, normalized to surface.	Nunn (2008).
Strength of precolonial political organization	Precolonial political institutions are measured by using the "jurisdictional hierarchy beyond the local community level". We obtain the present-day country value from the original, ethnicity-level data. In the paper, we rely on a dummy equal to 1 if the original variable is greater than 2, which identifies state organizations beyond the local level.	Michalopoulos and Papaioannou (2013).
State capacity	First of all, Borcan et al. take into account whether there was some governmental organization beyond the tribal level. Secondly, they assessed whether an own government existed or whether it was a foreign power that provided the government, via governors for example. Thirdly, they estimated how much of the territory was under the control of one government (rather than several), using modern size as a reference.	Borcan et al. (2018)

References

Alsan, Marcella. 2015. "The Effect of the TseTse Fly on African Development." *American Economic Review* 105 (1): 382–410.

Borcan, Oana, Ola Olsson, and Louis Putterman. 2018. "State History and Economic Development: Evidence from Six Millennia." *Journal of Economic Growth* 23 (1): 1–40.

Frankema, Ewout, and Morten Jerven. 2014. "Writing History Backwards or Sideways: Towards a Consensus on African Population, 1850–2010." *The Economic History Review* 67 (4): 907–31.

Frankema, Ewout, Jeffrey Williamson, and Pieter Woltjer. 2018. "An Economic Rationale for the West African Scramble? The Commercial Transition and the Commodity Price Boom of 1835–1885." *The Journal of Economic History* 78 (1): 231–67.

Henderson, Morgan, and Warren Whatley. 2014. “Pacification and Gender in Colonial Africa: Evidence from the Ethnographic Atlas.” 61203. MPRA Paper. University Library of Munich, Germany.
Jedwab, Remi, and Alexander Moradi. 2016. “The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa.” *The Review of Economics and Statistics* 98 (2): 268–84.
Michalopoulos, Stelios, and Elias Papaioannou. 2013. “Pre-Colonial Ethnic Institutions and Contemporary African Development.” *Econometrica* 81 (1): 113–52..
Mitchell, B. R. (Brian R.). 2007. *International Historical Statistics: Europe, 1750-2005*. Sixth edition. Palgrave Macmillan.
Nunn, Nathan. 2008. “The Long-Term Effects of Africa’s Slave Trades.” *The Quarterly Journal of Economics* 123 (1): 139–76.

Appendix A – Organizing the numeracy series by birth cohort

In the literature that uses age-heaping-based numeracy, data are typically analyzed through time series by organizing figures by birth cohorts, because very basic numeracy is typically acquired during the first decade of life. This has been systematically studied by using the same birth cohort in country, but comparing different census years and, correspondingly, different age groups. For example, in a hypothetical country, people born in the 1890s would be part of the age group 33 to 42 in a census taken in 1928, while being in the age group 43 to 52 in a different census taken ten years later (1938). If the birth cohort did not matter and forgetting one's own age at higher ages did, one would observe systematic deviations of numeracy figures concerning the exact same sample across census taken at different times. By contrast, in many empirical studies, the birth-cohort effect has been shown to be the strongest. For example, Manzel et al. (2012, Figure 2) show that people born in Argentina in the same decades throughout the 18th and 19th centuries had almost exactly the same numeracy independently of the censuses through which they were recorded. Crayen and Baten (2010) discuss similar issues for a large number of countries.

References

Crayen, Dorothee, and Joerg Baten. 2010. "Global Trends in Numeracy 1820-1949 and Its Implications for Long-Term Growth." *Explorations in Economic History* 47 (1): 82–99.

Manzel, Kerstin, Joerg Baten and Yvonne Stolz. 2012. "Convergence and divergence of numeracy: the development of age heaping in Latin America from the seventeenth to the twentieth century". *The Economic History Review* 65 (3): 932-960.

Appendix B – Alternative regions for border analysis

Figure B1 – South Omo (Ethiopia) vs Marsabit & Moyale (Kenya): Sources: Minnesota Population Center (2018).

