

University of Tübingen
Working Papers in
Economics and Finance

No. 20

Growth Effects of 19th Century Mass Migrations:
“Fome Zero” for Brazil

by

Yvonne Stolz, Jörg Baten & Tarcísio Botelho

Faculty of Economics and Social Sciences
www.wiwi.uni-tuebingen.de



GROWTH EFFECTS OF 19TH CENTURY MASS MIGRATIONS: “FOME ZERO” FOR BRAZIL?

Yvonne Stolz* Joerg Baten** Tarcísio Botelho***

We estimate a long-run trend of Brazilian human capital that extends back to the very beginning of the 18th century. With new data on selective immigration during the era of mass migrations at the end of the 19th century, we show that human capital endowment of international migrants can induce effects on economic development that persist until today. According to our estimations, the effect of selective immigration on real GDP per capita in the year 2000 is significant and equals around 75 US \$ overall. As a reference, this value equals the amount poor Brazilians get to supplement their subsistence in the “*Fome Zero*” (Zero Hunger) program. We argue that human capital formation is a highly path-dependent and persistent process.

JEL codes: F22, J40, I21, N30

Keywords: migration, economic growth, Brazil, human capital, path dependency

Acknowledgements: Comments by Luis Bértola, Stephen Broadberry, Ewout Frankema, Tommy Murphy, Jan Luiten van Zanden, Jeffrey Williamson, Financial Support of the German Science Foundation (DFG), ESF GlobalEuroNet program and the EU HIPOD program is gratefully acknowledged, and data support of IPUMS, NAPP and CEDEPLAR/UFGM.

*University of Tuebingen, yvonne.stolz@uni-tuebingen.de

**University of Tuebingen and CESifo, joerg.baten@uni-tuebingen.de

***Universidade Federal de Minas Gerais, tbotelho@fafich.ufmg.br

Introduction

19th century mass migrations drove some 55 million Europeans to the New World in search of a better life (Hatton and Williamson 1998). The magnitude of migratory flow has been unprecedented in history. This has been possible due to a conjuncture of events such as the transport cost revolution, a sharp rise in population growth in the old world and labor scarcity in the New World, as documented well in the literature (Hatton and Williamson 1998, Cohn 2009, Balderas and Greenwood 2010). Only recently, however, the skill selectivity of these migrants has shifted to the centre of attention to scholars (Cohn 2009, Murin and Viarengo 2010, Stolz and Baten 2011). Evidence that allows measuring the impact of selective migration on destination country's economies is therefore a clear research desideratum. The human capital which migrants incorporate, their individual skill-set, is an important channel through which migration can have persistent effects on destination country's economic prosperity. Human capital includes formal education, but also entrepreneurial spirit and motivation. These capabilities and skills will not be applied in the home country's economy, but in the economy of the receiving nation. Selective migration can have persistent effects on the process of human capital accumulation of receiving nations and thus for economic growth in general. The direction of this effect depends very much on how immigrants compare to the destination country's population. In this context, research providing long-run data on human capital endowment of both migrants and native population is particularly important, as it sheds light on the long-term relationships, persistence and path-dependency of these processes.

This article provides new evidence on Brazilian human capital formation in the very long-run, from the early 18th to the 20th centuries. Furthermore, we collected new data on human capital endowment of 19th century migrants to Brazil, whose educational level was higher than that of native Brazilians. We use the age-heaping method (Mokyr 1983, A'Hearn,

Baten and Crayen 2010), which provides a numeracy indicator particularly suitable for early stages of human capital formation. It exploits the tendency of less educated people to round their age instead of stating an exact age. Numeracy is highly correlated with other human capital output indicators such as literacy. Its main determinant is schooling investments (Crayen and Baten 2009). We argue that - apart from the migratory transfer of skills - the relatively high human capital levels of immigrants induced spillover effects to overall Brazilian human capital endowment which persistently enhanced economic growth in Brazil. Endogenous growth theory states a theoretical relationship between human capital development and economic growth (Lucas 1988, Romer 1994, amongst many others). Testing this relationship empirically is complex. Capolupo (2009) and others (for example, Cohen and Soto 2007) show that the impact of human capital might have even been underestimated in previous research due to a lack of consistent and long-run data on human capital. In this study we are interested in the long term impact of human capital formation on economic performance. Does 19th century human capital development still impact growth performance today? In a cross-section of 90 countries, we show that numeracy levels in 1900 have a significant impact on real GDP per capita in the year 2000, even after controlling for a set of control variables and after instrumenting numeracy to cope with potential endogeneity and measurement error. That means countries which already performed well in terms of human capital around the year 1900, still have a high real GDP per capita in the year 2000, whereas countries which performed badly in terms of numeracy around 1900, today are situated more at the bottom of the global income distribution. This exercise provides some evidence, that human capital endowment a hundred years ago still impacts growth performance today.

In a counterfactual, we show that the effect of migratory human capital transfer has also an economic impact. For illustration, consider that in the year 2002, 44 million Brazilians received between 50 and 95 Brazilian *Reais* per month to amend their subsistence with the “*Fome Zero*” program initiated by President *Lula da Silva*.¹ For illustration: the effect is

larger than one of these payments, namely around 75 US \$ or somewhat over 130 *Reais*. This would be the estimated magnitude of the persistent human capital effect on today's income that selective mass migrations around the turn of the last century still has.

When it comes to 19th century mass migrations, mainly the destinations of North America and Argentina have been at the centre of scholarly attention (Hatton and Williamson 1998). The Brazilian experience has somewhat been neglected, although it received over 4.1 million immigrants between its independence in 1822 and 1926, ranking after the US and Argentina among the top three most favorite immigration countries during the era of mass migrations (Ferenczi and Willcox, 1929, important exceptions are, for example, Klein 1995, Balderas and Greenwood 2010). Brazil was the last country in the Western World to abolish slavery. Consequently the demand for immigrant labor on the plantations of the thriving coffee sector created an immigrant flow that was attracted to a large extent by favorable immigration policies of the Brazilian government which paid passages or granted access to land. Also owners of large coffee plantations were willing to pay the travel costs for immigrants and their families, who then worked off their debts on the plantations. For Brazil it was necessary to be financially attractive, because the country had to compete with Argentina and the US, whose economies provided higher income per capita and hence higher wage gaps in relation to European source countries. This created stronger pull forces to these economies than to Brazil (Holloway 1980, Stolcke 1988).

The article is structured as follows: in the next chapter, we clarify the relation between human capital formation, selective migration and economic growth. We also propose possible channels via which human capital spillovers could have taken place. In the next section we review the literature on human capital formation in Brazil since the 18th century. Then, we introduce and discuss the data. The following section displays our regression results and the counterfactual. In a separate section, we look at regional human capital distribution and

immigration, and verify whether high absolute immigration in a state generally correlate with numeracy increases. We end with a conclusion.

Economic Growth, Human Capital Formation, and International Migration

Human capital plays a central role in endogenous growth theories. Generally, there exist three broad strands of literature. Lucas (1988) and others have provided models that incorporate human capital as a separate factor that becomes the driving force for economic growth. An additive recent strand of literature puts the process of knowledge creation via innovation and R&D into the centre of attention (Grossman & Helpman, 1991; for a comprehensive survey of the topic see Drinkwater et al 2003). All of these approaches have been applied to the field of international migration to answer the question of how migration affects economic growth in the long run. Central to these approaches is the assumption that migrants become a part of the destination country's human capital stock, which then affects economic performance.

Empirical studies using one of these models commonly conclude that international migration can affect economic growth through the skill composition of migrants which they apply in the destinations' labor market (Reichlin and Rustichini 1993, Lundborg and Segerstrom 2000, 2002 as in Drinkwater et al 2003).² In destination countries, the question is whether migrants are high or low skilled and how they compare to natives.³ Although migration is typically not included in growth regressions, there exists evidence from the field of economic history that international migration induces growth effects in receiving nations for the case of Argentina (Taylor 1997). Taylor uses a Computable General Equilibrium Model to capture the effects of migration on the Argentinean economy during the age of mass migration. He finds that immigration causes a significant increase in GDP.⁴ However, the outcome of the analysis is very sensitive to initial assumptions of a variety of model parameters that have to be chosen or estimated beforehand.

Recently, the Brazilian case of the state of São Paulo has been investigated by Carvalho Filho and Colistete (2010). Based on historical data from agricultural censuses and education statistics, these authors find a positive and enduring effect of the presence of foreign born immigrants on the supply of public instruction, as well as negative effects of land concentration.

In this study, we quantify the long-term impact of selective migration on human capital formation and economic growth in Brazil. Human capital formation is a highly persistent and path dependent process. Looking at numeracy levels in the year 1900 versus higher educational attainment of the total population provided by Barro and Lee (2001) (Figure 1), we find a strong positive relationship. Both are indicators of human capital, however, while numeracy captures a basic form of human capital, the latter is an indicator suitable for economies today, capturing a very advanced form of human capital, namely average years of higher schooling in the total population. It is evident that countries performing poorly in terms of human capital in the year 1900, like Egypt or Bangladesh, are still underperforming in the year 2000. On the other hand, countries like the US or Japan, which already had solved their numeracy problem around 1900, are among the top-performers in the year 2000 (Figure 2). The graph shows the path dependency of human capital formation in the very long run. Due to this path dependency, and the theoretical relation of economic growth and human capital, we would also expect the long-run relationship between early human capital formation and economic output to hold.

We first estimate the effect of human capital endowment in the year 1900 measured with age heaping techniques on log GDP per capita in the year 2000, controlling for a set of geographical and institutional variables.⁵ In a next step, we calculate a counterfactual, assuming that the migrants' human capital had not arrived in Brazil, to assess whether there are any economically important long-term effects. In a next section, we look at regional dispersion of human capital and correlate this with immigration statistics, to find out, whether

human capital enhancing effects took place mainly in the Brazilian states which were also predominantly destinations for the immigrants. This would give further support to the hypothesis that the human capital of migrants produced spillover effects to the overall Brazilian human capital formation.

Through which channels does the human capital of migrants affect economic growth? Migrants take their human capital endowment with them, which comprises their crude manpower, but also entrepreneurial spirit, innovative ideas and motivation, as well as formal and informal education and training. Migrants are not a random sample of the source country population (Todaro 1969, Hatton and Williamson 1998, 2008). Positive or negative selection in terms of formal education, relative to the source country's populations highly depends on skill premia or relative inequality of source and destination countries (Borjas 1994, Stolz and Baten 2011). It has been argued in the literature that the great majority of migrants however, are positively selected in terms of their willingness to take risks, entrepreneurial spirit and motivation. This leads to four channels, via which the migrants' human capital induces positive spillover effects to destination countries' human capital stock:

The first channel is a sheer level effect: 19th century immigrants exhibited a formal and informal education and training that was better than that of native Brazilians. The immigrants stemmed from societies that already had implemented school systems and provided compulsory formal education for most children (with the exception of Middle Eastern immigrants). They were also better trained for tasks in factories due to earlier industrialization in their home countries. Even though it is true that immigrants often were less educated than the people staying in Europe (Klein 1995), they often were more educated than native Brazilians (Merrick & Graham 1979), a fact that is supported by our data. Merrick and Graham (1979) state, in terms of literacy, that the average education level of immigrants was higher than that of native born Brazilians. They estimate that migrants had twice the literacy rate and three times the level of secondary and higher education than the native population.⁶

The first channel might therefore be simply the higher education of migrants compared to the destination country's population.

