Kerstin Manzel,* Joerg Baten**, and Yvonne Stolz***

CONVERGENCE AND DIVERGENCE OF NUMERACY: THE DEVELOPMENT OF AGE HEAPING IN LATIN AMERICA, 17TH TO 20TH CENTURIES[1]

* University of Tuebingen, Mohlstr. 36, 72074 Tuebingen, kerstin.manzel@uni-tuebingen.de,

** University of Tuebingen and CESifo, (corresponding author) joerg.baten@uni-tuebingen.de

*** University of Tuebingen,

---

[1] Acknowledgments:

Abstract

This study makes the first systematic attempt to trace the long-term development of Latin American numeracy: a phenomenon of great interest to economic historians in that it serves as an accurate gauge of human capital development. In order to approximate basic numeracy we use age-heaping techniques. We find that Latin America was on a path of convergence with Western Europe during the early 18th century. During the early 19th century, not only did numeracy development stagnate in some Latin American countries but differences among some of them actually increased. While numeracy rates in Argentina, Uruguay, and to a lesser extent Brazil, underwent, along with Europe, a significant increase in the late 19th century, they declined in Mexico, Ecuador, and Colombia. By performing a regression analysis, we find, even when we control for investment in education, that mass immigration contributed to human capital formation.

Introduction

When Pedro de Valdivia, a Spanish conqueror and the founder of Santiago de Chile, came to the New World in 1536, 89 of the 150 Spaniards who accompanied him could not sign their names, and only one of them had received any formal education.[2] The educational level of the indigenous population of Latin America was at least as low as that of these Spaniards.The Aztecs and Mayas had some schools, but attendance was restricted to the sons of their leaders, or *caciques.* How did education develop during and after the colonial era, once Latin America had gained independence from Spain? This study will trace long-term trends of human capital from the early modern period to the 20th century and address the issue of the timing of improvement of numeracy in the New World. Was there already an international convergence in numeracy in the 18th century? Was the mass immigration of Europeans in the late 19th century a driving force in improving human capital in Latin America?

Most studies of the development of human capital restrict their view to a single country or a short time span for two reasons: information on human capital measures tends to be scarce (this is the case for most Latin American countries prior to the second half of the 20th century) and those measures that are far apart in time or space do not permit comparison. We begin our comparative analysis of human capital development in the early 20th century and then move back in time to the 18th, and in some cases the 17th, century. One important component of human capital is numeracy, i.e., the "ability to count, keep records of these counts, and make rational calculations".[3] We employ the age-heaping technique, which permits one to calculate the proportion of a given population able to provide to census takers their exact age, rather then providing a rounded age**.** This indicator of basic numeracy is a precondition for developing more advanced skills, including literacy. The age-heaping approach thus captures a basic form of human capital that is useful in this context, primarily because it is not linked to an individual's mother tongue. Moreover, this proxy is

---

[2] Austin, *Education*, p. 1.
[3] Emigh, 'Numeracy', p. 653.

constructed coherently over time and space[4], thereby permitting the comparison of numeracy values from various sources and countries. Having been introduced into modern economic historiography by Mokyr[5], the concept of age heaping has recently led researchers to cultivate a new field of research.[6] The literature shows that numeracy is closely correlated with other human capital indicators, such as literacy and schooling. (A summary of the methodology can be found in Appendix C)

Latin America offers interesting sources that permit historians to trace numeracy over the long term. The colonial powers carried out population counts and regularly collected data on their colonies' populations. A great number of these sources have survived and can be used to estimate numeracy trends, despite biases, which need to be factored into any conclusions. We pose three fundamental questions: How did numeracy develop in Latin America from the 17th century on? Which Latin American countries led the field, at any given time, when it came to education? At what point in time may we speak, in reference to numeracy, of a convergence between Latin America and Western Europe?

Our findings reveal that during the 18th century many Latin American countries made such rapid numeracy progress that their levels converged with those of Western European countries, but that by the early 19th century this progress had begun to level off. Considerable differences emerged in the early 20th century, with the Southern Cone countries (chiefly Argentina, Chile, and Uruguay) at the upper end of the scale and Colombia, Ecuador and Mexico at the other. By performing a regression analysis, we find, even when we control for investment in education, that mass immigration contributed to human capital formation.

The remainder of the paper is structured as follows. Section 2 reviews the literature on education in Latin America. Section 3 presents the data sources we used to construct a new database

---

[4] Crayen and Baten, 'Numeracy', p. 89.
[5] Mokyr, *History*, p. 245.
[6] A'Hearn, Baten, and Crayen, 'Quantifying quantitative literacy; Baten, Crayen, and Manzel, 'Zahlendisziplin'; de Moor and van Zanden, 'Leeftijdstapelen'; Clark, *Alms*; Crayen and Baten 'Numeracy'; Manzel and Baten, 'Gender equality'; Cinnirella, 'British nutritional status'; Mironov, 'Novaya istoricheskaya demografia'; O'Grada, 'Jewish demography'.

of numeracy development in Latin America and discusses the representativeness of our samples. Section 4 shows the estimates for Argentina, Brazil, Colombia, Ecuador, Mexico, Peru, and Uruguay. Section 5 compares the estimated numeracy trends of these Latin American countries with those of the U.S. and Europe. Finally, Section 6 draws conclusions.

2. Literature review: Latin American human capital development in the very long run

Long before the arrival of the Spanish conquistadores three major cultures had devised systems by which to perform complex numerical calculations. The Mayas used a vigesimal -- that is, based on 20 -- number system, represented by bars and dots. Their time intervals were the *tun* (360 days), the *winal* (20 days) and the *k'in* (one day).[7] The Aztecs combined simple numbers to signify larger ones; for instance, 399 was represented by (15+4) x 20+15+4.[8] The Incas devised *quipo,* a calculation technique, based on knotted strings that permitted them to develop a sophisticated administrative system including population counts.[9]

During the colonial period, schooling was seen as a method to "civilise" the native elites, undermine indigenous customs, and spread the Catholic religion. Education was in the hands of various religious orders that had arrived in Latin America during the 16th century: primarily the Jesuits, but also some Franciscans and Dominicans.[10] The Jesuits established missions, primarily among the indigenous populations, and often learned their languages in order to facilitate communication and thus conversion. When, in 1767, the Spanish King, Charles III, expelled the Jesuits from the entire continent, they left tens of thousands of indigenous people unprotected in their missions, many of which decayed, but some of which were taken over by Franciscans or the Dominicans.[11] Schools were scarce in colonial Latin America, especially before the 19th century. Moreover, attendance was restricted to the sons of the European elite and of the *caciques,* and

---

[7] Closs, 'History of mathematics', p. 143.
[8] Conant, *Numeracy*.
[9] Julien, *Inka*.
[10] Leininger Pycior, *Education*; The Jesuits, for example, arrived in Peru and Mexico in 1568 and 1572 (Merino and Newsom, 'Jesuit missions') and the Franciscans in Mexico in 1524 (Vaughan, 'Mexican education').
[11] Merino and Newsom, 'Jesuit missions'.

classes were conducted almost exclusively in Spanish.[12] Apart from what scattered Jesuit missions provided before they were expulsed, there was virtually no formal education in Brazil in the 18th century.[13] What is more, the few schools that survived thereafter did not create much in the way of human capital, as the teachers were underpaid and the schools understaffed.[14]

Regional disparities tended to be considerable. For example, in post-independence Mexico the funding and administration of schools was often handed over to local municipalities, a policy that worsened regional disparities, with prosperous cities benefiting at the expense of rural areas.[15] A similar situation prevailed in Brazil, where the educational system was decentralised shortly after independence, resulting in a decline in the quality of elementary education in some regions.[16] Compulsory primary education was theoretically introduced over the late 19th and early 20th centuries throughout Latin America, although not reinforced consistently. Generally, schools were poorly endowed.[17] One reason often cited to explain this decline in the quality of the educational system is the reluctance of the ruling elite to finance public schools; keeping the labourers and peasants uneducated reinforced their hold on power.[18] Lindert has called primary public education "the kind of education that involves the greatest shift of resources from upper income groups to the poor."[19] He discusses a number of positive and negative factors that influenced the decision to introduce large-scale, tax-financed, universal primary schooling. Rural elites in countries like Argentina, Chile and Venezuela who benefited from independence objected modernization endeavours including the provision of rural schools for fear of losing their bases of power.[20] The point of view of a member of the landed elite was, 'Why should I pay taxes to provide for public schooling when it will only spread discontent among the poor, including the day labourers who work on my own estate and incite them to rebel?'.

---

[12] Bakewell, *History*, p. 90.
[13] Baer, *Brazilian economy*.
[14] Ibid.
[15] Vaughan, 'Mexican education'.
[16] Martìnez-Fritscher, Musacchio and Viarengo, 'Brazilian education'.
[17] Meyer Loy, 'Colombian education'.
[18] Mariscal and Sokoloff, 'Comparative education'.
[19] Lindert, *Social spending*.
[20] Morse, 'Urbanization'.