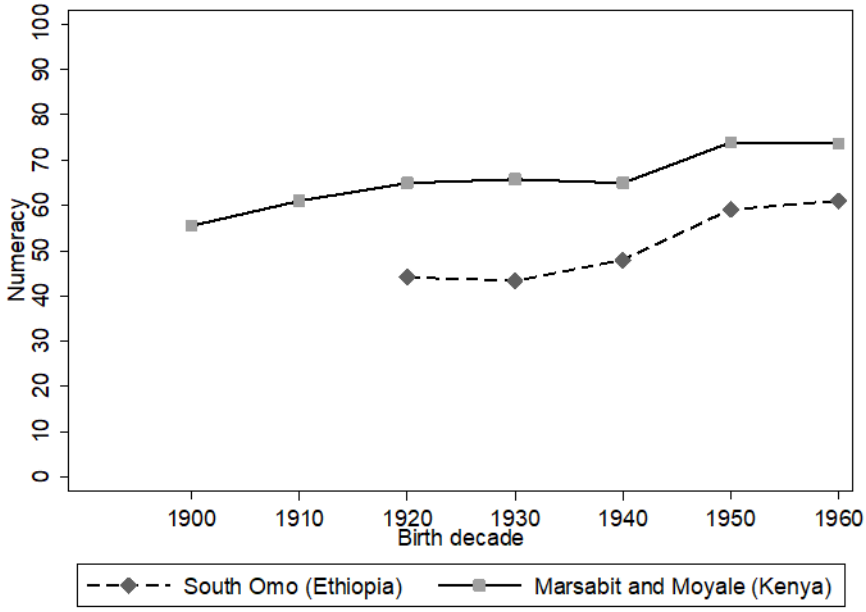
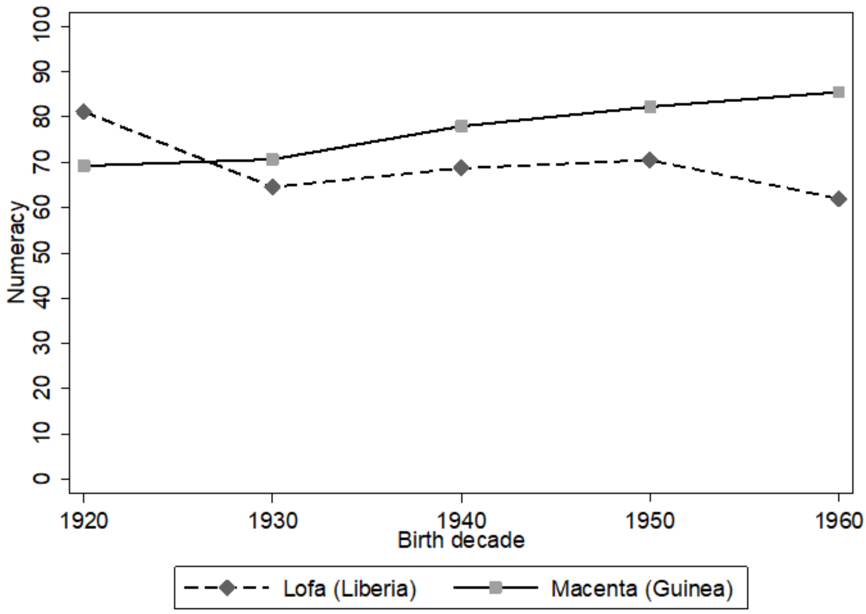


Figure B2 – Lofa (Liberia) vs Macenta (Guinea). Sources: Minnesota Population Center (2018).



References

Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.1 [dataset]. Minneapolis, MN: IPUMS, 2018.

Appendix C – The colonial status dummy based on Henderson and Whatley (2014)

Henderson and Whatley provide estimates of the date of colonization at the subnational (ethnic) level (Henderson and Whatley 2014). We rely on their dataset on the duration of colonial rule to code a country-level date of colonization: a country is defined as being colonized when the last territory within its present-day borders is classified as occupied. Like Henderson and Whatley (Henderson and Whatley 2014, pp. 9-11), we are interested in the beginning of systematic colonization, not "outside contact" or "commercial contact". In addition to this being important to allow missions to operate in the territory, the ending of colonial-expansion conflicts was linked to the demand for education by Africans: once it became clear that Europeans had come to stay in the medium or long term, the benefit of European education grew compared to its initial costs (Frankema 2012, p. 349).