The second channel is related to entrepreneurship. Campos Araújo and his co-authors (2006) stress analogies between the mindset of an entrepreneur and a migrant. The motivation to start a new life in a geographically and culturally remote target country is very similar to the willingness to risk and the ambition to start a new business. Also, immigrants are often driven to engage in entrepreneurial activities because of restrictive labor markets. Particularly in 19th and early 20th century Brazil, where there was no social security system whatsoever, immigrants had to start their own businesses to survive if they were not working on plantations or in manufacturing, or after they had paid off their debt of assisted passages to the plantation owners. Most of them successfully did so. For example, by 1934, over 40 percent of the coffee production in São Paulo was produced by the 14.5 percent foreign population of the state, which shows that, as Baer (2001) states, the immigrants were “economically ambitious people”. Additionally, if immigrants wanted to maintain their cultural identity and lifestyle and consume the services and products they were accustomed to, they had to produce them, provide them or initiate trade with the home country to obtain them. This channel would induce trade, craftsmanship and employment in both immigrant communities and among the native population.

A third possible channel could be human capital spillovers from migrant human capital to the local population. The higher education of immigrants could spill over to the native population as they initiated successful behavior. Immigrants built their own schools, hospitals and charitable institutions, in Rio de Janeiro and São Paulo (Klein 1995). These institutions were channels through which the native population might have enjoyed greater access to education and social care. Schwartzman and Brock (2003) also state that German, Italian, and Japanese immigrants created their own schools, sometimes with the support of the governments of their countries of origin, or foreign priests.

In the Brazilian case, one has to bear in mind that the majority of the migrants arrived in the great cities of *Rio de Janeiro* and *São Paulo* via the ports where they disembarked the steamships after their journey across the Atlantic. They mainly went to the center and the southern part of Brazil (Klein 1995). Hence, human capital spillovers would have predominantly taken place in and around these areas with high immigrant concentration, which is particularly strong in the São Paulo region due to the rise in the coffee industry. After 1880 around 70 percent of all immigrants arrived here (Merrick & Graham 1979). This uneven distribution of immigrant human capital spillovers might have contributed, at least partly, to the regional inequality in human capital that persists today. In 1924, the *Associação Brasileira de Educação* was established in Rio de Janeiro with members from different nations and played a very important role in bringing education on the national agenda (Paim 1981 as in Schwartzman and Brock 2003). Later human capital formation has, due to the country's colonial legacy, a long history, which we will now study in detail.

Finally, a fourth channel for spill over might be the fact that female human capital has positive implications for intergenerational transfer. The migration experience sometimes increases the relative importance of female human capital, because families might not keep their traditional gender roles as tightly as in the case of stayers. Moreover, in particular the basic numeracy that we measure with age-heaping method is a component of education which is taught primarily in the household during the first years of life. More self-confident migrant mothers might actually be able to improve this important component of human capital, even if school attendance (and hence literacy) might be not equally developed. In other words, while both education in school and in the household is important for numeracy, the latter might be the less important for literacy.

Human Capital Formation in Brazil 1700-1940: what do we know?

Baer (2001) argues that there was virtually no formal education in Brazil prior to the late 18th century, except for some regionally scattered efforts of Jesuits. They were expelled from the country in 1759, after conflicts with the new minister *Marquis of Pombal*. The Portuguese crown was in need of resources to reconstruct Lisbon after the devastating earthquake in 1755, which is why it confiscated the assets of the fraternity (Oliveira Marques, 1986). Even later, the few existing schools had little impact on the cultural level of the population as a whole, because the schools were understaffed and the teachers poorly paid (Prado Júnior, 1959). Carvalho (1982) stresses that Portugal, as a colonial power, refused systematically to allow the organization of any institution of higher learning in its colonies. In 1768, he further mentions the *Concelho Ultramarino* which was the administrative institution which dealt with overseas issues, denied a request of the captaincy of *Minas Gerais* to build a school of medicine. Only after 1808, when the Portuguese court fled to *Rio de Janeiro* due to the Napoleonic invasion in Portugal, two medical and two military academies were allowed in Brazil. But not only was higher learning severely underdeveloped, also, and more importantly, mass education was lacking. Schwartzman and Brock (2003) state that by 1794 only 179 appointed professional teachers existed in Brazil, whereas in the small country of Portugal the number was around 748. Martínéz-Fritscher, Musacchio and Viarengo (2009) repeat that around a century later, in 1890, Brazil still had one of the lowest levels of literacy in Latin America. Only around 15 percent of the population were able to read and write. Schwartzman and Brock (2003) also argue that throughout the First Republic (1889-1930), primary and secondary education remained the responsibility of local and state governments, and only about 25% of the population, at most, were literate. Large problems in this context were regional disparities in education. In 1834, the responsibility for education was assigned to the provincial governments and for several years the central government subsidized poorer provinces. However, in 1845 these subsidies were abolished and some provinces were unable

to provide an adequate schooling system. As a consequence, the number of children enrolled in public schools declined (Barman 1994, p. 242). These financial difficulties fostered regional disparities within Brazil even more. Most of the improvement in education during the late 19th century took place in Rio de Janeiro and São Paulo or the more prosperous southern regions while the northeast fell behind (Lewis 2006, p.125). These regions are also the ones that received most immigrants during the turn of the centuries. Martínéz-Fritscher et al. (2009) find that educational development was associated with trade and commerce, as states that received export tax revenues invested them in schooling and education. However, they do not account for immigration at the state level. We find numeracy improvements exactly in those states to be associated with knowledge spillover effects of international immigration. Given the fact, that between 1772 and 1900 almost one fourth of all population growth is due to immigration, it is quite plausible that human capital formation in Brazil was influenced by immigrant's human capital. Baer (2001) even argues that public investment in immigration had been a sensible, potent substitute for investing in education.

Data

We use the age-heaping approach to proxy early human capital formation in Brazil. Age-heaping exploits the fact that people living in societies at early stage of human capital formation typically tend to round their age to convenient multiples of five instead of reporting their exact age (see Crayen and Baten 2009). The share of multiples of five to other numbers in the age distribution is expressed in the Whipple Index (A'Hearn, Baten, Crayen 2010). In this study, we apply the ABCC-Index, which is a linear transformation of the latter. It ranges between 0 and 100, where zero stands for an age distribution where everybody states a rounded age and 100 represents no heaping at all. Its coherent construction allows comparability over space and time and different data sources (Crayen and Baten 2009,

Manzel, Baten and Stolz 2011). In recent studies, the indicator has proven very valuable in a wide range of topics in human capital research (see, for example Humphries and Leunig 2009, Stolz and Baten 2011, Prayon and Baten, 2010, for details, also see the internet appendix on the methodology age heaping). For a society at that early stage of human capital formation, the age heaping approach is a suitable indicator to proxy human capital. Usually, societies reach higher numeracy values before they reach better levels of literacy, as knowing one's exact age is a basic skill compared to reading and writing. One of the major determinants of numeracy is, of course, schooling investments as a major input to mass education (Crayen & Baten 2009). However, we should acknowledge some scholars being sceptic about the age heaping methodology, because it is a more indirect indicator of human capital than, say, literacy. Especially the question should be posed whether the introduction of birth registers, the number of censuses and other bureaucratic procedures might increase the knowledge of one's age, but not the education. Age heaping might measure, apart from numeracy, also family stability, the development of local civil government, the spread of vital registration, for example. For example, there is clearly a correlation with vital registration. But previous studies actually found that the residual of regressions of age heaping on schooling investments are not correlated with vital registration and other bureaucratic innovations (Crayen and Baten 2009). For example, the UK had vital registration since the 17th century and developed local government during the 18th and 19th centuries, but as late as 1851 and 1881, age heaping was still particularly strong in the UK. The advantage of age statements is that people did take it serious, but not sufficiently serious to go to the parish priest and counter-check with birth registers. In fact, in world wide comparison, numeracy correlates with the number of censuses and the introduction of birth registers. However Crayen and Baten have found in a regression of numeracy on school enrolment and the number of censuses, only the former has a consistent and significant impact.⁷

Brazilian Numeracy

To estimate the Brazilian numeracy trend, we use census data from five consecutive censuses, 1772, 1830, 1890, 1920 and 1950. For the first time, it is possible to extend the trend back to the very beginning of the 18th century, a period where quantitative data has been scarce so far. One census executed in the region of Sorocaba (São Paulo) in 1772 provides data for the birth cohorts 1700-1740. This is the earliest enumeration of the region analyzed in a quantitative study, so far. It has survived in the *Ultramarino* Historical Archive of Lisbon. Almost the entire São Paulo region was enumerated. We took a sample of 6,278 observations for this study.⁸

Unfortunately, there is no information on age statements of slaves or ethnic background of the free population, so we do not know anything about the ethnic composition of the sample. However, knowing that only free people are included in this enumeration, we had to adjust it, to secure comparability with the other sources, which all contain data on free and enslaved people. To this end, we looked at the Sorocaba district in the later, 1830 evidence, containing data on slaves and free people and we calculated the difference of slave and free ABCC levels, weighted by the number of observations. With this value, we corrected the Sorocaba 1772 evidence to make it comparable to our other sources.

Next, the 1830 evidence refers to enumerations in the states of *Paraná*, *São Paulo* and *Minas Gerais*. For Minas Gerais, there are two very important series of nominative lists: one for 1831-32 and another one for 1838-40. These sets of documents were processed by researchers of the Center of Regional Development and Planning, of the Federal University of Minas Gerais (CEDEPLAR/FACE/UFMG).⁹ They are the result of attempts at producing general censuses in the provinces and were organized by justices of the peace (the lowest judicial authority in post-colonial Brazil), and were commissioned by the Provincial Government. The lists of 1831-32 were used because they comprehend a larger share of the population of Minas Gerais at the time. They contain information about 234 out of the 410

districts (the territory under the jurisdiction of a justice of the peace) of the province, covering 57% of all localities. The data base contains 491,017 inhabitants, 356,267 (72.6%) of them were free and 134,750 (27.4%) were slaves, corresponding to approximately 71% of the estimated total population of the province for the first half of the 1830 decade. The nominative lists of inhabitants of São Paulo (that included the present state of Paraná) are part of a more substantial set of documents called *Maços de População*. These lists give us a large volume of information on the population of the province. The lists of 1836, which we use here, are the last link of a chain of assessments of data on the population, which started to be compiled in 1765. All the lists of the 38 municipalities of São Paulo have been assessed. The data base contains 257,751 inhabitants, 181,849 (70.6%) of them being free and 75,902 (29.4%) being slaves, corresponding to around 79% of the total population of the province. It contains regional data on both sexes and also on their status, i.e. slaves and free people, which is why we consider it to be representative at least for the southern areas of Brazil.

The later samples of the 1890 and 1950 nationwide censuses are representative of the entire Brazilian population per definition. Only the 1920 census might contain an upward bias due to enumerating strictly the district capitals and not the rural areas of the districts.¹⁰

To which degree are the early samples biased? Before the coffee boom and the large migration waves set in after the 1870s, the Southeast region did not have a reputation for being well-developed compared to the Brazilian average (except the urban Federal District of Rio de Janeiro), unlike the South. We can compare the regional numeracy for the birth decade of the 1830s. In this decade, the ABCC indexes in São Paulo and Minas Gerais were some 3 percent below the national average, whereas the one on Paraná was some 3 percent above the national average. Assuming similar regional differences for the inhabitants in the 1830 population lists, the slight negative regional bias cannot be very pronounced (around minus 1 percent). For the 1772 census, in which only São Paulo is represented, it might be of the order of a 3 percent negative bias (i.e., the Brazilian national figure might be even higher). A new

source of evidence even allows us to countercheck for the early modern period, whether the south was initially less numerate than the north. We look for regional differences in the earliest period using inquisition files of the Portuguese Sanctum Officium which include a sufficiently large number of cases with age-statements of the involved Brazilians from all parts of the country.¹¹ If we adjust the 1920 census which contains only provincial capitals, downward by 4 percent, the values for the birth decades of the 1860s and 1890s would also be identical to the respective birth decades in the censuses of 1890 and 1950.