Statistics on the development of education in the New World are scanty. Oxford University's Latin American economic-history database is limited to the second half of the 20th century. Earlier literacy estimates exist for selected countries[21], even though Javier Núñez criticises the estimates for the period around 1900, arguing that there was a lack of comparability between different definitions of literacy due to the fact that censuses were taken in different years and were confined to particular sectors of the population.[22] Moreover, in some countries "literacy" was defined as "the ability to read," in others as "the ability to read and write." Núñez's solution is to combine population censuses with marriage registrations (since they require signatures) and with crime records (since they indicate whether or not the criminal was literate). This methodology allows him to re-estimate literacy rates for the turn of the 20th century for a number of Latin American countries. He finds that literacy rates in Brazil in the first half of the 20th century were low, and concludes that this was due to the fact that Brazil abolished compulsory primary education in 1891 and reintroduced it only in 1934.

Benavot and Riddle[23] performed a wide-ranging assessment of late-19th- and early-20th-century school-enrollment rates. Their data show a moderate overall increase in Latin American primary-school enrollment rates, more pronounced in the urbanised countries such as Argentina, Uruguay, and Chile, where there was a relatively large proportion of European immigrants, and in the British colonies of the Caribbean as well. However, one has to bear in mind that school-enrolment rates measure input, not output, and that even today the quality of education in some Latin American countries is poor, as measured by high repetition rates and low share of pupils finishing school, in spite of rising enrolment rates.[24]

During the 20th century, major differences among Latin American countries persisted.

---

[21] Newland, 'Elementary education'; Engerman, Haber, and Sokoloff, *New world economies*; Astorga, Bérges, and FitzGerald, 'Living standard'.
[22] Núñez, 'Literacy evolution'.
[23] Benavot and Riddle, 'Education expansion'.
[24] For Mexico, see Palafox, Prawda, and Velez 1994; for Latin America and the Caribbean, see: Wolff, Schiefelbein, and Valenzuela, 'Primary education'.

Astorga, Bérges, and FitzGerald[25] estimate literacy rates among those above the age of 15 in the LA6 countries (i.e., Argentina, Brazil, Chile, Colombia, Mexico, and Venezuela) at 33% in 1900, 60% in 1950, and 89% in 2000. For the remaining Latin American countries literacy rates were 32% in 1920, 46% in 1950, and 82% in 2000.[26] Moreover, most of the improvement in literacy took place from 1900 to 1939 for the LA6, but among the others did not occur until 1940-1980.

Núñez estimated that educational development in 19th-century Latin America trailed that of Italy and Spain by three or four decades, not on account of a lack of financial resources and resistance on the part of the elite but rather on account of the Independence Wars.[27] During the early 19th century, the newly independent countries suffered from political instability along with a dropoff in international trade, and capital flight.[28]

European immigration was a decisive factor in the development of education in Latin America. Brazil subsidised immigration extensively in the late 19th century, a policy that Baer deems an effective alternative to investment in education.[29] Luis Bertola and José Antonio Ocampo[30] suggest a typology of Latin American countries: those in which indigenous and mestizo people compose a large proportion of the population, those with a relatively high proportion of African descendents and those with a relatively large proportion of European immigrants, such as Argentina, Chile, Uruguay, and Cuba. Such immigrants not only had a positive impact on the promotion of education[31] but also constituted a source of teachers.[32] In Brazil, which had a far lower immigration rate than did Argentina, human capital development in those states where the rate was relatively high was significantly greater than where it was not.[33]

In sum, several (mostly qualitative) studies on Latin American educational development exist, but estimates of human capital measures taken prior to the late 19th century, which would

---

[25] Astorga, Bérges, and FitzGerald, 'Living standard'.
[26] Cuba is excluded.
[27] Núñez and Tortella, *La maldición divina*.
[28] Haber and Klein, *Brazilian independence*.
[29] Baer, *Brazilian economy*.
[30] Bertola and Ocampo, *Desarollo,* p.21.
[31] This despite the fact they were not as well educated as immigrants to the United States, and in many cases had emigrated because they had not encountered success in their native countries.
[32] Thorp, *Economic history*, p.37.
[33] Stolz, Baten, and Botelho, 'Mass migration'.

permit long-range comparisons, are lacking. For instance, the study by Manzel and Baten of the differential between male and female age heaping among Latin American and Caribbean countries begins only with the post-1880s period.[34] This study fills the gap.

3. Data Sources

Beginning with that of the ambitious *Visita General* of the Viceroy Francisco de Toledo, in 1572, population counts in the New World were carried out regularly. The first were intended to provide a detailed overview of the territory and its inhabitants and did not contain systematic age statements. Most *visitas* focused on small regional units and were repeated every five to ten years. These were followed, in the next century, by the *empadronamientos de tributarios*, population counts to determine how much individuals owed in taxes. *Padrones de población* (population counts with a limited geographical or social scope) and partial censuses carried out during the 18th and 19th centuries covered larger regions and a larger portion of the population.[35] In 1776, the Council of the Indies started a series of systematic census records.[36] At the same time, Portugal, eager to learn more about the inhabitants of her colonies, carried out systematic enumerations.[37] By means of this information the colonial powers could better calculate their tax rates and learned about the total potential of their territorial possessions. For the post-colonial period, information on the first few decades is scarce; censuses of the republics are available, but most of them date back only to the late 19th and 20th centuries.

In fact, Latin America is the only continent in the developing world for which such early population counts are widely available.[38] These records are particularly valuable in that they provide a considerable quantity of detailed information; records for the birth cohorts of the late 17th and 18th centuries include a given individual's age, gender, birthplace (in most cases), ethnicity or

---

[34] Manzel and Baten 'Gender and numeracy'.
[35] Mellafe, *Social history*, pp. 148-170.
[36] Platt, *Census records*, p. 8.
[37] The only groups excluded from the Brazilian counts were regular troops, ecclesiastics, and those Indians who, having successfully resisted attempts at cultural assimilation, were excluded from these censuses (Alden, 'Brazilian population').
[38] Platt, *Census records*, p. 7.

caste, marital status, occupation, and in some cases an itemised list of the family property. In many cases only aggregated data were available for the 19th and early 20th century.[39]

Ethnic background played a crucial role in the allocation of rights and duties in colonial society. In New Spain, for example, Spaniards, Mestizos, and Indios paid tributes, each at a different rate, to their *encomenderos* (from the verb "to entrust"), landowners who functioned as trustees, and whose tax obligations were, in turn, partly calculated according to the tributes paid by the indigenous people who lived and worked on their estates.[40] The data currently available permit an in-depth study of age heaping from the colonial era to the middle of the 20th century. Our evidence, spanning Argentina, Brazil, Colombia, Ecuador, Mexico, Peru, and Uruguay,[41] represents a large part of the Latin American continent, around 85% of today's population.[42]

An important question is whether our various sources are representative of a given society during the period under study. Population censuses were in theory universal, representing all social strata in the area under consideration. However, the native population viewed officials of the Spanish Crown as invaders, and the purposes of such a census (chiefly tax collection and military recruitment) led to fear and distrust on the part of the populace. Many demographers are therefore convinced that most of the census data are flawed by undercounting. Gootenberg[43] goes so far as to assume that all of Peru's population estimates are inaccurate. Some men aged 16 to 36 attempted to avoid the military draft.[44] Women declared themselves to be widows or spinsters in order to protect

---

[39] For the 19th century, the information is even more abundant; however, the period of the wars of independence features, not surprisingly, several gaps.

[40] McAlister, 'Social structure'; Mellafe, *Social history*; throughout the 18th century, however, the New Spain's socio-ethnic stratification was modified by the relaxation of certain traditional ethnic barriers (McAlister, 'Social structure'). For example, the newly created militia regiments provided opportunities for professional, and therefore material and social, advancement not only to whites but also, for the first time, to Creoles and free coloured people. In 18th-century Mexico, however, Indians, continued to be banned from military service, and mestizos were discouraged from enlisting (Vinson, 'Mexican militia').

[41] While national borders changed during the colonial and post-colonial periods, we refer to national borders as they exist today.

[42] Because Central America was sparsely populated and had no mineral resources, the Spanish Crown deemed it of little interest (Mabry, *Colonial Latin America,* 58), and it was therefore rarely the object of a census.

[43] Gootenberg, 'Population'.