This colonialism-status dummy variable is a considerable improvement for our purposes compared to studies that focused on trading posts along the coast (Olsson 2009). Senegal is one example: our colonization dummy variable relying on Henderson and Whatley's work defines Senegal as being "systematically" colonized starting in the 1870s, when the majority of the inner (populated) regions of Senegambia were already occupied, whereas trading posts had existed since the 16th century. Our dataset codes Nigeria as being colonized by the British in 1910, which is consistent with the discussion by Henderson and Whatley (2014, pp. 10-11). The same holds true for Ghana: following a war with the Ashanti Kingdom, the British protectorate was extended to the northern areas of today's Ghana at the turn of the 20th century, when the construction of the railway also began (Jedwab et al. 2019).

References

Frankema, Ewout. 2012. "The Origins of Formal Education in Sub-Saharan Africa: Was British Rule More Benign?" *European Review of Economic History* 16 (4): 335–55.

Henderson, Morgan, and Warren Whatley. 2014. "Pacification and Gender in Colonial Africa: Evidence from the Ethnographic Atlas." MPRA Paper. University Library of Munich, Germany.

Jedwab, Remi, Felix Meier zu Selhausen, and Alexander Moradi. 2019. “The Economics of Missionary Expansion: Evidence from Africa and Implications for Development.” Working Paper.

Olsson, Ola. 2009. “On the Democratic Legacy of Colonialism.” *Journal of Comparative Economics* 37 (4): 534–51.

Appendix D – Discussion of the control variables added to the panel regressions

The average annual rate of population growth may correlate negatively with increases in numeracy, due to a quantity-quality trade-off; yet, on the other hand, population growth in preindustrial societies may indicate a general growth in wellbeing, which would affect schooling positively (Persson and Sharp 2015). The urbanization rate is included to control for the rising demand for skills associated with the presence of urban centers, together with the fact that commuting distance to colonial or missionary schools is expected to be reduced in urban regions compared to rural regions (Jedwab et al. 2019): hence, we expect a positive relationship between urbanization and the level of human capital accumulation in our within-country estimation. We admit that urbanization may be endogenous to human capital accumulation: we are confronted with the trade-off between variables that might be endogenous controls and omitted variable bias. Hence, we include such controls in most of our regression models to avoid omitted variable bias and exclude them in other specifications to avoid this bad-control problem: when doing so, the results remain virtually unchanged. Both indices measuring the intensity of domestic and international conflicts should correlate negatively with increases in numeracy. Cattle and cropland per capita are included as proxies for the availability of proteins and calories because of the potential relationship between nutrition and cognitive skills. As further controls coded from previous studies, we include the per-capita value of the cash crop trade (logarithm), the density of railways and state capabilities.

References

Jedwab, Remi, Felix Meier zu Selhausen, and Alexander Moradi. 2019. "The Economics of Missionary Expansion: Evidence from Africa and Implications for Development." Working Paper.

Persson, Karl Gunnar, and Paul Sharp. 2015. *An Economic History of Europe*. Cambridge University Press.

Appendix E – Panel evidence when relying only on census data

Table E1 – FE panel-data model: observations based only on census data.

Dep. Var.: Av. annual increases in numeracy	(1) 1730 – 1970	(2) Post 1800	(3) Post 1800 w/controls	(4) Post 1800 w/controls + interactions	(5) Post 1900 w/controls + interactions
Initial Numeracy	0.0436	0.0436	0.0421	0.0573	0.1292***
	(0.046)	(0.046)	(0.056)	(0.068)	(0.027)
Colony	1.2143***	1.2143***	0.7478	-0.9632*	-1.4599***
	(0.267)	(0.267)	(0.524)	(0.508)	(0.343)
Colony x UK				1.9471**	2.6852***
				(0.949)	(0.680)
Colony x France				-0.0121	-1.2619*
				(0.741)	(0.744)
Urbanization rate			-2.8968	-2.7279	3.0245
			(2.730)	(2.449)	(2.085)
Intensity of int. conflicts			-2.2485**	-1.8290	1.2739
			(1.035)	(1.400)	(0.841)
Intensity of dom. conflicts			-1.5929	-2.0048	-0.3775
			(1.804)	(1.602)	(1.177)
Cattle p.c.			0.3573	0.1748	-0.9135*
			(0.585)	(0.571)	(0.528)
Cropland (sq. km) p.c.			0.6217	0.9154	0.1086
			(2.001)	(2.260)	(0.961)
Annual growth population			0.5755**	0.5175**	0.1909
			(0.277)	(0.190)	(0.145)
Cash-crop-trade value per capita (log)			-0.8764***	-0.8845***	-0.0153
			(0.271)	(0.253)	(0.150)
Km railways per sq. km (Mitchell)			-6.5340*	-4.1391	2.5968
			(3.430)	(4.332)	(1.731)
Observations	194	194	188	188	160
Country and Time FE	Y	Y	Y	Y	Y
N. of countries	38	38	36	36	36
Adjusted R-squared	0.474	0.474	0.557	0.559	0.543

Robust standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix F – Did the boundary problem of human capital indicators influence the results?