Our numeracy estimates for Brazil during the 18th and 19th century are arranged by birth decade averages and allow us to shed light on human capital formation in Brazil from 1710 onwards. We use birth decades because basic numeracy is normally acquired during the first decade of life (Crayen and Baten, 2009).¹² Hence, every census provides us with four decades of numeracy values. We start to observe Brazilian numeracy in 1700 at a level of 63 ABCC points (Figure 3). This is similar to the value of Hungary at that time and better than other Latin American countries like Argentina or Mexico (Manzel, Baten and Stolz 2011). Between 1740 and 1760, the level of numeracy increases somewhat, and remains high during the later period of the 18th century. The level of the 1830 birth cohort is significantly lower than the peak in 1790, suggesting a deterioration of human capital formation during the Napoleonic Wars and the following independence of Brazil in spite of being a quite peaceful event, by Latin American standards. This stagnation in numeracy trends can also be observed in other Latin American countries (Manzel, Baten and Stolz 2011). From the 1840s onwards, Brazil started a sustained upward trend, which accelerated between 1870 and 1910, a period with very high immigration rates.¹³

Does this correspond to what we would have expected from Brazilian economic history? The eighteenth century in Portuguese America was characterized by the rise of south-central region of the colony, thanks to the discoveries of gold in Minas Gerais and, to a lesser extent, Mato Grosso and Goiás. Minas Gerais was still the major producer of gold until at

least 1760s. The importance assumed by the gold mining led to a complete redefinition of the colonial economic space. The challenge to supply the mines promoted the articulation of regions that previously had restricted economic activities and was focused in subsistence production. Thus, the captaincies of the South have become producers of food and animals to the mines. The interior of the Northeast also turned to the supply of Minas Gerais, especially cattle. Hence yes, the mild increase of 18th century numeracy corresponds to the general development.¹⁴

From the 1770s, this situation began to change. On the one hand, the production of gold in Minas Gerais (the one that remained significant in the second half of the eighteenth century) began a process of decline. According to statistics of collected taxes on gold by Wilhelm Ludwig von Eschewege the year 1771 can be considered one in which the decrease in revenue was found to be irreversible. Thus, the peak reached in 1754, when it was collected 1770 kilos of gold, came to 1200 kilos in 1771, reaching 450 kilos in 1800 and 30 kilos in 1820 (Eschewege, 1944, p. 364-369). The decline of gold was not the economic decline of the region, but the return to production for the domestic market, which then sustained its economic growth, including the continuation of significant purchases of African slaves.¹⁵ During the 19th century, the city of Rio de Janeiro became the capital of the Empire of Brazil and the growth of coffee cultivation in the province of Rio de Janeiro ensured the continuity of economic growth in this region. In 1850, the slave trade was forbidden, which sharply reduced their volume to the complete disappearance after 1856.

The decades from 1850 onwards witnessed the strong expansion of coffee plantations. Such economic growth was initially supported by slave labor that had accumulated in the region after the sharp rise in Atlantic slave trade in the last two decades it endured. Later there was the use of internal slave trade, with the provinces of Rio de Janeiro and Sao Paulo absorbing much of the slaves from the northeastern provinces. From the last years of the 1880s, at the end of Brazilian slavery, European immigration came up with an alternative

supply of labour to the strong expansion that coffee was known in its new frontier, the west of Sao Paulo. A positive development of numeracy is visible in Figure 3, after immigration replaced slave imports and interregional slave trade.

Immigrant numeracy

The immigrant data was preserved in the National Archive of Rio de Janeiro. The archive contains passenger lists of ships that arrived in Brazil from the beginning of the 19th century. However, until 1822, Portuguese immigrants were not explicitly declared as Portuguese; because Brazil was still part of the Portuguese empire and the Portuguese king even resided there after Napoleon invaded Portugal in 1808. We took samples of various arrival years between 1812 and 1932. The largest part of the sample stems from the period between the 1870s to the 1920s. This is the period in which mass migration to Brazil really took off. The immigrants of our sample were recorded at the ports in *Rio de Janeiro* and *Santos*, which were the most important ports. According to Ferenczi and Willcox (1929), around 85 percent of all immigrants entered Brazil via those ports. Italians were the largest immigrant group, followed by the Portuguese, who have an exceptionally long migration history with Brazil due to the colonial era. Third was Spain, followed by German speaking immigrants and migrants from Eastern European countries. There was also some immigration from Middle Eastern countries during the increasingly turbulent years before and after the fall of the Ottoman Empire (Klein, 1995). Our data set reflects this distribution quite well (Figure 4). In addition, the port and year of arrival, origin, ship and age were given. Numeracy estimates for immigrants are also arranged by birth cohorts. They show that except for Lebanese migrants, all immigrant groups are more numerate than native Brazilian birth cohorts (Figure 5).¹⁶

Between 1851 and 1920, Brazil received more than 3.5 million Europeans and other immigrants. Of this total, 39% were Italian, 29.6% were Portuguese and Spaniards were 3.6%.

This volume of immigrants makes Brazil the third largest destination of European mass migration, behind only the United States and Argentina.

What about the skills of European immigrants? There is a historiographic debate about the role of immigration in the development of the states that received the larger number of immigrants in Brazil (São Paulo, Rio Grande do Sul, Paraná and Santa Catarina). Merrick and Graham (1996, p. 93-4) report that Italians, for example, were predominantly recruited from São Paulo's farmers in the Northern Italy, given the belief that there was a more promising human capital in this region. But Leff (1982, p. 69-70) argues that although the immigrants had better literacy levels than Brazilians, this was not a crucial factor in the role played by immigrants in the development of such regions. This debate remained unsolved before the present study, because the literacy evidence allowed both interpretations.¹⁷

How path dependent is human capital formation?

Our strategy to assess possible human capital spillovers econometrically follows a two-step approach. First, we investigate the impact of 1900 human capital on real GDP 2000 in a cross-country setting. This tests our hypothesis of a long-run path dependency of human capital formation on economic performance of an economy as a whole. After verifying the relationship, we calculate a counterfactual in which we simulate a human capital situation of Brazil in 1900 without the immigrant's human capital, allowing us to compare the migration versus non-migration scenario. In a second step, we run a regional cross-section of immigrant share increase across Brazilian states on ABCC increase. The hypothesis is that the immigrant's share is positively correlated with ABCC increase after controlling for other influences such as educational expenditures at the state level. To this end, we use data from the period of 1890 to 1920, the time in which Brazilian mass immigration reached its peak.

In the first step, our argument is that economic growth is *ceteris paribus* determined by current human capital stock, which in turn is determined by human capital in the past. This argument points to the direction of the work of Acemoglu, Johnson and Robinson (2001) that stresses the importance of past events on economic performance today, although they focus on institutions rather than human capital. Glaeser, La Porta, Lopez-de-Silanes and Shleifer (2004), disagree on the fact that institutions should be at the focus of the analysis and point to the importance of human capital, being a “more basic source of growth”, which in turn shaped institutions and brought about the path of dependency for economic development. In a more recent study, Prayon and Baten (2010) test this relationship with an extended data base for human capital. They find that it is indeed the human capital channel that enhances institutions which affect economic output in the very long run. Figure 6 shows the long-run relationship between real GDP per capita and numeracy performance in the year 1900. With the exception of some oil-exporting countries, we find a stable linear relationship. Those, who performed well in terms of human capital at the beginning of the twentieth century, are among the top-performers a century later.

What other variables should be of interest? Aggregate output is a function of physical capital, labor and human capital at a given technological level. Labor can be ignored, as we look at per capita values of GDP. What about physical capital? Baier, Dwyer and Tamura (2006) provide data on the physical capital endowment for a broad range of countries. Their work suggests that the levels of physical capital across world regions varied considerably in 1900.¹⁸ Sokoloff and Engerman (2001) argue that geography and factor endowment play a crucial role in determining development outcomes across countries. Therefore, we include a set of geographical and institutional variables, such as a variable for the political system, civil war, oil production, longitude, latitude and a dummy variable for landlocked geographical situation.¹⁹

We regress log GDP per capita on ABCC values for 90 countries in the year 1900 (Table 1).²⁰ The numeracy level in 1900 is highly significant. An increase in numeracy in 1900 by one ABCC point leads to a three percent increase in GDP per capita in the year 2000. Institutional aspects and geography render significant results, as well. We explain around 60 percent of the inter-country variance in real GDP with our model. Comparing Column 1 and 2, only some 10 percent are explained by world region fixed effects. There is an unexplained part, as, of course; today's economic policies and factor endowments should also play a crucial role in determining real GDP per capita today. Yet, we find considerable evidence for path dependency of human capital formation. This result implies that the countries, which had resolved their numeracy problem by 1900 like Norway or the US, are the states that are doing well in terms of economic performance in the year 2000, as well. Cases with higher GDP than expected are states like Indonesia, India, Kuwait, Bahrain, Egypt and Morocco.

An important issue is endogeneity. It is quite likely that wealthy economies can afford better educational systems and therefore achieve a more rapid human capital accumulation. Instrumental variable techniques could be helpful, if there might be (a) potential endogeneity (which might be somewhat less likely, as the explanatory variable is measured 100 years before the dependent variable) and (b) measurement error, which can of course be a problem here. We included as an instrumental variable for numeracy the number of books per capita in the world region around 1750. Although the number of cases is quite small for this regression (which would increase the likelihood of insignificant coefficients), the human capital variable is still significant (Column 4).

How well does our estimation perform in predicting real GDP per capita in the year 2000 in Brazil? If we consult the Penn World Tables, we find that our estimation is quite close to the actual value of Brazil's real GDP per capita in the year 2000. Whereas we would predict 7584.65 US\$, the PWTs display an actual value of 7455.9 US\$.

In a next step, we calculate a counter-factual to answer the question whether the human capital incorporated in Brazil's immigrants had any long-term effects on Brazilian long-run economic performance. In 1900, the Brazilian population consisted of 17.4 million people. According to Klein (1995), seven percent of them were immigrants. The ABCC for the 1900 birth cohort of the whole population, including migrants is 89.7 percent. From our data, we are able to calculate the total average ABCC for immigrants, born in 1900, which yields 94.1 percent. This means that the average immigrant numeracy was around five percentage points higher than numeracy levels for the entire population. With these figures, we can calculate the numeracy value for native Brazilians only, conditional on their population share, which renders 89.3 percent. Filling this into our estimated relationship between ABCCs in 1900 and real GDP per capita in the year 2000, we find that average real income per capita would have been around 75 US \$ lower, if the migrants had not come to Brazil and brought their numeracy skills with them. 75 US \$ look like a small amount at first glance. How much is that worth in Brazil? 75 US \$ equaled 134 Brazilian *Reais* in the Year 2000.²¹ For comparison: the minimum wage was around 260 *Reais*, so this amount would have been roughly one half of a minimum wage. Or put differently, this equals more than the amount that ca. 44 million Brazilians receive with the "*Fome Zero*" program, which was introduced in 2002 by the Brazilian president *Inacio Lula da Silva*. It is aimed at the very poor of the Brazilian society and provides them with money or coupons to buy foodstuffs. Hence, this money would be a modest amount in general, but still a quite significant share of real income to many poor Brazilians, even today. Moreover, this is a lower bound estimate, because first, before 1900, the gap between natives and immigrants was larger. In the early 19th century, immigrants had an ABCC advantage over natives of around 20 percent (Figure 5b). Second, we measured only the level effect of one generation of migrants. Third, we do not capture any spillover effects to native numeracy in this regression. And last but not least,

many events in the turbulent twentieth century certainly distorted the picture and added to the downward bias.