[44] Orozco y Berra, *Historia Mexico*; Alden, 'Brazilian population'; generally, census takers in English, Spanish, and Portuguese colonies were confronted with such problems as the dispersion of the population over a huge area and passive resistance" (Alden 'Brazilian population', p. 181).

their husbands and sons.[45] In the *Censo de Revillagigedo* (1790-1794) in Mexico City, for example, widows are over-represented. While this and other draft-dodging techniques led to an undercount of the overall population, they do not bias our numeracy estimates as long as the data concerned are not strongly correlated with the men's educational status. The available evidence suggests that military-draft avoidance was widespread.[46] Did the opportunity costs of more educated persons lead to a downward bias in numeracy in the colonial period? Given the higher opportunity costs of richer and more educated persons, we would expect that our samples would be negatively biased where army service was perceived as an activity which should be avoided particularly by the upper social strata. On the other hand, human history is full of examples in which belonging to warrior castes or other types of military activity was not avoided in favour of economic and trading activity, especially in societies in which returns to land were partly appropriated in a quasi-feudal way by those who were powerful. Rich Latin American landowner's sons would enter the militia or the army normally as officers, because they would otherwise lose contact with the center of power which granted their landownership.

Given the fact that its financial resources were limited and that it had reason to doubt the fidelity of Latin American military forces, the Spanish colonial government's attitude toward them was cautious.[47] Until the 1760s, it often opted for small, part-time militias; but toward the end of the century it reversed its policy, and armies began to expand. One would expect this trend would trigger an increase in avoidance behaviour on the part of the educated classes, which in turn would translate into a downward bias in our numeracy estimates. Can we observe this in the evidence? When we compare the 1744 military and the 1771 non-military Buenos Aires samples (Figure 2), we find no significant difference in the overlapping birth decade, that of 1710-1719. If anything, the early, military motivated sample had a slightly higher numeracy in this birth decade, but the difference is very small. Similarly, the numeracy trend in the 1818 military sample of Buenos Aires is situated at a level that would have been reached had the trend of the earlier samples continued in

---

[45] Arrom, *Mexican women*.

[46] With the possible exception of the very rich; Orozco y Berra, *Historia Mexico*, p. 72.

[47] We thank an anonymous referee for this hint.

a linear way. To conclude, comparing the different samples of the Argentinean case does not support a strong downward bias for militarily motivated samples, not even during the late 18th century, when army service might have been extended over previous militia service. Evidently the sons of landowners and wealthy farmers, who constituted the majority of Argentina's socioeconomic elite, did not avoid the military draft in significant numbers.

Similarly, censuses for the purpose of tax collection prompted only moderate avoidance behaviour: that is, the wealthy tended to conceal a portion of their wealth, instead of trying to avoid being registered at all. For example, it is hard to imagine that the authorities allowed Buenos Aires's upper class to avoid being listed among the city's inhabitants, whereas it is easy to imagine that by means of a bribe a rich man could get his tax base modified in his favour. Such avoidance behaviour would therefore not lower the numeracy rate, since it did not cause a distortion in the population base of the tax-census data, and therefore does not falsify our numeracy estimates. Alden describes similar avoidance behaviour in Mexico, Brazil and elsewhere in Latin America.[48]

A further strategy to assure the comparability of our sources is to run a panel regression of our numeracy estimates on a dummy variable that controls for the specific purpose of certain censuses: specifically, prison and military ones, with others, taken for other purposes, used as a reference category (Table 3). The 1700 Peru census, the Argentina ones of 1744 and 1818, and the 1846 and 1868 Uruguay ones were undertaken for military purposes. The 1871 Peru sample is taken from prison records. We control for country- and time-fixed effects. We find that neither the institutional context of prison nor military motivation was significantly different from the general-purpose census that served as reference. Consequently, even data flawed by an undercount can provide reliable information on the development of numeracy, as an element of education, in the colonial societies of Latin America. Some bias is unavoidable; we minimize its impact by comparing samples whenever possible.

An important issue is whether each member of a given household in the age range under

---

[48] Alden, 'Brazilian population'.

study (23-62) reported his own age or whether it was the household head who reported the ages of all the household members. In the latter case, since the household head may not have been certain of their exact ages, age heaping, except in the case of the heads themselves, may have occurred. When we compared male household heads and non-heads registered in the 1744 Buenos Aires census, we found, in fact, no substantial differences, even though one might expect that the educational status of the heads would be higher.[49] We conclude that the possibility that a household head was the source of data on other household members does not pose a serious problem. In addition, census takers sometimes noted that a certain person claimed to be 30 years of age but "but looked considerably older".[50] Such notes are strong evidence that census takers collected their data directly from each individual concerned, instead of accepting second-hand information provided by the household head, and that they resisted the temptation to adjust obviously erroneous statements of age.

Table 1 and Appendix A contain additional details on the sources that we used to construct a numeracy series for colonial and post-colonial Latin America. Some of our sources were restricted to a capital city, others to a region. The most challenging data sets derive from prison records; the concentration of prisoners on 23 and 24 years of age results in a skewed age distribution, and therefore in some cases in an upward bias for the youngest prisoners in the ABCC index. (For a discussion of the ABCC index, see Internet Appendix C.)

In order to assess the representativeness of our samples in detail, we compare the ethnic composition of our census samples with that of the entire population for all cases in which the literature provided the necessary information to do so (Table 2). It is noteworthy that classification by skin colour was abolished in most countries after independence; our comparisons according to skin colour are therefore limited to the colonial-era samples. The 1744 and 1771 Argentine samples provide an accurate representation of the population of Buenos Aires with the exception of its

---

[49] The difference in the ABCC index, not counting slaves, is only 48 vs. 46. We compared age heaping for the entire age range, 23-62, only, because the data set was too small to permit us to compare them by decade. Fortunately, household heads were not as a rule older than other household members, so age plays no role here.

[50] Cook, *Numeración*, p. 34.

blacks and mulattos, slightly underrepresented in the first of the two.[51] As for the Mexico census samples, those of both Mexico City and Oaxaca in 1777 underrepresent the indigenous population, whereas that of Durango errs on the other side. The 1700 Lima sample overrepresents Spaniards, while the proportion of Spaniards in the prison sample for 19th-century Lima is smaller than that of the Lima population as a whole.[52] While we have no information on Brazil's ethnic composition prior to the 1830s, according to the 1830 census slaves constituted 36% of the population in the districts in Minas Gerais, São Paulo, and Paraná, a figure that seems reasonable in light of the fact that slaves constituted 38% in the 1838 Rio de Janeiro census.[53] Thus it is safe to conclude that, while not all census samples reflect the exact share of each ethnic group in colonial society, our data are free of any serious bias.

Another important issue is whether migrants should be included in the individual samples. Since the focus of this study is the long-term formation of human capital in Latin America during the colonial and postcolonial periods, when immigrants were an important part of the population, they are indeed included. This decision enables us to compare individual and aggregated samples, since immigrants were always included in the latter. However, some shifts in the numeracy graphs during periods of mass immigration are to be expected, since the numeracy rate was generally higher among immigrants to Latin America than it was among the native population.[54]

Can we also compare age heaping and literacy in Latin America to double-check the methodology for the sources just described? To show explicitly that age heaping is a good indicator of educational status, we compare age-heaping and literacy data from samples representative of the 1869 and 1895 Argentinean censuses. These samples contain information on 38,776 and 43,897 inhabitants, respectively, from 23 to 62 years of age. In 1869, roughly 78% of the population were

---

[51] During the 19th century, the percentage of blacks and mulattos in the Buenos Aires censuses plummeted, from approximately 25% in those of 1810, 1822, and 1838 to 1.8% in that of 1887 (Andrews, *Afro-Argentines*).

[52] Fisher's data (Fisher, *Peru*) refer to the late 18th century. Afterwards, Lima's population evolved, the proportion of mestizos to whites and Indians gradually increasing.

[53] Graden, 'Slave trade'; the proportion in the northeastern states, however, tended to be higher than elsewhere in Brazil; estimates for Salvador are around 42% (ibid.). The 1838 census of Rio de Janeiro only distinguishes between slaves and free people. The figure refers to slaves, one has to bear in mind that not blacks were slaves and that there was an important share of mulatoes.

[54] Stolz, Baten, and Botelho, 'Mass migration'.

unable to read and write, while 21% declared themselves to be literate (1% gave no answer). By 1895 the literate share had had soared to 49%.[55] Figure 1 shows that age-heaping was far more extreme among illiterates than among literates; the spikes of their rounded ages are much higher, indicating that they were less willing or able to report their exact age. Literates in the 1840s birth decade had a ABCC rate of exact-age reporting of 81.5%, illiterates an index of only 60.5%.[56] The correlation on a provincial level between the share of illiterates and the Whipple index yields a highly significant coefficient, of 0.87. We can therefore safely conclude that in this context age heaping is a valuable, informative human capital indicator.

4. Age-heaping trends from colonial times to the 20th century

In order to assess the long-term development of numeracy in Latin America, we begin with a review of Argentina, for which we possess the most comprehensive data. Afterwards, we will proceed to discussing the other country cases, for which the data set is still substantial, but has some gaps and selectivities.