The analysis of human capital indicators over time is always challenged by the upper-bounded nature of school-related indices, particularly in the past – e.g. both literacy and numeracy rates have a maximum value of 100 percent – even though human-capital disparities across countries might still be quite relevant. This upper boundary is also related to the problem of declining growth rates as the value of e.g. literacy or numeracy approach 100.

However, in the context of 19th-century African countries, one may argue that other indices or indicators of human capital and education are even more problematic: if the number of school years for African countries were available throughout the 19th century, most values would probably be close to zero and thus hide potentially-relevant disparities in human-capital endowments.

A first way to explore if the boundary issue really is problematic for the validity of our results is to explore whether African countries were about to reach 100-pecent numeracy in the second half of the period that we study – e.g. around 1900, 1930, or 1960. We find that not even a single country was within the 95-100 range until well into the 1900s. Rwanda and Senegal were the only African countries to have reached virtually universal numeracy in the 1930s; Botswana joined the club in the following decade. Finally, even in the 1960s (the last decade in our study), only 6 countries out of the 43 countries (with non-missing numeracy series) were characterized by 95 to 100 percent numeracy. In summary, the upper-bounded nature of numeracy does not weaken our identification strategy, because most countries were far from the upper bound even in the last decade considered, which is the 1960s.

We also experimented with SRM (Shortfall-Reduction Method) which explicitly considers, and adjusts for, boundary problems. Contrary to annual growth, the SRM reflects the effort to close the gap with the highest possible value by assuming that a given percentage of reduction is equally viable at different initial levels of development (Equation 1 below). For example, numeracy changing from 90 to 95 means that there is a 50 percent reduction of the shortfall (distance from max value), which – in the SRM framework – is the same of numeracy changing from 50 to 75 percent. Using the standard growth-rate method, the above cases would be 5.5 percent and 50 percent, respectively (Gidwitz et al. 2010; Felice and Vasta 2015).

(1) $SRM = actual\ (x) value\ - past\ (x) value\ /\ maximum\ (x) value - past\ (x) value$

We replicate Table 4 in the paper here, but we use the Shortfall-Reduction (%) over decades as an alternative dependent variable instead of the average annual increase in numeracy. The results are consistent with those reported in the paper – if anything, the coefficients of interests are, in the fullest specifications, more statistically significant, with a similar magnitude relative to the SR mean, which is equal to 10 percent. The disadvantage of SMR model results is that they are more difficult to read, hence we report them below (Table F1).

We present results based on a fixed-effect panel regression with heterogeneous colonial rule where the numeracy data that are not based on censuses are excluded. By doing so, we explore the possibility that historical sources such as anthropological reports and slave records might be more prone than census data to selection bias. As the table below shows, the main results do not change.

References

Felice, Emanuele, and Michelangelo Vasta. 2015. "Passive Modernization? The New Human Development Index and Its Components in Italy's Regions." *European Review of Economic History* 19 (1): 44–66.

Gidwitz, Zachary, Martin Philipp Heger, José Pineda, and Francisco Rodríguez. 2010. "Understanding Performance in Human Development: A Cross-National Study." HDRP-2010-42. Human Development Research Papers (2009 to Present). Human Development Report Office (HDRO), United Nations Development Programme (UNDP).

Table F1 – Colonial rule and the Reduction in the Shortfall % (SRM). Panel data, 1730 – 1970.