Numeracy spillovers on a regional Level

How do we measure human capital spillovers? In this case, we look at the regional distribution of immigration and analyze whether the immigrant increase that Brazilian states experienced between 1890 and 1920 is correlated with the regional ABCC-increases during this period. Using data from the 1890 and 1920 censuses of Brazil, we generate a regional distribution of the evolution in numeracy between those two points in time (Figure 7).²² This pattern suggests that immigration had something to do with the evolution of human capital in Brazil at the regional level. The biggest ABCC increases we find are in the states that were the main destinations for immigrants like Rio de Janeiro, São Paulo and Minas Gerais. In contrast, states with low immigration increase had also low ABCC increase, such as Rio Grande do Norte and Piauí. We now analyze whether the ABCC increases are correlated with regional immigrant increase over the respective period. To control for educational expenditures, we use the data provided by Martínéz-Fritscher, Musacchio, Viarengo (2009). These authors provide a very useful data base on a regional level for Brazil around the turn of the century. We use their data on education expenditure. Martínéz-Fritscher et al. argue that an improvement in education was predominantly possible in states with higher income from commodity exports. They further state that immigration could not have played a role, because the immigrants were poorly educated and the great improvement in education took place only after the bulk of immigrants already had entered Brazil. We do not argue that their main argument of commodity export taxes improving state educational expenditures does not hold, but we argue that states which generated such income were also states that experienced significant immigration, because export economies were the labor markets targeted by the

immigrants. Moreover, a significant proportion of European immigrants came from countries that already had a schooling tradition, which was reflected in literacy rates higher than those observed in Brazil. Finding themselves in a situation where there were no public schools, immigrants were involved in movements for public education or in the organization of networks of community-based ethnic schools (Kreutz 2007). Community schools were more developed in the homogeneous nuclei of national communities in the larger cities (Kreutz, 2007). Ethnic schools were especially important for German, Italian, and Japanese immigrants. These national groups also had the support of their countries of origin, which provided educational materials and even financed the operation of many schools. In urban areas ethnic schools grew until the beginning of the 1920s. Later, their decline began in the cities of São Paulo and Minas Gerais. In the states of Rio Grande do Sul, Santa Catarina, Parana and Sao Paulo, which had large concentrations of immigrants in rural communities, ethnic schools continued to grow until the 1930s. Hence, we argue that in states that were open and integrated into the world market because they had a valuable commodity to trade, migration also played an important part in boosting education. Both, trade and migration take place at the same time, because they are both characteristics of open economies.

There is a correlation between the growth rate of education expenditures between 1890-1920 and immigrant increase across states, the correlation coefficient is as high as 0.55 (significant at the one percent level, p-value: 0.015) shows. In our regressions, we run the ABCC increase between 1890 and 1920 and between 1890 and 1950 on the log of the absolute immigrant increase per state (Table 2 and 3). Additionally, we control for log population density, because more densely populated regions might also have a better infrastructure to provide education to their citizens. Furthermore, we control for initial ABCC levels, because the more numerate the people in the state already are, the less room might be available for improvement, since the ABCC index is an indicator of basic human capital. Of course, state educational expenditures are also important in this context. We control for this in

two different ways, first, we take average state educational expenditures between our two time points and in another specification, we only control for the initial value of state expenditures in 1890, as the increase in this variable is correlated with the increase in immigrants throughout states. In our third regression, we additionally included a set of dummy variables that control for certain export commodities like coffee, cattle and cotton or mate.

We ran three different specifications, where we varied the measures of educational expenditures of the states and additionally included the export industry dummy variables. Throughout all specifications, the log of immigrant increase in the states has a significant effect of substantial and robust magnitude. If we calculate standardized beta coefficients, the effect is the most important in the first two specifications.²³ None of the commodity expenditure dummies becomes significant. The log of population density is only significant in the first two specifications, whereas the initial ABCC level is always significantly and negatively correlated with the ABCC increase of the respective state, a result we would have expected. To test the robustness of these results, we ran the same regression on the ABCC increase between 1890 and 1950 across states (Table 3). The results stay robust; immigrants seem to have produced an impact on human capital development across Brazilian states.²⁴

Conclusion

In this article, we studied the long-term human capital formation in Brazil. We constructed a numeracy trend for Brazil that extends back to the beginning of the 18th century. With a new sample on human capital endowment of 19th century mass migrants, we showed that international migration had a positive effect on the stock of human capital in Brazil. Looking at regional data for Brazil, we showed that human capital spillovers took place predominantly in those states that received the most migration in absolute numbers. Increases in numeracy are significantly positively correlated with immigration, even after controlling for educational

expenditures on the state level. This suggests that human capital grew strongest in those states where most immigrants arrived. Could there be crowding-out effects in a way that Brazilians do not have education incentives, because the education-intensive jobs are increasingly filled with immigrants? We actually found no evidence for this theoretical possibility. Human capital intensive migration in contrast might stimulate firm creation and also provide teachers and institutions to provide education even for native Brazilians.

We also found that a modestly sized, but appreciable significantly positive effect of this improvement of the human capital stock on economic performance persists until today. Hence, human capital in its various dimensions incorporated in international migrants can affect human capital endowment, which in turn determines economic development and performance of destination countries. This relationship is crucial in understanding today's human capital selective immigration policies of major immigrant countries like the US, the UK, Canada, or Australia.

References

- Acemoglu, D., Johnson, S. and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review* 91(5), 1369-1401.
- Andrade, C. V. de (2001). *Domicílios mineiros oitocentistas: uma aplicação do método Grade of Membership (GoM)*. Belo Horizonte: UFMG, (Unpublished MA. Dissertation).
- Andrade, C. V. de (2009). *Population in Nineteenth-century Vila do Conde: the demographic dynamics of a northwestern Portuguese urban parish*. Cambridge: (Unpublished PhD Dissertation).
- A'Hearn, B., Baten, J., and Crayen, D. (2010). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. *Journal of Economic History* 47(1), 82-99.
- Baer, W. (2001). *The Brazilian Economy: Growth and Development*. Sixth edition. Boulder.
- Baier, S.L., Dwyer, G.P., and Tamura, R. (2006). How important are Capital and Total Factor Productivity for Economic Growth? *Economic Inquiry* 44(1), 23-49.
- Balderas, J.U., and Greenwood, M.J. (2010). From Europe to the Americas: a comparative Panel-data analysis of migration to Argentina. *Journal of Population Economics* 23, 1301-1318.
- Barman, R.J. (1994). *Brazil: The Forging of a Nation, 1798-1852*. Stanford, Stanford University Press.
- Barro, R.J. and Lee, J.W. (2001). International data on educational attainment: updates and implications. *Oxford Economic Papers*. Oxford University Press.
<http://oep.oxfordjournals.org/content/53/3/541.short>, last accessed, May 2nd, 2011.
- Basile, M. (2006). Festas cívicas na Corte regencial” *Varia Historia* 22(36), 494-516.

- Bassanezi, M. S. C. B., Scott, A. S. V., Bacellar, C. A. P., Truzzi, O. M. S. (2008). *Atlas da imigração internacional em São Paulo, 1850-1950*. São Paulo: Editora da UNESP, FAPESP.
- Beine, M., Docquier, F. and Rapoport, H. (2008). Brain Drain and Human Capital Formation in Developing countries: Winners and Losers. *The Economic Journal* 118 (April), 631–652.
- Bergad, L. (1999). *Slavery and the demographic and economic history of Minas Gerais, Brazil, 1720-1888*. Cambridge: Cambridge University Press.
- Borjas, G. J. (1994). The Economics of Immigration. *Journal of Economic Literature* 32(4), 1667-1717.
- Breu, C. de (1976). *Ensaio e Estudos*. Rio de Janeiro: Civilização Brasileira; Brasília: Instituto Nacional do Livro.
- Bruecker H, and Defoort, C. 2006. „The Self-Selection of International Migrants Reconsidered: Theory and New Evidence.” IZA Discussion Paper Series. IZA DP 2052.
- Campos Araújo, J.R. de, Cruz Paiva, O. de and Rodriguez, C.L. (2006). Immigrants and Entrepreneurs in São Paulo, Brazil: Economic Development in the Brazilian ‘Melting Pot’. In: Galbraith, C.S. and Stiles, C.H. (eds): *Developmental Entrepreneurship: Adversity, Risk and Isolation. International Research in the Business Disciplines*, Oxford, Vol. 5, 297-322.
- Capolupo, R. (2009). The New Growth Theories and Their Empirics after Twenty Years. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 3.
- Carvalho, J. M. de (1982). Political elites and States Building. The Case of 19th Century Brazil. *Comparative Studies in Society and History* 24(3), 378-399.

- Carvalho Filho, I. and Colistete, R.P. (2010). Education Performance: Was it all determined 100 years Ago? Evidence from São Paulo, Brazil. <http://mpra.ub.uni-muenchen.de/24494/>MPRA Paper No. 24494.
- Cohen, D. and Soto, M. (2007). Growth and Human Capital: Good Data, Good Results. *Journal of Economic Growth* 12, 51-76.
- Cohn, R. L. (2009). *Mass Migration under Sail. European Immigration to the Antebellum United States*. Cambridge University Press, Cambridge.
- Crayen, D., d Baten, J. (2009). Global Trends in Numeracy 1820–1949 and its Implications for Long-term Growth. *Explorations in Economic History* 47(1), 82-99.
- Drinkwater, S., Levine, P., Lotti, E. and Pearlman, J. (2003). The Economic Impact of Migration: A Survey. Hamburg Institute of International Economics FLOWENLA Discussion Paper No 8. ISSN 1616-4814.
- Englebert, P. (2000). Solving the Mystery of the AFRICA Dummy. *World Development* 28(10), 1821-1835.
- Eschwege, W. L. von (1944). *Pluto Brasiliensis. memórias sobre as riquezas do Brasil em ouro, diamantes e outros minerais, v. 1*. Rio de Janeiro: Companhia Editora Nacional.
- Fausto, B. (ed.) (1999): *Fazer a América: a imigração em massa para a América Latina*. São Paulo: Edusp.
- Ferenczi, I and Willcox, W. F. (1929). *International Migrations. Volume I. Statistics*. In: Easterlin, Perlman and Thomas (eds): *Demographic Monographs*. A series of economics reprints. New York.
- Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2004). Do institutions cause growth? *Journal of Economic Growth* 9(3), 271-303.
- Grossman, G.M. and Helpman, E. (1991). Quality Ladders in the Theory of Growth. *Review of Economic Studies* 58, 43-61.