Argentina

For Argentina, we provide the first long-term estimate of numeracy values, beginning in the 17th century. The values for Buenos Aires in the late 17th and early 18th centuries indicate a very low numeracy rate (Figure 2). Initially, less than 40% of the population of Buenos Aires reported an exact age. The 1744 and 1771 censuses of the capital are mutually coherent, the value of the youngest cohort of the first similar to that of the oldest cohort of the second. Overall, there is a remarkable, 30%, increase in numeracy between the earliest cohort and the mid-18th century. A sample from the 1818 Buenos Aires census covers the birth decades of the late 18th century. Although its main aim was to assess military capacity, it included females and those who were not household heads. The fact that the line is flat suggests stagnation during the late 18th century, but at

---

[55] Only native-born Argentines are considered.
[56] These figures are comparable to those of Newland (Newland, 'Elementary education'), who calculates that in 1900 52% of Argentines over the age of 10 were literate.

a high level, of around 75-80%. For the Buenos Aires birth cohort of the 1810s the numeracy levels are about 70%, far below the one for the late 18th century and almost as low as the one for the 1750s. The birth cohorts of the period of the wars of independence (roughly 1808-1829) indicate an increase, but from a very low level, below that of the mid 19th century. We would characterise the increase between the lows of the 1810s through the 1840s as recovery because only the late 18th century levels were reached again.

For the 1810s, we also have the first nationwide estimate for Argentina: lower than in Buenos Aires (where in 1810 it was approximately 9%), suggesting that those who lived in the capital had greater access to the educational system and benefited from its economy, which offered more opportunities for skilled workers than were available elsewhere. After the 1840s, a continuous increase in numeracy took place, both in the capital and the rest of the country. The upward trend continued steadily until 1880, when, according to 1947 census data, age heaping ended. However, some age heaping occurs in the overlapping birth decade of the 1880s, as one can observe by examining the data for the youngest age group in the 1914 census. This gap of numeracy for the same birth cohort between two census years is typical for periods of mass immigration of more skilled migrants, relative to the native population. Although immigration rates declined after 1914, the influx remained considerable; the proportion of numerates -- that is, of those who declared their exact age -- in the census of 1947 would have been smaller than it was if it had not been for this selective immigration.

In order to estimate Argentina's national numeracy trend, we carried out an OLS regression, controlling for the share of males in each census (between 0.47 and 0.77) and controlling for a census of the capital with a dummy variable (Table 4). The coefficient of the capital and the male share is statistically significant and positive. The resulting time dummies, based on an assumption that males constitute 50% of the population, indicate that there was a positive development that persisted over time, with a steep upward trend from 1680 to 1760 (Figure 3). However, the regression results also indicate that there may have been a temporary stagnation, or even a

temporary deterioration, in the numeracy rate from the late 18th to the mid 19th centuries, even after controlling for gender and for the greater access to schools in the capital. During the second half of the 19th century, however, numeracy development in Argentina was clearly positive, thanks in large part not only to the influx of substantial numbers of numerate Western European immigrants but also to the relatively early introduction, in 1884, of compulsory primary schooling.[57] Although the educational level of these immigrants was generally lower than that of those who headed to the United States, it was superior to that of the native Argentines, and thus fuelled the rapid growth in human capital there.[58]

Brazil

Our earliest evidence regarding numeracy in Brazil dates from the birth decade of 1700 (Figure 4), our most recent was a census taken in 1950. The sample for the 1700-1740 birth cohorts, from a census of the São Paulo district of Sorocaba, does not contain data on ethnicity; that for the 1770-1800 birth cohorts, derived from 1830 census data collected in the states of São Paulo, Paraná, and Minas Gerais, contains data on men and women from both the slave and the free populations. In 1830 slaves constituted over 36% of Brazil's south-eastern population: a far larger percentage than that found in Argentina and Uruguay at the time. A'Hearn, Baten and Crayen showed that during early educational development, ABCC values of numeracy begin to increase first, before literacy values move upward.[59] In other words, ABCC values of around 50 percent often correspond with literacy values below 10 percent. Knowing one's exact age is a much more basic skill than reading and writing. This is also true for the slave population, which was largely illiterate. A'Hearn et al. studied U.S. African Americans who were born in slavery.[60] With data taken from the 1870 US census the authors show that numeracy values of native blacks is around 70%, while literacy only reaches a value of 35. If we apply this difference to the 1830 Brazil evidence, where some groups of

---

[57] Cortés Conde, *El progreso*.
[58] Núñez and Tortella, *La maldición divina*, p. 371.
[59] A'Hearn, Baten, and Crayen, 'Literacy implications'.
[60] Ibid.

slaves exhibit an ABCC of only 45, their literacy must have been below 10 percent, hence almost illiterate.[61] In all three states, slaves show significantly less numerical abilities than the free population.

How do our two early samples compare? Together they describe an overall upward development in numeracy throughout the 18th century. What exactly happened during the 30-year period between our two samples remains unclear, but their levels suggest that numeracy first increased and then stagnated or even decreased during the Napoleonic wars, which prompted the Portuguese court to flee from Lisbon to Rio de Janeiro, where it remained in power for over a decade. Brazil achieved independence quite peacefully, at least by Latin America's standards.[62] The 1890 census sample, which can be regarded as representative of the Brazilian population, indicates that numeracy levels rose modestly through the birth decade of the 1830s to that of the 1860s. The 1920 census sample is slightly positively biased as detailed age information is only reported for the regional capitals of the Brazilian districts. The census data of 1950 is again nationally representative. The trend of the 1830s to 1860s continues almost linearly to the trend of the 1890s to 1920s. The slightly upward biased urban sample in between confirms this development. Although there was a solid improvement, Brazil had not yet overcome age heaping by the 1920s, in contrast to Argentina and Uruguay. However the improvement in numeracy during the era of mass migration is considerable and as already in the Argentinean case, one can speculate about human capital enhancing effects from international migrants who came to Brazil.

Colombia

Evidence regarding Colombian numeracy is scarce. Calculations from our primary sources suggest

---

[61] This value refers to the 1770 birth cohort in Paraná. Slaves in the other birth cohorts were for the most part in the range of 35 to 60 ABCC points.

[62] To which degree are the early samples biased? Before the coffee boom and the large migration waves set in after the 1870s, the Southeast was actually not a region with higher human capital than the Brazilian average (except the urban Federal District of Rio de Janeiro), whereas the South was. We can consider the regional numeracy of the birth decade 1830s based on the 1890 census. In this decade, the ABCC indexes in Sao Paolo 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 large (around minus 1 percent). For the 1772 census, in which only Sao Paolo is represented, it might be in order of the 3 percent negative bias (i.e., the Brazilian national figure might be even higher).

that numeracy levels were relatively high during the early 18th century throughout the country, and especially so in the merchant city of *Cartagena de las Indias* (see Appendix B). The fact that ABCC indexes for a broad sample of provinces in the first half of the 19th century were all of the same level suggests that numeracy stagnated or improved only slightly between the early 18th and early 19th centuries. The 1928 census reveals a low numeracy rate for the birth decades of the 1880s and 1890s.

To what degree is each of the three early Colombian samples representative of the country? In Table 5 we report by region the number of observations for the 1777, 1870, and 1928 samples. These are sorted by the ABCC numeracy value in the 1930s birth decade, for which we possess data from all regions. Bogotá has the highest numeracy levels, with 98% of the population reporting an exact age, the centre and the Amazonas regions the lowest. Neither Bogotá nor the second-most numerate region, Eje Cafetero, is represented in these three early census samples, whereas the two least numerate regions are, suggesting that there may be some downward bias among them. The 1870 census represents a broad mix of regions. In 1777 and 1928, a higher share of regions fell into the lower half of the 1930s numeracy spectrum. Of course, the relative ranking probably changed between 1777 and 1930, but the regional composition may at least partly account for differences among the three samples. However, that there was stagnation between the 18th and late 19th centuries is quite likely, since the first and third samples were similar to one another in their regional composition.

Can this poor performance during the second half of the 19th century be explained by negative events that hampered efforts to improve Colombia's educational system? In fact, the many changes of government during this period and, more important, the civil war of 1876-1877 had a negative influence on educational levels. Ramirez and Salazar[63] find that funding of public education fluctuated with each change of regime. Moreover, during the civil war educational reforms were abandoned. Schools were transformed into hospitals, teachers were turned into

---

[63] Ramirez and Salazar, *Educaiòn en Colombia.*

soldiers, and governmental spending on education was suspended.

To summarise: Colombia's numeracy levels compared well with those of other Latin American countries at the beginning of the 18th century but stagnated during the 19th, resuming their upward course, toward convergence with those of its more advanced neighbours, only after 1900.