Dep. Var.: Av. annual increase in numeracy	(1) 1730 – 1970	(2) Post 1800	(3) Post 1800 w/controls	(4) Post 1800 w/controls + interactions	(5) Post 1900 w/controls + interactions
Initial Numeracy	0.7827***	1.0340***	1.0796***	1.2722***	2.5759***
	(0.158)	(0.250)	(0.322)	(0.397)	(0.357)
Colony	4.9843	5.7186	15.5200**	-7.6893	-14.8733**
	(5.623)	(6.053)	(7.458)	(10.800)	(7.307)
Colony x UK				26.7660**	31.9798***
				(10.574)	(11.153)
Colony x France				8.6251	2.6730
				(12.861)	(22.985)
Urbanization rate			239.5577***	235.5412***	274.1006***
			(45.023)	(46.485)	(42.125)
Intensity of int. conflicts			2.2769	5.9022	-13.9122
			(6.350)	(5.899)	(15.270)
Intensity of dom. conflicts			13.5329	10.8147	3.3883
			(14.985)	(15.546)	(38.908)
Cattle p.c.			-11.0644	-13.0420*	-14.9088
			(7.932)	(7.458)	(13.005)
Cropland (sq. km) p.c.			-24.2778*	-20.3132	-55.3041***
			(13.785)	(16.086)	(18.667)
Annual growth population			7.5092**	6.8818**	5.3713
			(3.662)	(3.314)	(4.139)
Cash-crop-trade value per capita (log)			-3.0395	-2.9370	-0.4781
			(1.926)	(1.768)	(3.851)
Km railways per sq. km (Mitchell)			-39.6811	-10.6808	-31.7344
			(26.184)	(31.142)	(42.153)
Observations	261	214	207	207	165
Country & Time FE	Y	Y	Y	Y	Y
N. of countries	42	40	38	38	37
Adjusted R-squared	0.153	0.144	0.340	0.350	0.480

Robust standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix H – Historical evidence on the British education system in Africa, vs others

Cogneau (2003) has argued that "the French free and secular education system […] seems to have been less of a success than the British system based on a partnership between the missionaries and government. The Belgian and Portuguese systems relying completely on missionary work were no more of a success than the French system […]" (Cogneau 2003, p. 28). According to the historical literature, the British colonial education system in Africa was different from the others: as argued by White (1996), not only did British colonial administrators provide freedom to missions and the ground for denominational competition, they chose to cooperate with them. Furthermore, schooling was seen as something that would need to adapt to local conditions. Content was also important. Although British colonial education was dual and segregated—divided between more vocational education for Africans and more academic formation for Europeans and elites—according to some (Madeira 2018), it allowed native students to further their studies. Most importantly, the British did not place any religious bar to admission, despite the supply of schooling being linked to missions.

Importantly, the British subsidized African schooling more heavily than other colonial administrations in the period of 1900–1940 did, often based on taxes paid by Africans (see data for the Gold Coast, pre-WWI German Togoland and post-WWI French Togo in Cogneau and Moradi 2014, Table 1, p. 701). Furthermore, as Meier zu Selhausen (2019) argues, European missionaries soon started to train local, native African missionaries, given the lower mortality rates for natives compared to Europeans as well as their lower cost—which was also mainly paid through local contributions (Frankema 2012). This choice was cost effective, provided that native missionaries could teach in local languages. The ratio of African to Western missionaries around 1911 was larger in Protestant Africa compared to Catholic places (Meier zu Selhausen 2019, Table 2 and Figure 7).[10]

The different stance towards missions by the British administrators as compared to other metropoles in Africa is useful to understand the evolution of colonial school policy in other African colonies. Colonial education policy in Portuguese Mozambique borrowed from the French "assimilation"— although Portuguese policy was a more "adaptive" approach to schooling that was more similar to the British policy of indirect rule. According to Madeira (2018), Mozambique was 30 years behind in terms of public primary school enrollment compared to French and British colonies in the same region. Interestingly, consistent with Madeira's claim, our data show that, while average numeracy in East Africa was 75 percent in 1920, the value in Mozambique was 48 percent—close to the mean value of East Africa in 1880 (54 percent). Following Madeira's work, one may put forth that the disadvantage of Portuguese colonial schooling was due to the weakness of missionary activity towards the end of the 19th century, which could not compete with foreign—particularly Protestant—missions in the colony. This feature made support from the colonial administration crucial; again, however, an important difference existed between the British and Portuguese administrations. Instead of being characterized by cooperation, the relationship between the missions and the local administrators in Portuguese Africa was one of contrast, at least in the period following the Scramble. On the eve of the 20th century, local colonial officials blamed the insufficient number and quality of churches, priests and missionaries on the ecclesiastical administration; at the same time, the Catholic Church equally blamed the local government for disregarding the commitment that had stipulated the appointment of "a missionary with a school for each circumscription". Madeira (Madeira 2005, p.51) cites Sebastião José Alves, head of the

[10] In Meier zu Selhausen's data, the different ratios between Protestant and Catholic missionaries seem to be somewhat driven by South Africa. However, our results remain virtually unchanged if we exclude South Africa from the cross-section regression with interaction terms.