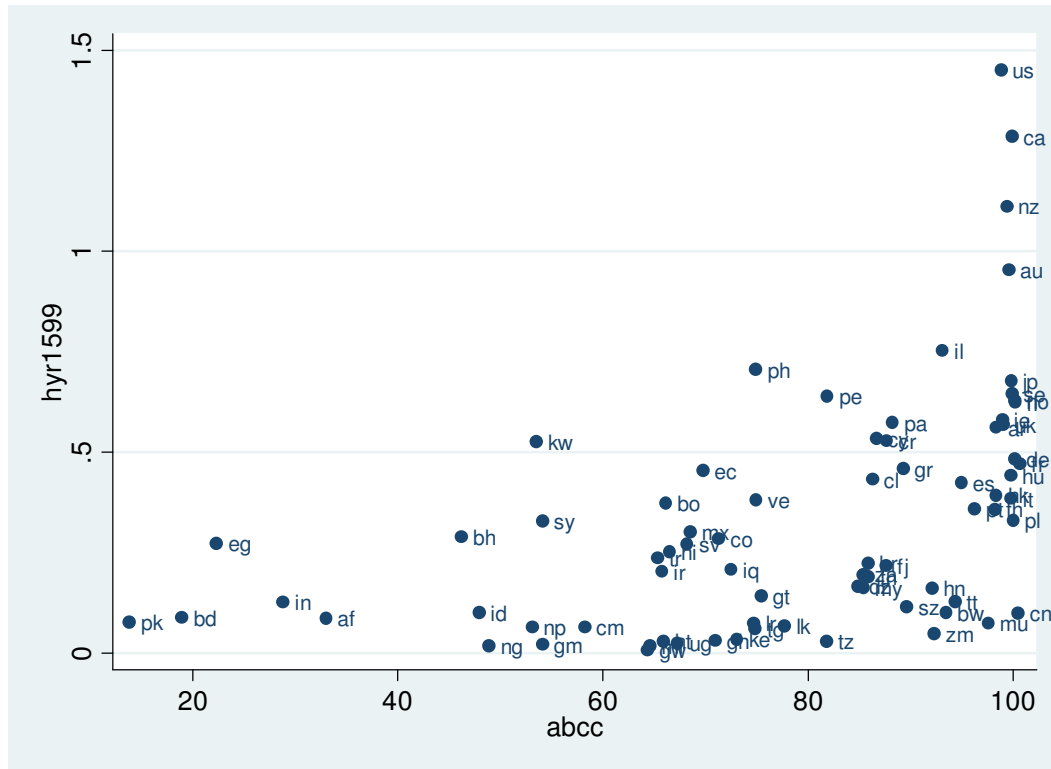
- Hatton, T. J. and Williamson, J. G. (1998). *The Age of Mass Migration: Causes and Economic Impact*. Oxford University Press. New York.
- Hatton, T. J. and Williamson, J. G. (2008). *Global Migration and the World Economy: two Centuries of Policy and Performance*. MIT Press. Michigan.
- Holloway, T. H. (1980). *Immigrants on the land: coffee and society in São Paulo, 1886-1934*. Chapel Hill, NC: University of North Carolina Press.
- Humphries, J. and Leunig, T. (2009). Was Dick Whittington Taller than those left behind? Anthropometric measures, migration and the quality of life in early 19th century London. *Explorations in Economic History* 46(1), 120-131.
- Juif, D.; Baten, J. (2011) On the Human Capital of 'Inca' Indios before and after the Spanish Conquest, *Working Paper University of Tuebingen*.
- Kang, T. (2010). *Instituições, Voz Política e Atraso Educacional no Brasil, 1930-1964*. (Unpublished M.A. Diss., Universidade de São Paulo).
- Klein, H. S. (1995). European and Asian Migration to Brazil. In: Cohen (ed): *The Cambridge Survey of World Migration*. Cambridge University Press.
- Kreutz, L. (2007). A educação de imigrantes no Brasil. In: Lopes, F. F., and Veiga (eds.). *500 anos de educação no Brasil*. Autêntica. Belo Horizonte.
- Leff, N. (1982). *Underdevelopment and development in Brazil: economic structure and change, 1822-1947, vol. 1*. London: Allen & Unwin.
- Lewis, W.A. (2006). *Tropical Development, 1880-1913*. London and New York, Taylor and Francis.
- Lucas, R.E., jr. (1988). On the Mechanics Economic Development. *Journal of Monetary Economics* 22, 3-42.
- Lundborg and Segerstrom (2000) "International Migration and Growth in Developed Countries: A Theoretical Analysis" *Economica*, November 2000, 579-604.

- Lundborg and Segerstrom (2002). The Growth and Welfare Effects of International Mass Migration, *Journal of International Economics* 2002, 177-204.
- Manzel, K.; Baten, J., Stolz, Y. (2011). Convergence and Divergence of Numeracy: The Development of Age Heaping in Latin America, 17th to 20th Centuries. *Economic History Review*, forthcoming.
- Martínéz-Fritscher, A. C., Musacchio, A., Viarengo, M. (2009). The Great Leap Forward in Education in Brazil: The Political Economy of Education in Brazil, 1890-1940. *Harvard Business School Working Paper 10-075*.
- Merrick, T.W. and Graham, D.H. (1979). *Population and Development in Brazil 1800 to the Present*. Johns Hopkins University Press. Baltimore and London.
- Messias, E. (2003). Income Inequality, Illiteracy Rate, and Life Expectancy in Brazil. *American Journal of Public Health* 93(8), 1294-1296.
- Mokyr, J. (1983). *Why Ireland Starved*. London: George Allen & Unwin.
- Monteiro, M. (2000). *Migrantes, emigrantes e brasileiros (1834-1926)*. Fafe.
- Murtin, F. and Viarengo, M. (2010). American Education in the age of mass migrations 1870-1930. *Cliometrica* 4, 113–139.
- Oliveira Marques, A. H. (1986). *História de Portugal*. Lisboa.
- Paiva, C. A. (1986). Minas Gerais no século XIX: aspectos demográficos de alguns núcleos populacionais. In: COSTA, I. D. N. da (org.). *Brasil: história econômica e demográfica*. São Paulo: IPE/USP, 173-187.
- Paiva, C. A. (1996). *População e economia nas Minas Gerais do século XIX*. São Paulo: USP, (Unpublished PhD. Dissertation).
- Paiva, C. A., Arnaut, L. D. H. (1990). Fontes para o estudo de Minas oitocentista: listas nominativas, V Seminário sobre a Economia Mineira, Anais, Belo Horizonte: CEDEPLAR/FACE/UFMG, 85-106.
- Prado Júnior, C. (1959). *História Econômica do Brasil*. São Paulo. Fifth edition.

- Prayon, V. and Baten, J. (2010). Human capital, institutions, settler mortality, and economic growth in Africa, Asia and the Americas. *University of Tuebingen Working Paper*.
- Reichlin, P. and A. Rustichini, 1993, Diverging patterns in a two country model with endogenous labor migration, Discussion paper no. 9332 (CORE, Louvain-la-Neuve).
- Reis, J. (2008). How Helpful is Age Heaping? An Exploration using the Portuguese Census of 1940. *Working Paper Instituto de Ciências Sociais, Univ. Lisbon*.
- Rodarte, M. M. S. (2008). O trabalho do fogo: Perfis de domicílios enquanto unidades de produção e reprodução na Minas Gerais. Belo Horizonte: UFMG, (Unpublished PhD. Dissertation).
- Rodrigues, H. (2006). *Emigração e emigrantes: Vale do Lima no século XIX*. Viana do Castelo.
- Romer, P.M. (1994). On the Origins of Endogenous Growth. *Journal of Economic Perspectives* 8(1), 3-22.
- Sachs, J.D. and Warner, A.M. (1997). Fundamental Sources of Long-Run Growth. *The American Economic Review* 87(2), 184-188.
- Schwartz, S. (1992). *Slaves, peasants and rebels: reconsidering Brazilian slavery*. Illinois: Illinois University Press.
- Schwarzman, S. and Brook, C. (2003). *The challenges of Education in Brazil*. Oxford University, Centre for Brazilian Studies.
- Slenes, R. W. (1999). *Na senzala, uma flor: esperanças e recordações na formação da família escrava (Brasil Sudeste, século XIX)*. Rio de Janeiro: Nova Fronteira.
- Sokoloff, K. and Engerman, S. (2000). History Lessons. Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives* 14(3), 217-232.
- Stark, O., Helmenstein, C. and Prskawetz, A. (1997). A brain gain with a brain drain, *Economics Letters* 55, 227–34.

- Stolcke, V. (1988). *Coffee planters, workers and wives: class conflict and gender relations on São Paulo, 1850-1980*. Oxford: Macmillan, St. Antony's College Oxford.
- Stolz, Y. and Baten, J. (2011). Brain Drain in the Age of Mass Migration: Does Relative Inequality Explain Migrant Selectivity? *Working Paper Univ. Tuebingen*.
- Taylor, A. (1997). Peopling the Pampa: On the impact of mass migration to the River Plate, 1870-1914. *Explorations in Economic History* 34, 100-132.
- Todaro, M.P. (1969). A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *The American Economic Review* 59(1), 138-148.

Figure 1: Numeracy 1900 versus higher educational attainment to total population in the year 2000 (Barro and Lee 2000, Prayon and Baten 2010, Crayen and Baten 2009)



Notes:

All country abbreviations are ISO 3166,

(http://www.metatab.de/meta_tags/laenderkuerzel.htm)

Figure 2: Higher educational attainment 2000 (Barro and Lee, 2000) as a major determinant of log GDP per capita 2000 (Penn World Tables)