Ecuador

Our Ecuador data, derived from an 1871 census, provide complete coverage of the western provinces of Manabi, on the coast, and Azuay and Pichincha, inland: age, gender, occupation, and birthplace. There is no information on race or ethnicity in the tradition of all republican censuses. The sample comprises 71,545 observations. The birth decades, ranging as they do from 1800 to 1840, cover the critical period of independence. The 1800-1809 birth decade attains a numeracy level of only 54 ABCC points, much lower than the levels recorded elsewhere in Latin America (Appendix B). Between 1810 and 1830 there is a very slight improvement, and during the 1830s birth decade Ecuador's numeracy level rises sufficiently to converge with that of another Andean state, Colombia, but their levels remain the lowest in our panel. The numeracy level of the 1880s birth decade, represented by the 1950 census, is slightly below that of the 1840s birth decade, suggesting that levels stagnated throughout the 19th century. It is only after 1890 that Ecuador's level begins to rise, overtaking that of Mexico during the 1910-1919 birth decade.

Mexico

Our Mexican numeracy estimates begin in 1680 for the provinces of Hidalgo, Guanajuato, and Oaxaca, in central and southern Mexico (Figure 5).[64] The ABCC index suggests that only 40% of those born in these provinces in the 1680s were able to state their exact age. A number of sources indicate that by the mid 18th century basic numeracy had improved, with values varying between about 60 and 65%, and our sources for Coahuila (a northern province) during the late 18th century

---

[64] In order to have an adequate number of observations per birth decade, we were obliged to pool regions thus. For further details, see the Internet appendix B.

suggest that this trend continued. The census of Guadalajara (1821/22) in central Mexico reports considerably lower numeracy for the late 18th century but this was a regional phenomenon. Since northern Mexico's provinces were sparsely populated, but decently well-off in per capita terms, and therefore spent more on education than did those in the centre and the south,[65] it is possible that Guadalajara's lower numeracy level reflects this regional disparity. Because data for the birth decades from 1800 to 1860 are missing, we cannot prove that numeracy levels stagnated around the time of independence, but the fact that the 1870s birth decade's numeracy levels were low suggests that this was the case. For the birth decades from 1880 on (as represented by the 1930 and 1950 census samples), numeracy indexes remained quite low in Latin American comparison. To estimate the numeracy in Mexico we use a regression analysis similar to the one we used for Argentina, controlling for the share of males in each census and including a dummy variable for a census of the capital (Table 4, Columns 2 and 3). We also controlled for the indigenous percentage of the Mexican population, although it turned out to be insignificant. The trend in Mexico City, likewise, was not significantly different from that of the country as a whole, whereas the male share was significant for Mexico, indicating that in the 19th century gender differences in numeracy were greater in Mexico than in Argentina; results in line with those for the early 20th century.[66] Many of the time-dummy coefficients were insignificant (Table 4, for a graphical representation of time-dummy coefficients see Figure 3). Until 1750, Mexico's numeracy levels were higher than those of Argentina, but then they stagnated, whereas Argentina's soared.

Peru

The Spanish conquest of South America began with Peru in 1532. For at least the next 200 years thanks mostly to Bolivia's silver mines and Lima's proximity to the sea, the capital remained the continent's most important city. Lima's residents were mostly European, only a small portion of the

---

[65] With the exception of the federal district of Mexico City (Vaughan, 'Mexican education').
[66] Manzel and Baten, 'Gender and numeracy'.

indigenous population living within the city walls.[67] We have data for very early birth cohorts, those of the 1640s and 1650s, which are characterised by very low numeracy levels; those of the next two decades show some improvement. Age statements are of male household heads in Lima only and might therefore actually overestimate the general numeracy of the population. Because the data were limited to age statements made by male household heads, the estimates may indicate a higher level of numeracy than in fact existed among the general population.

Because we do not have age data for the birth decades of the 18th century, there is a 160-year gap, terminating in the 1820s with a sample of Lima prison inmates, many of whom were natives of the capital or the coastal region.[68] The numeracy estimates, especially those derived from data on the youngest prisoners, those born in the 1850s and 1860s, appear to be upwardly biased, when we compare them with the first birth cohort of the 1940 representative census of Lima. This 1940 census also allows us to estimate the numeracy gap between Lima and the rest of Peru: an unusually high difference of 17% for the 1880s birth decade. There are not sufficient census data from earlier decades to permit us to estimate Peruvian numeracy in a regression analysis; instead, we provide the available data in a table in the Internet Appendix B. For the same reason, we restrict our interpretation to the city of Lima.

Uruguay

Most of our Uruguay data derive from Montevideo. Numeracy in the capital in the early 18th century is at the same high, 60%, level as that for Buenos Aires: not surprising when one considers that Montevideo was founded in 1726 mainly by *Porteños* (inhabitants of Buenos Aires).[69] Two early-19th-century samples from the city's prison population display some random fluctuation, but their rates are not as low as those of rural Soriano and Maldonado (Appendix B). As we explained in the "Data Sources" section 3, such random fluctuation is associated with prison data. The fact

---

[67] Mabry, *Colonial Latin America*, p. 59.
[68] Twrdek and Manzel, 'Peruvian living standards'.
[69] After initially six families had arrived from Buenos Aires, they were followed by a second and third group of 13 and 25 families, respectively, who came from the Canary Islands. We thank an anonymous referee for his hint.

that Europeans constituted as much as 60% of this population helps to explain the relatively high level of numeracy in those samples.

However, the evidence suggests that throughout the late 18th century numeracy trends were nearly flat. If we assume that prison samples from the mid 19th century are, despite some variability, representative of the general population, then the numeracy level of the latter was about 80 to 85%. For 20th century birth decades there is the nationwide census of 1963. Age heaping has disappeared; this means that between the mid 19th century and 1900 Uruguay's overall numeracy rate must have risen sharply. To summarise: numeracy in Montevideo is at a high level in the early 18th century, stagnates in the late 18th century, and then, along with the rate in the rest of the country, soars during the late 19th century, perhaps thanks to mass immigration. If we assume that the mid-19th-century numeracy levels in rural regions were lower than those in the capital, then it follows that this increase would have been from a lower starting point and therefore even steeper if we had been able to calculate earlier rates for the entire country as well as for Montevideo.

Because we have little data for rural Uruguay, and the other explanatory variables are not reported systematically we cannot apply a regression analysis in order to adjust our results to account for the capital effect and other variables. We therefore decided to study the available data on Montevideo and Soriano/Maldonado up until the mid 19th century and display the trend for the whole country around 1900 graphically and in a table provided in Internet Appendix B.

Comparison of Latin American countries

We will now use a LOWESS regression, carrying out a weighted linear least-square fit of the data points, to estimate the general trends of average numeracy, thereby reducing short-term fluctuations in order to compare the the seven countries under study (Figure 6).[70] Data points near the point for which the response is being calculated are given greater weight than are the others.[71]

Argentina's numeracy levels are relatively low at the start of the period under study, but during the early 18th century they rise faster than do those of the other six countries, and by the

---

[70] LOWESS stands for "Locally Weighted Sum of Squares."
[71] Cleveland, 'Smoothing scatterplots'.

1750s reach those of Mexico, Uruguay (data for this period are mostly for Montevideo), and Colombia. After 1810 Mexico and Colombia began to lag behind the other five (but the Colombian lag may have been partly due to a regional bias). The numeracy levels of these two countries seem to have suffered the most from the wars of independence, the political instability of the early republican governments, capital flight, and the disruption of internal trade. The early-19th-century numeracy levels of Ecuador, for which we provide fairly reliable evidence, are extremely low until the end of the century.

In Mexico the conflicts between centralists and federalists (who will evolve into the conservative and the liberal parties) lasted until the 1860s, when President Benito Juárez, introduced a law that made primary education mandatory as well as free. During the *Porfiriato* (the presidency, spanning the years 1876-1911, of the war hero José de la Cruz Porfirio Díaz Mori), two decades later, primary schooling began to spread.[72] But especially the *Porfiriato* was also characterized by high inequality. Nevertheless, our trend for Mexico shows a small improvement in numeracy during these later decades of the 19th century, although we might want to interpret this as a convergence from an actually very low starting level.

The Brazilian region around São Paulo had favourable levels during the early 18th century. Around 1850, Brazil was on the fourth place, behind Lima, Uruguay, and Argentina, even if Brazil achieved independence relatively peacefully, and suffered less economic dislocation than did most other Latin American countries during their wars of independence. From the mid 19th century on, Brazil remains in the middle, trailing these three but trailed by Mexico, Colombia, and Ecuador until they start to catch up, after 1900. The Mexican Revolution of 1910, however, causes the Central American country to fall back even behind Ecuador and Colombia and to drop into last place.