Mozambique Ecclesiastic Government, who complained about the lack of state support: "Our own missions in Zambezi have not been the object of any visits for the past six years […] Our colonies are escaping us because we do not have national elements and we do not accept foreign [Catholic] missionaries to work under the Portuguese ecclesiastic authority". Indeed, Catholic missions in Portuguese Africa were not able to compete with Protestant foreign missions "with regard to missionary training (more theoretical than practical); in the organization of the missions (dispersed and sometimes in inhospitable places); in the constitution of its personnel (secular, isolated, left to fend for themselves); in the facilities and material resources (threadbare, derisory, precarious premises); and so on" (Madeira 2005, p. 53). The Protestant missions, however, were not allowed to take root. Finally, the implementation of a dual system of schooling in the French model, with schools for Europeans and assimilated Africans and schools for indigenous people, may have further depressed numeracy: the Portuguese system was more segregated than the British system, according to Madeira (2005).

VanderPloeg (1977) has argued that Cameroon—a German colony later taken over by the British and French—was regarded as a model of good colonial schooling by the British in particular. Two government-led schools were opened around 1890, being met by a high demand for schooling by the natives due to the spread of cash crops. By 1911, there were four government schools, but the number of (primary) mission schools was an impressive 495, according to Shibata (2005). Thus, it does not come as a surprise that numeracy in Cameroon grew remarkably with the colonial education system, from a flat trend of approximately 40 percent in the 19th century to approximately 70 percent in 1920 (Figure H1 below).

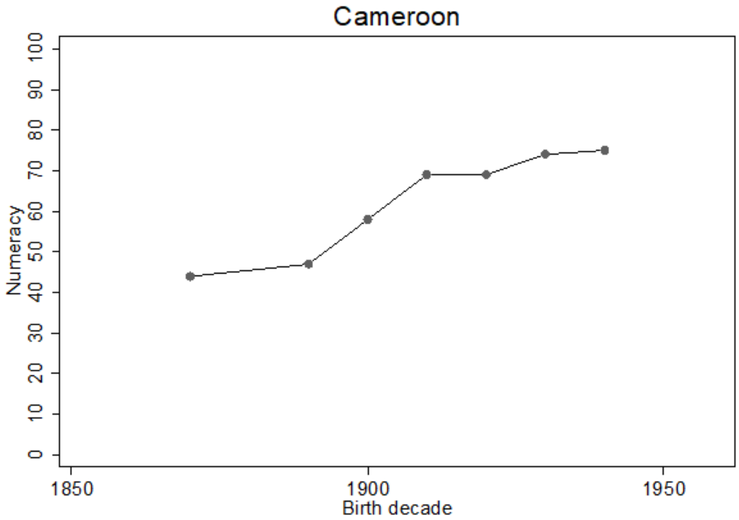


Figure H1: Numeracy in Cameroon. Sources: see main text.

VanderPloeg (1977, p. 96) argues that the Germans themselves were surprised by the progress made by African children. A young missionary teacher, Theodorus Christaller, was sent to Kamerun to establish Germany's governmental schooling in the colony and "was forced to admit to the remarkable mathematical ability demonstrated by his pupils. […] The teachers accepted that the Kamerun Africans' academic performance was comparable to that of German youth of similar age". Our evidence suggests that this was an overstatement, but the positive trend was impressive. VanderPloeg does not thoroughly assess what factors led to a rapid spread of education in the colony. According to him, competition for political and economic supremacy in Douala by different denominations and Christian factions may have strengthened the race towards schooling: "One can view the schools as pawns in the struggle for political and economic supremacy between the two factions and in the Douala assertion of their rights relative to the Germans" (VanderPloeg 1977, p. 105). Therefore, "the Christians of the Kamerun coast were quite willing to set up their own churches in opposition to those of the European missionary societies" (idem), implicitly stressing a "competition" mechanism that has been deemed to have fostered African education in the past (Gallego and Woodberry 2010).

Furthermore, the effort to contain the German "economic expansion" in the coastal areas of Cameroon "depended on educated Africans" (idem).