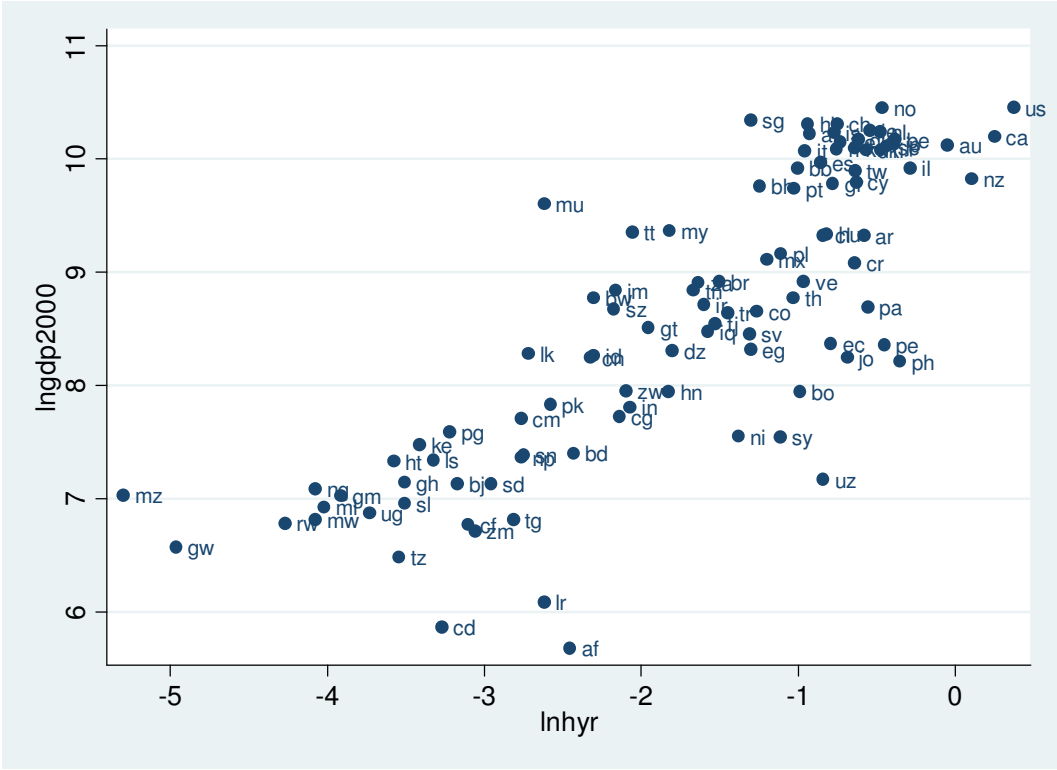
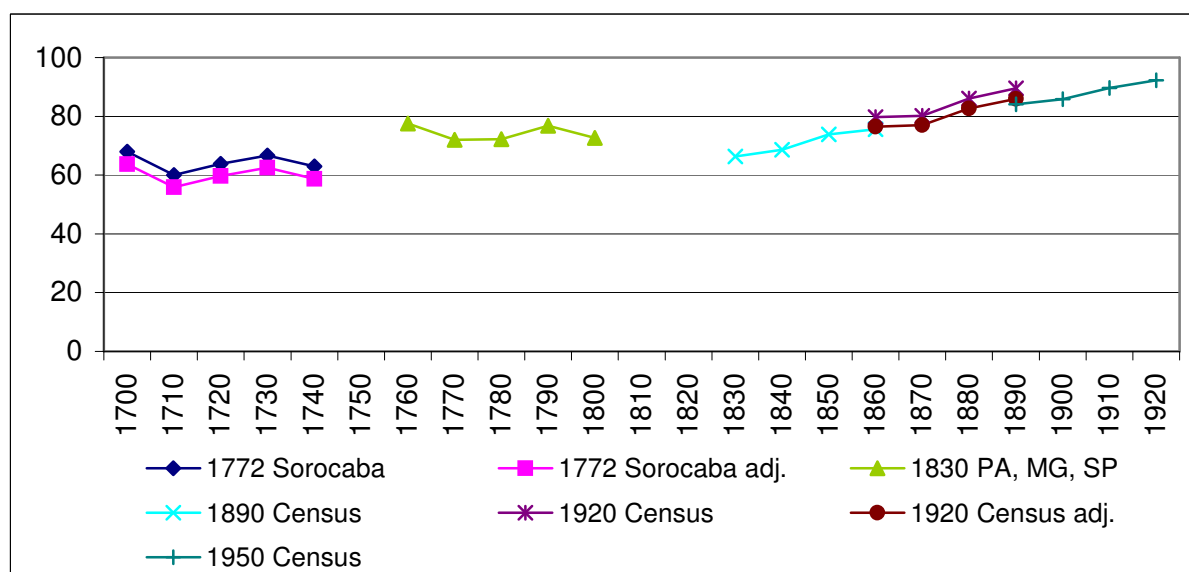


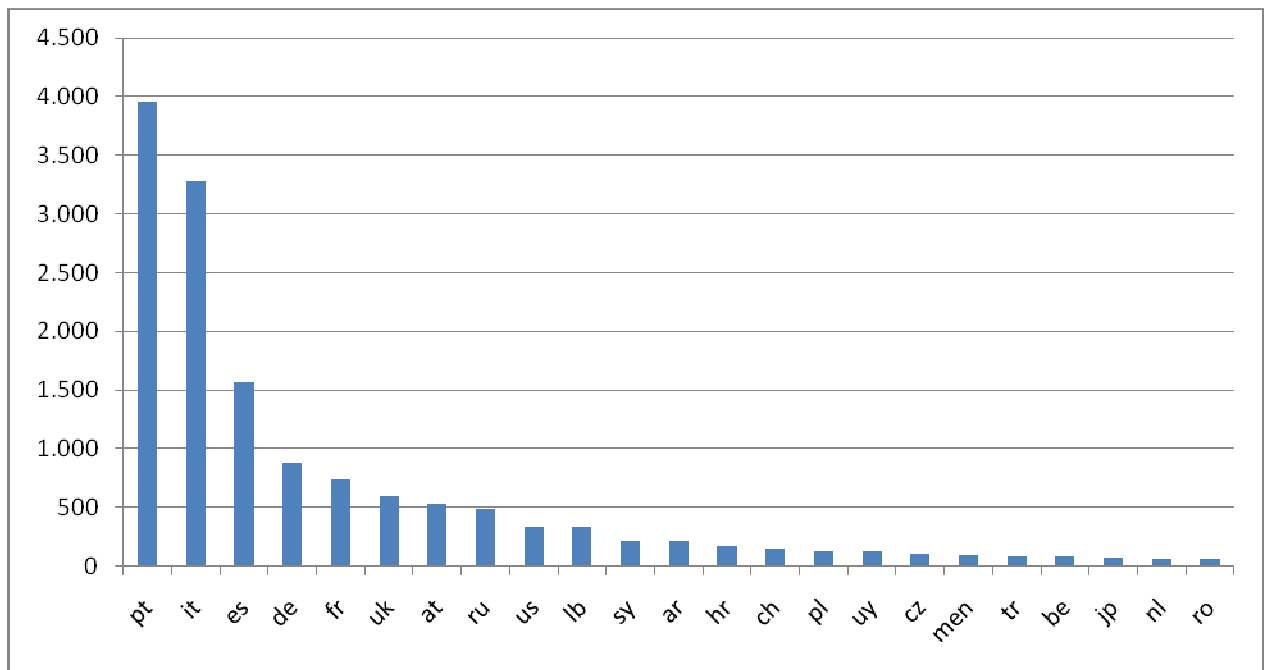
Figure 3: Numeracy estimates for Brazil 1700-1910. Sources see Text. Author's calculations.



Data: Author's calculations, sources, see text

Note: We looked at the Sorocaba district in the later, 1830 evidence, containing data on slaves and free people and we calculated the difference of slave and free ABCC levels, weighted by the number of observations. With this value, we corrected the Sorocaba 1772 evidence to make it comparable to our other sources.

Figure 4: Immigrants by Source Country n>50.



Notes:

Author's calculations. Data from archival records of the National Archive in Rio de Janeiro and <http://www.an.gov.br/sian/inicial.asp>. MEN stands for immigrants that were listed to be "arabe", "ottoman", "orient", "oriental" with no source country more specifically given.

Figure 5: Immigrant Numeracy versus Brazilian Numeracy 1770-1910. (Selected Countries)

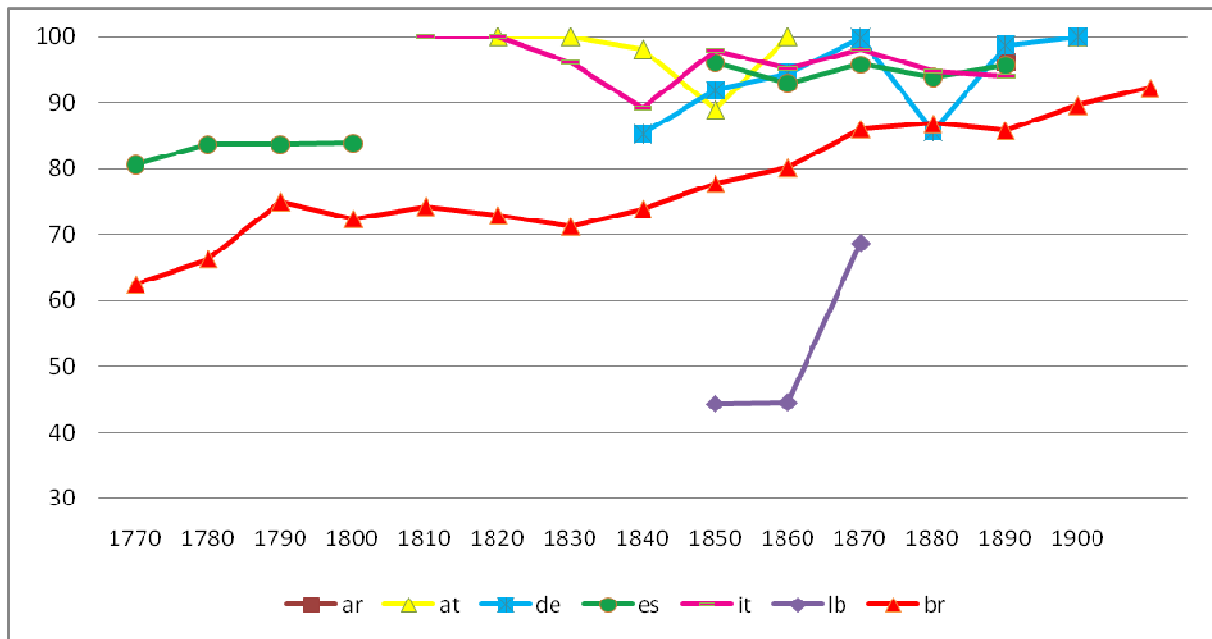


Figure 5b: Immigrant Numeracy versus Brazilian Numeracy 1770-1910. (Selected Countries)

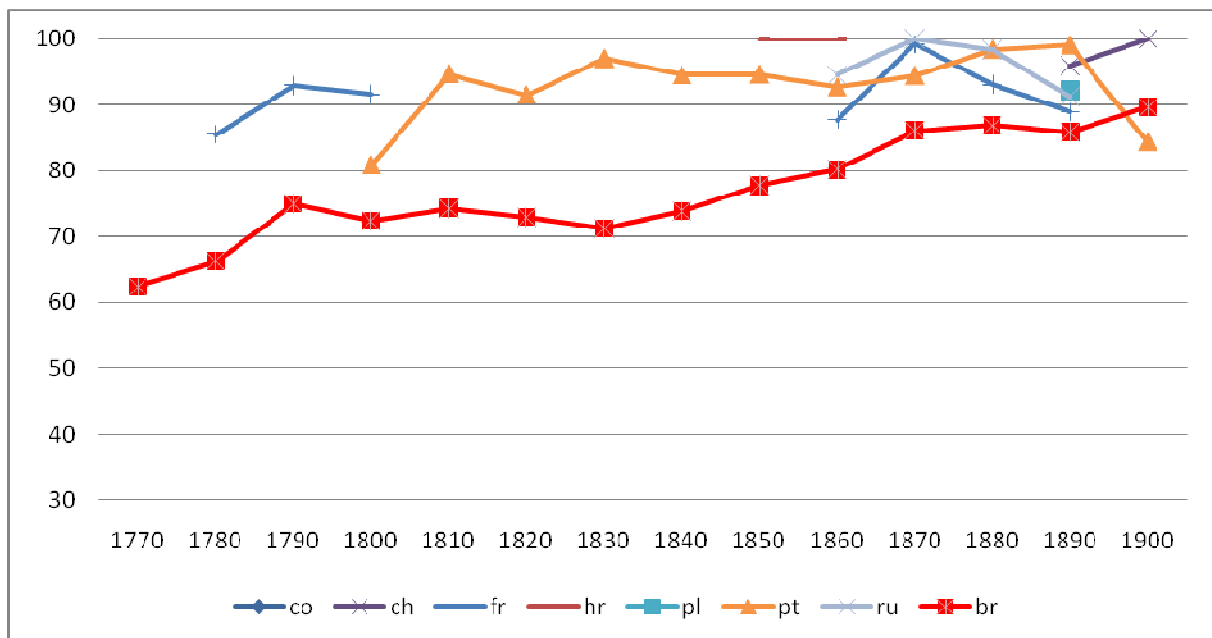
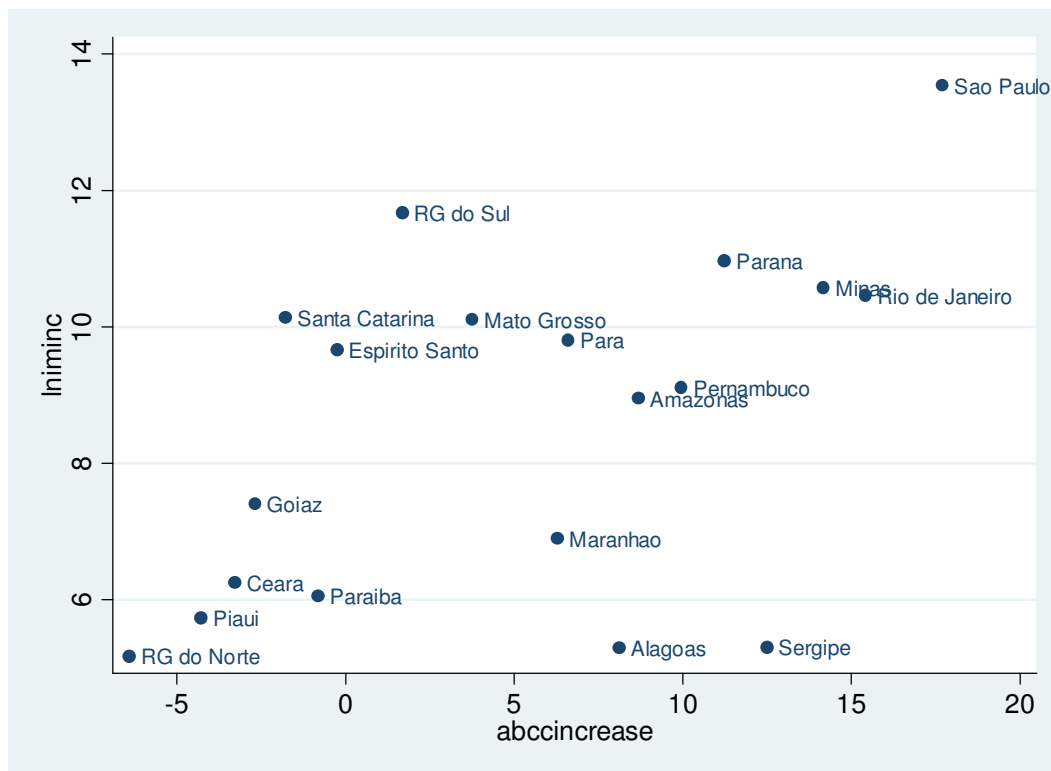