How did European immigration to Latin America affect numeracy trends there? A study by Klein describes immigrants that raised the levels in Argentina and Brazil, often constructing their

---

[72] Andrade de Herrera, 'Education in Mexico'.

own schools and charitable institutions.[73] In addition, their entrepreneurial activities had a significant socioeconomic impact on Argentina, Brazil, and Uruguay. In order to assess whether the impact of immigration on numeracy development in Latin America was systematic or not, we constructed a small panel using the numeracy time series presented here (Figure 8). When we regressed these numeracy values on schooling estimates[74] and a dummy variable which controlled for substantial immigration that surpassed a certain threshold, we obtained a significantly positive relationship between schooling and numeracy (Table 6).[75] Even after we have controlled for schooling investments, the immigration variable has a p-value of 0.005. The R^2 suggests that about 40% of the numeracy variation is explained with these two variables (Adj. R^2=0.38). We assume that many other variables are reflected by the schooling investments, such as the quality of institutions and ethnic variation. Thus it seems safe to say that mass immigration from Europe to a given Latin American country had a positive impact on human capital formation there.

To summarise: numeracy levels improved in all seven countries in our sample, but less so in Mexico, Ecuador, and Colombia than in the other four, and almost all of them suffered a stagnation or even a decline during the late 18th or early 19th century. Mass immigration from Europe provided human capital enhancing effects, even when schooling investments are controlled for.

5. Comparison of Latin American countries with European countries and the US

How do the trends compare with those in Western Europe, Eastern Europe, and the US (Figure 7)? For Western Europe we use the ABCC index for the UK reported in A'Hearn, Baten, and Crayen (2009), and for East-Central Europe we use comprehensive Hungarian data from the same source.[76] As for the US, we draw on the research of Fischer[77] into the 17th-century colonies that would later become the US. Working with a sample of some 4,000 individuals in Essex County, Massachusetts, in the years 1636-72 (that is, the 1620s and 1630s birth cohorts, for the most part), Fischer

---

[73] Klein, *Migration to Brazil*.
[74] Benavot and Riddle, 'Education expansion'.
[75] We define it as 'more than 5,000 per decade and per country'.
[76] Moreover, we use some additional early-19th-century data as well (Baten and Benyus, 'Hungarian numeracy').
[77] Fischer, *Ageing*.

calculates the ratio between those colonists reporting ages with multiples of ten and a 10-year moving average. This ratio can be transformed into an ABCC index of 69% reporting exact ages. Another sample features adult males in Westchester County, New York (N = 259), most of them members of the 1670s birth cohort; their age-numeracy level was roughly 73% (Wells 1975).

Beginning with the birth decade of the 1800s we are on firmer ground, thanks to the censuses of 1850, 1870, and 1900. Because age heaping was greater in the South than elsewhere in the US during the 1800s, the increase in numeracy between the 1670s and the 1800s is slightly underestimated.

As for Western Europe, 16th-century English emigration records suggest an ABCC index of about 76% for the birth cohorts around 1600.[78] By about 1700 it would be close to 93%, on its way to 100% by the end of the next century.

During the 18th century Hungary's numeracy levels were on a par with Mexico's: slightly superior, that is, to Argentina's. However, by about 1780 trends in all seven Latin American countries under study, along with Hungary, had improved sufficiently to reduce the gap separating them from Western Europe and the US, which had been at 50% in 1650, to about 30%. However, Mexico's and Argentina's levels stagnated in the early 19th century, whereas Hungary's continued to rise, not only converging with but in fact surpassing, for a short while, those of the US, which stagnated, particularly in the South, until the middle of the century. During the late 19th century Argentina's trend resumed its upward course, reaching the basic-numeracy levels of Western Europe around 1880, whereas Mexico fell farther behind.

6. Conclusion

This paper makes the first systematic attempt to estimate the long-term development of human capital, measured in terms of basic numeracy, for seven Latin American countries from the 17th to the 20th century. Despite the existence of a large number of sizable 19th and 20th century censuses

---

[78] It should be noted, however, that the numeracy levels of emigrants tend to have an upward bias.

and of more limited ones from all four centuries, the data set suffers from some gaps and flaws, which we discussed intensively in this study. The large time span covers part of the colonial rule, the early post-independence period, the era of the first globalization as well as the first half of the 20th century. It thus permits new insights into the socioeconomic development of the New World. We discuss in detail the issue of source selectivity, and where possible adjust for regional composition before estimating long-term trends.

There was an overall increase in numeracy throughout the entire region and the entire period under study. Argentina, Mexico, and Peru (Lima) started, in the 17th century, with relatively low numeracy levels but by the 1780s they had reduced the gap separating them from Western Europe and the US from 50 to 30%. It is thus safe to say that until the late 18th century numeracy levels in Latin American countries were quite good. However, during the late 18th and early 19th centuries -- particularly the first two decades, marked by the wars of independence -- numeracy levels stagnated in many of these countries even as Western Europe's soared, and so the gap widened once again. In the late 19th century, numeracy gaps within Latin America increased, as well, Argentina, Uruguay, and Peru (Lima) at the upper end of the scale and Ecuador, Mexico, and Colombia at the lower. Brazil's numeracy levels stagnated until the 1860s but then began to improve. Immigration from Western Europe was positively associated with numeracy (even after we controlled for investment in public education). The fact that Ecuador, Mexico, and Colombia benefited relatively little from immigration and invested less in schooling helps to explain the fact that their numeracy levels stagnated in the late 19th century, and contributes to our understanding of their long-term histories.

What are the wider implications of the results? First, Latin America, which is partly still a developing region of the 21st century, was on a stable human capital growth path during the 18th century. We can quantify the long-term human capital retardation due to political conflicts of the early 19th century, which helps us to understand the more general relationship between conflict and human capital formation. Second, the historical evidence of immigration effects can be cautiously interpreted as a policy option for less-developed regions to invite skill-selective immigration if they

are able to integrate (and attract) those immigrants.

References

A'Hearn, B., Baten, J., and Crayen, D., 'Quantifying quantitative literacy: age heaping and the history of human capital', *Journal of Economic History*, 69, (2009), pp. 783-808.

Andrade de Herrera, V., 'Education in Mexico: historical and contemporary educational systems', In Le Blanc Flores, J., ed., *Children of La Frontera: Binational Efforts to Serve Mexican Migrant and Immigrant Students*, ERIC Clearinghouse on Rural Education and Small Schools, (Charleston, SC, 1996), pp. 25-29.

Andrews, G. R., *The Afro-Argentines of Buenos Aires, 1800-1900*, (Madison, 1980).

Alden, D., 'The population of Brazil in the late eighteenth century: a preliminary study', *The Hispanic American Historical Review*, 43 (1963), pp. 173-205.

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

Arrom, S. M., *The women of Mexico City 1790-1857*. (Stanford, CA, 1985).

Astorga, P., Bérges, A. R., and FitzGerald, E. V. K., 'The standard of living in Latin America during the twentieth century', *The Economic History Review*, 58 (2005), pp. 765-796.

Austin, R., *The state, literacy, and popular education in Chile, 1964-1990*. (Lanham, MD, 2003).

Bachi, R., 'The tendency to round off age returns: measurement and correction', *Bulletin of the International Statistical Institute*, 33 (1951), pp. 195-221.

Baer, W., *The Brazilian economy: growth and development* (Boulder, 6th edn. 2001).

Bakewell, P. J., *A history of Latin America: c. 1450 to the present* (Malden, MA, 2004).

Barman, R. J., *Brazil: the forging of a nation, 1798-1852* (Stanford, CA, 1994).

Baten, J., and Benyus, R., 'Numeracy and protestantism in Hungary', Univ. of Tuebingen working paper (2009).

Baten, J., Crayen, D., and Manzel, K.,'Zahlenfaehigkeit und Zahlendisziplin in Nord- und Westdeutschland, 16.-18. Jahrhundert', *Jahrbuch fuer Wirtschaftsgeschichte*, 2 (2008), pp. 217-229.

Baten, J., Pelger, I, and Twrdek, L., 'The anthropometric history of Argentina, Brazil and Peru during the 19th and early 20th centuries', *Economics and Human Biology*, 7 (2009), pp. 319-333.

Benavot, A., and Riddle, P., 'The expansion of primary education, 1870-1940: trends and issues', *Sociology of Education,* 66 (1988), pp. 191-120.

Bertola, L. (2004) An overview of the economic history of Uruguay since the 1870s [WWW document]. URL http://eh.net/encyclopedia/article/Bertola.Uruguay.final, [accessed on 29 September 2010].

Bertola, L., Ocampo, J.A. *Desarrollo, vaivenes y desigualdad. Una historia económica de América Latina desde la Independencia.* Madrid, Segib, 2010.