Dunkerley (2009) argues that the spread of education in the Belgian Congo starting in the 1920s was comparatively favourable, following the atrocities perpetrated by Leopold II : a back-of-the-envelope calculation provided by the author and based on colonial figures suggests that 37% of school-age children were enrolled in the colonial school system around 1940 (Frankema 2013). Besides a small minority in secondary schools, most of the pupils enrolled were in the first tier of the colonial education system, assimilable to primary education. Furthermore, according to the author, most pupils were enrolled in subsidized Catholic missionary schools, or non-subsidized Catholic and Protestant rural missionary schools. Given the prominence of Catholic missions, the gender gap was very large (according to the author, Protestant schools could be more easily attended by girls). Despite this, such an enrolment rate is high compared to other African colonies in the same period (Frankema 2012). For example, estimates based on conscripts in the Gold Coast (Ghana) suggest a value of c. 20 percent throughout most of the Interwar years. According to Frankema (2013), this was due to the success of missionary schooling in the territory: the number of foreign missionaries was impressive in a comparative perspective, and its growth parallels the one experienced by schooling (Frankema 2013, p. 164). This, in turn, was prompted by the absence of institutional entry barriers, which fostered denominational competition (Gallego and Woodberry 2010) and by the Africanization of the mission in the early 20th century (Meier Zu Selhausen 2019). One thing that the work by Dunkerley (2009) and Frankema (2013) makes clear is that school enrolments and the functioning of the school system in Belgian Congo cannot be explained only by the way that missions operated (e.g. Protestant vs. Catholic) or just by the metropoles' top-down policy choices (e.g. assimilation and adaptation), but rather, in the way that the two interacted and developed on the ground.

In sum, this qualitative evidence lends support to the idea that the interaction between specific education systems (focusing on African teachers and liberal competition) and missionary activity was a crucial determinant of the pace of human capital formation in Africa. If the British system prompted more rapid numeracy increase, for example, historical evidence suggests that Portuguese colonies lagged behind. The system in German colonies is under-researched and probably heterogeneous, but our data are consistent with the view expressed above of rapid increases in numeracy in the 20th century.

References

———. 2013. “Colonial Education and Postcolonial Governance in the Congo and Indonesia.” *Colonial Exploitation and Economic Development: The Belgian Congo and the Netherlands Indies Compared*, 153–77.

———. 2018. “Comparing Colonial Education Discourses in the French and Portuguese African Empires: An Essay on Hybridization.” *Revista Española de Educación Comparada* 0 (31): 130–46.

Cogneau, Denis, and Alexander Moradi. 2014. “Borders That Divide: Education and Religion in Ghana and Togo Since Colonial Times.” *The Journal of Economic History* 74 (3): 694–729.

Cogneau, Denis. 2003. “Colonisation, School and Development in Africa. An Empirical Analysis.”

Dunkerley, Marie Elizabeth. 2009. “Education Policy and the Development of the Colonial State in the Belgian Congo, 1916-1939.,” September.

Frankema, Ewout. 2012. “The Origins of Formal Education in Sub-Saharan Africa: Was British Rule More Benign?” *European Review of Economic History* 16 (4): 335–55. https://doi.org/10.1093/ereh/hes009.

Frankema, Ewout. 2012. “The Origins of Formal Education in Sub-Saharan Africa: Was British Rule More Benign?” *European Review of Economic History* 16 (4): 335–55.

Gallego, Francisco A., and Robert Woodberry. 2010. “Christian Missionaries and Education in Former African Colonies: How Competition Mattered.” *Journal of African Economies* 19 (3): 294–329.

Madeira, Ana Isabel. 2005. “Portuguese, French and British Discourses on Colonial Education: Church–State Relations, School Expansion and Missionary Competition in Africa, 1890–1930.” *Paedagogica Historica* 41 (1–2): 31–60.

Meier Zu Selhausen, Felix. 2019. “Missions, Education and Conversion in Colonial Africa.” In *Globalization and The Rise of Mass Education*, edited by David Mitch and Gabriele Cappelli. London: Palgrave Macmillan.

Shibata, Atsuko. 2005. “The ‘Germanization’ of Africa: White Education InGerman South-West Africa.” Hitotsubashi University.

vanderPloeg, Arie J. 1977. “Education in Colonial Africa: The German Experience.” *Comparative Education Review* 21 (1): 91–109.

White, Bob W. 1996. “Talk about School: Education and the Colonial Project in French and British Africa, (1860-1960).” *Comparative Education* 32 (1): 9–25.