Figure 7: Log immigrant increase versus increase in ABCC levels between 1890 and 1920.



Data: ABCCs: Crayen and Baten 2009, Prayon and Baten 2010, Log immigrant increase:

Anuario estatístico do Brazil 1908-1912. Rio de Janeiro: Directoria Geral de Estatistica, v. 1-3, 1916-1927 and Anuário estatístico do Brasil 1938. Rio de Janeiro: IBGE, v. 4, 1939.

Table 1: Long term impact of human capital formation on economic performance

	(1)	(2)	(3)	(4)
Numeracy in 1990	0.029*** (0.000)	0.039*** (0.000)	0.027*** (0.000)	0.023*** (0.000)
Democracy			0.020 (0.293)	
Longitude			0.003 (0.281)	
Latitude			0.011* (0.082)	
Landlocked			0.295 (0.244)	
Civil War			-0.504*** (0.004)	-0.676** (0.028)
Oil			0.028 (0.552)	0.126* (0.061)
World region fixed effects	NO	YES	YES	NO
Constant	6.314*** (0.000)	5.210*** (0.000)	6.155*** (0.000)	6.153*** (0.000)
Observations	90	90	53	22
R-squared	0.294	0.408	0.663	0.593

Notes: robust p-values in parentheses

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

Oil defined as log of barrel per 1000 population (Source: "Oil reserves (per capita) by country", CIA World Factbook, 28 July 2005. Retrieved from

http://www.NationMaster.com/graph/ene_oil_res_percap-energy-oil-reserves-per-capita)

Civil War is 1 if a country experienced a violent inner-country conflict with 1000+ battle deaths during the last 50 years (Source: COW)

In Column 4, numeracy in 1900 is instrumented with books per capita in 1750 (Source: Baten and van Zanden 2008).

For books per capital, some country values are interpolated: for Southeast Asia the value of Indonesia is assumed,

for central-eastern and southeastern Europe Poland, for Eastern Europe Russia

Table 2: Regional regressions of ABCC increase 1890-1920

	(1)	(2)	(3)
Log absolute immigrant increase	1.15*** (0.006)	1.58*** (0.000)	0.98* (0.084)
Log population density	1.46*** (0.009)	1.19** (0.045)	-0.37 (0.867)
Initial ABCC	-0.52*** (0.000)	-0.63*** (0.000)	-0.59*** (0.009)
Average educational expenditure	4.96*** (0.009)		8.20** (0.046)
Commodity export: cattle			6.48 (0.151)
Commodity export: coffee			2.47 (0.660)
Commodity export: cotton			4.68 (0.174)
Commodity export: mate			3.99 (0.344)
Commodity export: rubber			-6.16 (0.425)
Commodity export: sugar			6.02 (0.367)
Commodity export: tobacco			0.00 (.)
Initial education expenditure in 1890		3.41* (0.063)	
Constant	16.80* (0.059)	29.22*** (0.006)	31.64* (0.090)
Observations	19	19	19
R-squared	0.85	0.83	0.89

Robust in parentheses

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

Table 3: Robustness-test – do the results hold also for the 1890 - 1950 ABCC increase?

	(1)	(2)	(3)
Log absolute immigrant increase	1.19*** (0.006)	1.65*** (0.000)	1.02* (0.084)
Log population density	1.52*** (0.009)	1.24** (0.045)	-0.38 (0.867)
Initial ABCC	-0.50*** (0.000)	-0.62*** (0.000)	-0.57** (0.013)
Average educational expenditure	5.17*** (0.009)		8.55** (0.046)
Commodity export: cattle			6.75 (0.151)
Commodity export: coffee			2.57 (0.660)
Commodity export: cotton			4.88 (0.174)
Commodity export: mate			4.15 (0.344)
Commodity export: rubber			-6.41 (0.425)
Commodity export: sugar			6.28 (0.367)
Commodity export: tobacco			0.00 (.)
Initial education expenditure		3.55* (0.063)	
Constant	17.50* (0.059)	30.43*** (0.006)	32.96* (0.090)
Observations	19	19	19
R-squared	0.84	0.82	0.89

Robust in parentheses

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

Internet Appendix: Methodology and basic concepts of age heaping

Numeracy is an important component of overall human capital. In order to provide estimates of its basic components, we apply the age-heaping methodology.²⁵ The idea underlying this methodology is that traditionally in underdeveloped countries when census takers, army recruitment officers, or prison officials asked an individual to state his or her exact age only a certain number were able to do so, the others rounding theirs off -- for example, to 40, when, in fact, the correct answer was 39 or 41. In today's world of mandatory schooling, passports, universities, birth documents, and bureaucracy, it is hard to imagine such a situation, but it prevailed before the modern era, in less information-oriented ones than ours. The typical result of age heaping is an age distribution with spikes at ages ending in a five or a zero and an underrepresentation of other ages, creating a distortion of the true age distribution. The fact that there was also some heaping on multiples of two, especially among children and teenagers and to a lesser extent young adults, indicates that most undereducated individuals knew their age as teenagers but had lost the ability to recall or calculate it by the time they reached adulthood.²⁶

The 1790 Mexico City census offers an example of rounding off to multiples of five, recording as it does 410 people aged 40, but only 42 aged 41: clearly an example of age heaping. Apolant (1975, 333) offers some particular examples: for instance, Joseph Milan, appearing in February 1747 as a witness in an Uruguayan court, was 48 years old according to one judicial record, but declared himself (in the same year) to be 45 in another. Such misreporting is a problem for demographers trying to calculate life expectancies and other population statistics, but, paradoxically, it is thanks to age heaping that we can extrapolate numeracy rates in various populations over the course of history.

The ratio between the preferred ages and the others can be calculated by means of any number of indices, the Whipple index being considered the most reliable.²⁷ The number of persons in a given census reporting a rounded age ending in 0 or 5 is divided by the total number of persons in the census, and this sum is then multiplied by 500. Thus the index measures the percentage of those who state an age ending in a five or zero, assuming that each terminal digit appears with the same frequency in the "true" age distribution.²⁸

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

For an easier interpretation, A'Hearn, Baten, and Crayen (2010) suggested another index, which we call the ABCC index.²⁹ It is a simple linear transformation of the Whipple index and yields an estimate of the share of individuals who correctly report their age

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

This share turns out to be closely correlated with other measures of human capital, such as literacy and schooling, at the local level, nationwide, and over various spans of time (Bachi 1951, Myers 1954, Mokyr 1983, A'Hearn, Baten, and Crayen 2010). A'Hearn, Baten, and Crayen (2010) found that after 1950 there was a close correlation in less developed countries (LDCs) between illiteracy and age heaping. They calculated the rates of age heaping and illiteracy among approximately 270,000 individuals representing 416 regions, ranging from Latin America to Oceania.³⁰ The correlation coefficient with illiteracy was high: 0.7. The correlation with the PISA results for mathematical abilities was even higher: 0.85; this means that the Whipple index is more closely correlated with mathematical abilities than with literacy.

In addition, A'Hearn, Baten, and Crayen (2010) also used a large US-census sample to perform a very detailed analysis of this relationship, subdividing the sample according to race, gender, and educational status, among other criteria. In each case, they obtained a statistically significant relationship. It is worth noting that the samples' coefficients are quite stable: that is, a unit change in age heaping is associated with similar changes in literacy across the various tests. The results are valid not only for the US but also for all of those countries in which age heaping has been found to occur, the correlation there, as in the US, being both statistically and economically significant.

In order to assess the robustness of those US-census results and the similar conclusions drawn from late-20th-century LDCs, A'Hearn, Baten, and Crayen (2010) also assessed age heaping and literacy in 16 European countries between the Middle Ages and the early 19th century. Again, they found a positive correlation between age heaping and literacy, although the relationship was somewhat weaker than for the 19th or 20th century data. It is likely that the unavoidable measurement error when using early modern data caused the lower statistical significance.

Age-heaping data have also been compared with other human capital indicators, such as primary-education rates. To date the broadest such geographical sample studied is one designed by Crayen and Baten (2009): age-heaping and schooling data (and other explanatory variables as well) from 70 countries. They found in a series of cross-sections between the 1880s and the 1940s that primary schooling and age heaping were closely correlated, with R-squares between 0.55 and 0.76 (including other control variables; see below). Again, the coefficients were quite stable over time.

This large sample also made it possible for us to examine other potential determinants of age heaping. To assess whether the extent of bureaucracy, birth registration, and government interaction with citizens are predictors of whether an individual knows his or her

own exact age, (i.e., is numerate) they used the number of censuses performed in a given country during the period under study as an explanatory variable for its age-heaping rate. Except for countries with a very long history of census-taking, all variations of this variable turned out insignificant, which would suggest that an independent bureaucracy effect was rather weak. In other words, it is sometimes the case that societies with a high number of censuses had high age awareness. But, at the same time, these societies were also early in introducing schooling and this variable clearly had more explanatory power in a joint regression than the independent bureaucracy effect. In addition, using height as well as per capita GDP as a proxy for welfare, Crayen and Baten also determined that the effect of the standard of living on age-heaping rates varies considerably, being significant in some decades, insignificant in others. Cultural determinants of age heaping were considerable in East Asia but not in the Latin American countries under study in this article.