Caillet-Bois, R. R., *Documentos para la historia argentina,* vols X and XI, (Buenos Aires, 1920-55)

Calvo, H. and Meisel, A., *Cartagena de Indias en el siglo XVIII* (Catagena, 2005), p. 326.

Cinnirella, F., 'Optimists or pessimists? A reconsideration of nutritional status in Britain, 1740–1865', *European Review of Economic History,* 12 (2008), pp. 325–354.

Clark, G., *A farewell to alms. A brief economic history of the world.* (Princeton, NJ, Oxford, 2007).

Cleveland, W. S., 'LOWESS: a program for smoothing scatterplots by robust locally weighted regression', *The American Statistician,* 35 (1981), p. 54.

Closs, M. P., 'Maya Mathematics', in Grattan-Guiness, I., ed. *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, 2 vols, (2003).

Coatsworth, J. H., 'Economic and Institutional Trajectories in Nineteenth-Century Latin America', in J .H. Coatsworth, and A. M. Taylor, eds., *Latin America and the World Economy Since 1800* (Cambridge, MA, 1998), pp. 149-175.

Conant, L. L., *The number concept: its origin and development* (New York, 1923).

Cook, S. F., and Borah, W., *Ensayos sobre la historia de la población: México y el Caribe* (Madrid, 1999)

Cook, N. D., *Padrón de los indios de Lima en 1613*, N.D. Cook ed. (Lima 1968).

Cook, N.D., *Numeración general de todas las personas de ambos sexos, edades y calidades q se hà echo en esta ciudad de Lima, año de 1700*. N.D. Cook, ed., (Lima 1985).

Corona Baratech, C. E., 'Notas para un estudio de la sociedad en el Río de la Plata durante el Virreinato', *Anuario de Estudios Americanos*, 8 (1951), pp. 59-167.

Cortés Conde, R., *El progreso Argentino, 1880-1940* (Buenos Aires, 1985).

Crayen, D., and Baten, J.,'Global trends in numeracy 1820-1949 and its implications for long-run growth', *Explorations in Economic History*, 47 (2010), pp. 82-99.

De Moor, T., and Van Zanden, J.-L., 'Uit fouten kun je leren. Een kritische benadering van de mogelijkheden van 'leeftijdstapelen' voor sociaal-economisch historisch onderzoek naar gecijferdheid in het pre-industriële Vlaanderen en Nederland', *Tijdschrift voor Economische en Sociale Geschiedenis,* 5 (2008), pp. 55-86.

Emigh, R. J., 'Numeracy or Enumeration? The uses of numbers by states and societies', *Social Science History*, 26 (2002), pp. 653-698.

Engerman, S. L., Haber, S., and Sokoloff, K. L., 'Inequality, institutions, and differential paths of growth among New World economies', in *Institutions, Contracts, and Organizations*, C. Menard, ed., (Cheltenham, 2000).

Fischer, D. H., *Growing Old in America*. (New York 1977).

Fisher, J. R., *Bourbon Peru, 1750-1824*, (Liverpool 2003).

Graden, D.T., 'An act "even of public security": slave resistance, social tensions and the end of international slave trade to Brazil, 1835-1856', *The Hispanic American Historical Review*, 76 (1996), pp. 249-282.

Gootenberg, P., 'Population and ethnicity in Peru: some revisions', *Latin American Research Review*, 26 (1991), pp. 109-157.

Haber, S., and Klein, H. S., 'The economic consequences of Brazilian independence', in S. Haber, ed., *How Latin America fell behind: Essays on the economic histories of Brazil and Mexico, 1800-1914*, (Stanford, CA, 1997).

Hernández Sánchez-Barba, M., 'La población hispanoamericana y su distribución social en el siglo XVIII', *Revista de estudios políticos*, 78 (1954), pp. 111-142.

Instituto Nacional de Estadística, Geografía e Informática. *Censo de población de la Ciudad de México, 1790* [CD ROM]: *Censo de Revillagigedo*. (Aguascalientes, 2003).

Julien, C. J., *Die Inka : Geschichte, Kultur, Religion*. (Munich 2001).

Klein, H., 'Asian and European migration to Brazil', in R. Cohen, ed., *The Cambridge Survey of World Migration*, (Cambridge 1995), pp. 208-214.

Leff, N. H., 'Long-term viability of slavery in a backward closed economy', *Journal of Interdisciplinary History,* 5 (1974), pp. 103-108.

Leininger Pycior, J., 'The history of education in Latin America: an overview', in P. J. Rodolf, ed., *History of Education*, (New York 1984), pp. 33-49.

Lewis, W. A., *Tropical development, 1880-1913*, (London and New York 2006).

Lindert, P., *Growing public. Social spending and economic growth since the eighteenth century*, 2 vols. (New York, 2004).

Mabry, D. J., *Colonial Latin America*, (Coral Springs, FL, 2002).

Maddison, A. (2007) World population, GDP and per capita GDP, 1-2001 AD, [WWW document]. URL http://www.ggdc.net/maddison/, [accessed on 29 September 2010].

Manzel, K., and Baten, J., 'Gender equality and inequality in numeracy: the case of Latin America and the Caribbean, 1880-1949', *Revista de Historia Económica*, 27 (2009), pp. 37-73.

Martinéz-Fritscher, A.C., and Musacchio, V., 'The great leap forward in education in Brazil: the political economy of education in Brazil, 1890-1940' Working paper, Harvard Business School 10-075, (2009).

Mariscal, E., and Sokoloff, K. L., 'Schooling, suffrage, and inequality in the Americas, 1800-1945', in Haber, S., ed., *Political Institutions and Economic Growth in Latin America*, (Stanford, 2000).

McAlister L. N., 'Social structure and social change in New Spain', *The Hispanic American*

Historical Review, 43 (1963), pp. 349-370.

Mellafe, R., *Historia social de Chile y América* (Santiago de Chile, 2004).

Merino, O., and Newsom, L. A., 'Jesuit missions in Spanish America: The aftermath after the expulsion', *Revista de História de America,* 118 (1994), pp. 7-32.

Meyer Loy, C., 'Primary education during the Colombian Federation: the school reform of 1870' *Hispanic American Historical Review*, 51 (1971), pp. 275-294.

Mironov, B. N., 'Novaya Istoricheskaya Demografia Imperskoy Rossii: Analycheski obsor Corremennoy Istoriografii', *Vestnik Sankt-Peterbursgskovo Universiteta*, 4 (2006), pp. 62-90.

Mokyr, J., *Why Ireland starved: a quantitative and analytical history of the Irish economy*, *1800-1850* (London and Boston 1983).

Morse, R. M., 'Trends and patterns in Latin American urbanization, 1750-1920', *Comparative Studies in Society and History*, 16 (1974), pp. 416-447.

Myers, R., 'Accuracy of age reporting in the 1950 United States census', *Journal of the American Statistical Association*, XLIX (1954), pp. 826-831.

Newland, C., 'The estado docente and its expansion: Spanish American elementary education, 1900-1950', *Journal of Latin American Studies*, 26 (1994), pp. 449-467.

Núñez, C. E., and Tortella, G. (eds.). *La maldición divina: ignorancia y atraso en perspectiva histórica* (Madrid 1993).

Núñez, J., 'Signed with an X: methodology and data sources for analyzing the evolution of literacy in Latin America and the Caribbean, 1900-1950', *Latin American Research Review*, 40 (2005), pp. 117-135.

O'Grada, C., 'Dublin Jewish demography a century ago' *UCD Centre for Economic Research Working Paper Series,* (WP 06/01 2006).

Orozco y Berra, M., *Historia de la ciudad de México, desde su fundación hasta 1854*, Volume 112 of the Series "Sep-Setentas" published by the Secretaría de Educación Pública (SEP) (Mexico City. 1980).

Palafox, J. C., Prawda, J., and Velez, E., 'Primary schooling in Mexico', *Comparative Education Review*, 38 (1994), pp. 167-180.

Parro, A. A., 'Census of Peru: 1940', *Geographical Review*, 32 (1942), pp. 1-20.

Pérez Cantó, M. P., *Lima en el siglo XVIII* (Madrid, 1985).

Platt, L. D., *Census records for Latin America and the Hispanic United States* (Baltimore, MD, 1998).

Ramirez, M.T., and Salazar, I., 'Surgimiento de la educación en la República de Colombia, ¿En que fallamos?' *Borradores de Economia*, 454 (2007).

Reis, J., 'How helpful is age heaping? An exploration using Portuguese data', Univ. of Lisbon working paper (2008).

Ruggles, S., Trent Alexander, J., Genadek, K., Goeken, R., Schroeder, M. B., and Sobek, M., Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis, MI: University of Minnesota, (2010).

Sianesi, B., and Van Reenen, J., 'The returns to education: macroeconomics', *Journal of Economic Surveys,* 17 (2003), pp. 157-200.