In order to apply the ABCC age-heaping index to several countries and birth decades, we use the age groups 23-32, 33-42, etc.³¹ We omit the age range from 63 to 72 because it provides an insufficient number of observations, especially in the case of the 17th and 18th centuries, when mortality was higher than in subsequent centuries.³²

An advantage of the age-heaping methodology is that age statements are more widely available than other human capital proxies, such as signature ability or school attendance. As Reis (2008) argues, age heaping is a very basic measure of human capital. It is therefore especially useful in the study of human capital development in Latin America during the 17th and 18th centuries, when more advanced human capital indicators were quite scarce and instead of offering insights into a broad spectrum of a given population provided only a limited one, of the socioeconomic elite.

References (not in main list of references)

Apolant, J. A. (1975) *Génesis de la familia uruguaya*, 4 vols. Montevideo.

Bachi, Roberto (1951). "The Tendency to Round off Age Returns: Measurement and Correction."

Bulletin of the International Statistical Institute, 33: 195-221.

Mokyr, Joel (1983). *Why Ireland Starved*. London: George Allen & Unwin.

Myers, Robert J. (1954). "Accuracy of Age Reporting in the 1950 United States Census." *Journal of*

the American Statistical Association XLIX (1954): 826-31.

¹ The „*Fome Zero*“ program was introduced 2002, by *President Luis Inácio Lula da Silva*. It assists ca. 44 million Brazilians with monthly payments between 50 and 95 Brazilian *Reais* to amend their nutrition expenditures. In Brazil, nowadays the share of people living from less than 1\$ per day is well over 20 percent.

² For source countries, emigration of highly skilled individuals, the phenomenon of a brain drain, can be problematic, however the possibility of migration might also induce further investment in human capital and therefore affect factor endowment positively. Beine et al. (2008), Stark et al. (1997), and Bruecker and Defoort (2006) conclude that effects on the source country's human capital endowment are ambiguous.

How was the brain drain situation in the Brazilian case? Stolz and Baten (2011) found that typically most migrants had numeracy levels in the middle between European source country and Brazilian destination levels of natives. Hence there was probably not such a brain drain for the source countries.

³ The question is whether the immigrants are substitutes or complements to the native population, because this determines if there will be a crowding out of natives on the labor market, or migrants' skill composition is complement natives so that growth effects can be realized (Borjas 1994).

⁴ CGE models have the advantage that possible internal migration due to international immigration can be controlled for and multiplier effects can be captured.

⁵ Alternatively, we tried a set of world regional fixed effects, the results were similar.

⁶ However, Martín-Fríscher, Musacchio and Viarengo (2009) argue that the increase in human capital in the period 1890-1940 cannot be explained by European immigration, because these migrants were among the least educated in Europe. Still, they do not provide data on immigrant selectivity. In our data, selectivity shows a mixed picture. Whereas migrants from Latin America, southern European countries and Syria were mainly positively selected in terms of numeracy from the source country population, US-Americans and Western Europeans were less numerate than their compatriots at home. Moreover, the education of some of the immigrant nationalities might have even exceeded their home countries' average. Portuguese immigration, the second most important national group arriving in Brazil in the end of 19th century and the beginning of 20th century, is an example in line. Using passport registers generated in Portugal, some studies showed that emigrants were more qualified and better educated than the average in Portuguese population (Monteiro, 2000; Rodrigues, 2006; Andrade, 2009).

⁷ Other potential caveats are discussed in an internet app and in related literature, see Manzel et al. 2011, A`Hearn et al. 2010.

⁸ The census contains information at the household level. All household members were asked for their names and ages, and there is information if they actually belonged to the family or were “*agregados*”. This word refers to people who were like servants in pre-industrial rural Northern Europe. They were free, lived in the master's household, performed a wide range of services, and received board and lodging, and sometimes also some money. Furthermore, the census provides information on the possessions of a household.

⁹ This work was developed by the team coordinated by researchers Clotilde Andrade Paiva, Roberto Borges Martins and Maria do Carmo Salazar Martins, who kindly allowed us to use their database. Some results reached with the use of this material can be found in Paiva (1986, 1996), Paiva and Arnaut (1990), Rodarte (2008), Andrade (2001).

¹⁰ All of the published volumes of these censuses can be found in the digital library of the Brazilian bureau of statistics (<http://biblioteca.ibge.gov.br/>).

¹¹ Using data from Juif and Baten we find that numeracy was even lower in the northern parts of Brazil than in the southeast. However, this difference was statistically significant only without controlling for accusation and gender. With these controls included, southeastern numeracy was not significantly different from that in the north. A reason for this may be that many Jews, with higher levels of human capital, fled to the southern pioneer regions whilst women tended to aggregate in the capitals of the north. The total sample size of the Portuguese inquisition cases with age statements is N=10316, but the number of Brazilian individuals is only around five percent of that number (N=493). Nevertheless, the authors recently showed that even quite small sample sizes yield surprisingly accurate results (Juif and Baten 2011).

¹² We only use the age ranges of the 23-62 year olds, to avoid age effects caused by selective mortality of the older age groups.

We made one exception for the 1750s birth cohort, represented by the 63-72 year-olds for the 1830 nominative list in the Southeast, because we were curious whether the higher ABCC level of the 1830s compared to the 1772 was caused by selectivity, or by an increase between 1740s and 1760s. Given that the birth cohort of the 1750s falls almost exactly between the 1740s and the 1760s values, this might be an indication of an increase of ABCC during this period.

¹³ The geographical, gender and ethnic distribution of human capital was, however, extremely unequal in Brazil. Colonial legacy had produced a society with a very unequal distribution of income and human capital (Sokoloff

and Engerman 2000, Fritscher et al. 2009). Thus, one has to keep in mind that numeracy levels for slaves or women or people from rural areas are far lower than those of city dwellers and white or upper class men.

¹⁴ At the same time, the sugar-producing regions in the northeast (Bahia and Pernambuco, especially) had been suffering a steady decline in market share of the Atlantic sugar, mainly a result of competition from the West Indies sugar. The opening of the mines made the situation even more delicate to the sugar areas since they started to compete in the transatlantic slave trade with the mining proprietors. However, throughout the eighteenth century sugar was the main product of the Brazilian export basket.

¹⁵ To a review of the debate about this question, see Bergad 1999.

¹⁶ The Portuguese continued to predominate in the migratory flow to the second half of the 1870s, followed by the Germans. From 1876, there was a strong increase in Italian immigration that began to dispute the primacy with the Portuguese. From 1882 the Italians have dominated the main migratory flow, which will expand in subsequent years until 1903. Thereafter, the Portuguese again predominate. The Spaniards formed the third largest group, with significant participation especially in the years 1889 to 1897 and from 1904 to 1917. The Germans also formed a significant group, especially in the 1880s, and from 1890 to 1908. It should be noted as the beginning of Japanese immigration from 1908, but which has taken over significant amounts only after the First World War.

¹⁷ For a panoramic and more recent view of this discussion in Brazil, see Fausto (1999), Bassanezi et al. (2008).

¹⁸ In a regression of GDP 2000 on physical and human capital in the year 1900, which is possible only for a number of 23 countries, both determinants are, of course, highly significant. We tested this with the human capital indicator incorporated in the Baier et al. (2006) data set and also with our numeracy estimates.

¹⁹ Geographical data is taken from www.cepiis.fr, political and institutional variables are from the PolityIV project.

²⁰ In the version with the world regional fixed effects, the only world region which displays a coefficient statistically different from zero is Sub-Saharan Africa. This result goes in line with previous studies from acknowledged scholars on economic growth which all detected a somewhat peculiar negative growth pattern for African countries (i.e. Sachs & Warner 1997, Englebert 2000).

²¹ <http://www.brasilien.li/artikel/wechselkurs-real-euro-1980>, last accessed March 24, 2011.

²² The ABCC levels in 1890 are smallest for north-eastern Brazilian states like *Alagoas* or *Sergipe*. *Rio Grande do Sul* and *Santa Catarina* display very high numeracy values in 1890. *Rio de Janeiro* and *São Paulo* are

situated somewhere in the middle. In 1920, however, these primary immigrant destinations display the highest numeracy values, together with the southern states of *Paraná* and *Rio Grande do Sul*

²³ If we include the set of export dummy variables, the beta coefficient of average education expenditure becomes larger than the one on immigrants.

²⁴ We could also have imagined that were changes in the number of days worked, but the changes were actually quite modest. Until the political independence of Brazil in 1822, the holidays were linked to the Catholic worship. Sundays were kept as holy days, when Catholic Church recommended not working. However, slaves were allowed by their masters to take advantage of this day to work for themselves, cultivating pieces of land, fishing, hunting or selling their products. Religious people had always condemned the practice, asking the slave owners to appoint other days for the slaves were dedicated to their own work. However, the reiteration of these claims indicates that the slaves who could took advantage of Sunday to work on their own. [About these practices see, among others, Schwartz, 1992, Slenes, 1999.]

In addition to Sundays, Brazil had an extensive religious calendar that was observed as days of rest. There were about 10 religious days observed by all (Holy Week, Corpus Christi, St. Anthony, St. John, Assumption, All Souls, All Saints, Immaculate Conception, Christmas), and other holy days that could be observed only locally, as the day of the patron saints of the parish or other saints worshiped there. In addition, there were the occasional celebrations related to the royal family, the arrival of governors and bishops, and the festivities of "particular character" as baptisms, birthdays and blessings of mills (Breu, 1976). Therefore, it is plausible to consider that there were at least 20 days per year considered as holidays in colonial period.

After the political independence of Brazil in 1822, new dates have been introduced as civic holidays. During the First Empire (1822-1831), for example, there was the commemoration of March 25 (oath of the Constitution), April 7 (abdication of Pedro I), September 7 (Independence Day), December 2 (birth of Pedro II) and others (Basile, 2006). Later, other days were added to the civic calendar. While those days were mainly observed in urban areas, they added new holidays for a significant portion of Brazil, making the total number of days without work arisen to about thirty days a year.

²⁵ For more detailed surveys on the age-heaping methodology see A'Hearn, Baten, and Crayen (2010).

²⁶ At more advanced ages this heaping pattern tends to be negligible, but is, interestingly, somewhat more prevalent among populations sufficiently numerate not to round off ages to multiples of five.

²⁷ A'Hearn, Baten, and Crayen (2010) found that this is the only index that fulfils the desired properties of scale independence (a linear response to the degree of heaping) and that it offers a reliable ranking of samples among which the degree of heaping varies.

²⁸ A value of 500 means an age distribution with ages ending only in multiples of five, whereas 100 indicates no such heaping patterns on multiples of five: in other words, 20% of the population reported an age ending in a multiple of five.

²⁹ The name results from the initials of the authors' last names plus Greg Clark's, who suggested this in a comment on their paper. Whipple indexes below 100 are normally caused by random variation of birth rates in the 20th century rich countries. They are not carrying important information, hence normally set to 100 in the ABCC index.

³⁰ See A'Hearn, Baten, and Crayen (2010). Appendix available from the authors.

³¹ This method has two advantages: it spreads the preferred ages, such as 25 or 30, more evenly within the age groups; and it adjusts for the fact that more persons will be alive at age 50 than at age 54 and at age 55 than at age 59 (Crayen and Baten 2010).

³² Given that adults aged 23 to 32 round on multiples of two as well as five, we use the adjustment method suggested by Crayen and Baten (2010) to increase the Whipple value (minus 100) by 24% before calculating the ABCC value.