Somoza, J., and Lattes, A., 'Muestras de los dos primeros censos nacionales de población, 1869 y 1895', Documento de Trabajo Instituto T. Di Tella, 46, (Buenos Aires, 1967).

Stolz, Y., Baten, J., and Botelho, T., 'Growth effects of 19th century mass migrations: the case of Brazil', Univ. of Tuebingen working paper (2010).

Thorp, R., *Progress, poverty and exclusion: an economic history of Latin America in the 20th Century* (Baltimore, 1998).

Twrdek, L., and Manzel, K., 'The seed of abundance and misery: Peruvian living standards from the early republican period to the end of the Guano era (1820-1880)', *Economics and Human Biology*, 8 (2010), pp. 145-152.

UN: *Demographic Yearbook, Special Topic Population Censuses*. (New York 1955).

Vaughan, M.K., 'Primary education and literacy in nineteenth-century Mexico: research trends, 1968-1988', *Latin American Research Review,* 25 (1990), pp. 31-66.

Vinson, B., 'Race and Badge: Free-Colored Soldiers in the Colonial Mexican Militia', *The Americas* 56 (2000), pp. 471-496

Wells, R. V., *The population of the British colonies before 1776: a survey of census data* (Princeton, NJ, 1975)

Werner, M. S. *Concise Encyclopedia of Mexico* (Chicago, 2001)

Wolff, L., Schiefelbein, E., and Valenzuela, J., 'Improving the Quality of Primary Education in Latin America and the Caribbean', *World Bank Discussion Papers*, 257 (1994), Washington.

Table 1: Data Sources

Country/Region	Year	Observations (age range 23-62)	Indiv. data?	Bias relative to total population?
Argentina				
1744 Buenos Aires	1744	3,179	yes	urban, military census, incl. slaves
1771 Buenos Aires	1771	11,140	yes	urban, incl. slaves
1778 Buenos Aires	1778	895	yes	urban
1818 Buenos Aires	1818	890	yes	capital, perhaps no slaves
1869 Argentina	1869	38,776	yes	no
1869 Buenos Aires	1869	5,005	yes	capital
1887 Santa Fé	1887	1,102	yes	regional
1895 Argentina	1895	43,897	yes	no
1895 Buenos Aires	1895	7,818	yes	capital
1914 Argentina	1914	3,286,844	no	no
1947 Argentina	1947	7,729,939	no	no
Brazil				
1772 São Paulo	1772	2,347	yes	regional, excl. slaves
1830 Brazil	1830	267,595	no	regional, incl. slaves
1870 São Christovão	1870	691	yes	urban upper-class, incl. slaves
1890 Brazil	1890	5,536,449	no	no
1920 Brazil	1920	1,010,056	no	no, representative sample
1950 Brazil	1950	13,798,696	no	no
Colombia				
1777 Cartagena	1777	2,431	yes	merchant city
1777 Colombia	1777	1,554	yes	regional
1870 Colombia	1870	2,387	yes	various regions
1928 Colombia	1928	567	yes	various regions
1963 Colombia	1963	6,058,045	no	no
Ecuador				
1871 Ecuador	1871	29,151	yes	regional
1950 Ecuador	1950	2,211,838	no	no
Mexico				
1740-44 Hidalgo/Guanajuato/Oaxaca	1740-44	1,228	yes	regional
1777 Mexico - Central	1777	4,379	yes	regional
1777 Mexico - City	1777	608	yes	capital
1777 Mexico - North	1777	705	yes	regional
1790 Mexico - City	1790	4,212	yes	capital, all households headed by Spaniards, and mestizos
1821 Guadalajara	1821	16,625	no	regional
1823 Coahuila	1823	1,598	yes	regional
1930 Mexico	1930	4,967	yes	regional
1950 Mexico	1950	9,934,234	no	no
Peru				
1700 Lima	1700	2,797	yes	capital, household heads
1866-1909 Lima prison	1866-1909	4,392	yes	capital, prisoners
1940 Lima	1940	352,755	no	capital
1940 Peru	1940	2,370,166	no	no
Uruguay				
1772 Montevideo	1772	1,362	yes	capital
1791 Montevideo	1791	2,96	yes	capital
1834-36 Soriano/Maldonado	1834-36	1,166	yes	regional
1846 Montevideo prison	1846	1,565	yes	capital, prisoners
1868 Montevideo prison	1868	1,268	yes	capital, prisoners
1963 Uruguay	1963	1,290,319	no	no

Table 2: Ethnic Composition of the Early Census Samples

Ethnic group	Composition estimates in previous literature	Composition estimates in our census samples
Argentina: Buenos Aires 1744	Corona Baratech (1951), Bs As 1744	
White/Spanish	80.2	87.9
Indigenous and Mestizo	2.9	3.73
Black and Mulatto	16.9	6.56
Argentina: Buenos Aires 1771	Corona Baratech (1951), Bs As 1770	
White/Spanish	66.8	66.8
Indigenous and free Black	4.8	5.8
Slave	28.4	21.9 (Black slaves 13.58)
Mexico 1777: Oaxaca		
White/Spanish	12.33	5.45
Pardo	3.75	5.29
Indigenous	83.92	12.68
N/A	0	64.44
Mexico 1777: Durango	Cook and Borah (1999), Table 20b	
Mestizo	81.61	33.3
Indigenous	18.39	66.3
Mexico 1777: Mexico City	Hernández Sánchez-Barba (1954), 127, urban population in Mexico, ca. 1794	
White/Spanish	49.27	48.91
Casta	26.64	34.58
Indigenous	24.0	5.30
Other	0	11.21
Mexico 1790: Mexico City	Hernández Sánchez-Barba (1954),127, urban population in Mexico, ca. 1794	
White/Spanish	49.27	48.40
Casta	26.64	21.12
Indigenous	24.0	22.37
Black	0	0.56
Other	0	7.55
Peru 1700: Lima	Pérez Cantó (1984), Lima 1700	
White/Spanish	56.5	97.11
Indigenous	11.7	1.5
Mulatto	9.7	0.75
Black	22.1	0.3
Peru 1871: Lima prison	Fisher (2003), 56, Peru in late 18th century	
White/Spanish	38.46	17.69
Indigenous	7.69	21.31
Casta	9.6	25.22
Black	44	13.90

Table 3: Regression of ABCC Outcomes on Reasons for Enumeration, All Countries

Dependent Variable	ABCC
Military	4.98
	(0.403)
Prison	4.88
	(0.198)
Country-Fixed Effects	included
Time-Fixed Effects	included
Constant	27.70***
	(0.004)
Observations	171
R-squared	0.74

in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Reference Category is an enumeration reason other than military or prison, the country is Argentina, and the birth decade is the 1640s.

Table 4: Regressions of numeracy (ABCC Index) on Male Share, Capital Effect, and Birth Decade

for Argentina and Mexico

Dependent Variable	(1) ABCC	(2) ABCC	(3) ABCC
Country	Argentina	Mexico	Mexico
Capital	6.41***	-0.43	-1.37
	(0.000)	(0.917)	(0.812)
Male share	0.31**	1.59**	1.52*
	(0.037)	(0.024)	(0.059)
Indigenous share			-0.06
			(0.789)
b1680	-80.30***	-30.25***	-31.06**
	(0.000)	(0.005)	(0.022)
b1690	-82.35***	-19.18*	-19.99
	(0.000)	(0.054)	(0.118)
b1700	-74.16***	-14.12	-14.94
	(0.000)	(0.145)	(0.234)
b1710	-64.16***	-15.72**	-16.38
	(0.000)	(0.047)	(0.127)
b1720	-56.28***	-19.07**	-19.69*
	(0.000)	(0.033)	(0.090)
b1730	-52.03***	-14.40*	-14.89
	(0.000)	(0.071)	(0.164)
b1740	-41.13***	-14.28*	-14.77
	(0.000)	(0.073)	(0.167)
b1750	-43.30***	0.51	0.38
	(0.000)	(0.960)	(0.976)
b1760	-29.70***	-8.99	-8.88
	(0.000)	(0.218)	(0.384)
b1770	-33.38***	-6.87	-6.65
	(0.000)	(0.377)	(0.534)
b1780	-26.68***	-4.21	-3.98
	(0.000)	(0.585)	(0.708)
b1790	-27.80***	-10.26	-10.03
	(0.000)	(0.194)	(0.353)
b1800			
b1810	-40.45***		
	(0.000)		
b1820	-34.65***		
	(0.000)		
b1830	-29.51***		
	(0.000)		
b1840	-27.12***		
	(0.000)		
b1850	-19.67***		
	(0.000)		
b1860	-17.81***		
	(0.000)		
b1870	-9.34*	-13.25	
	(0.060)	(0.170)	
b1880	-3.42	-7.71	
	(0.409)	(0.416)	
b1890	-0.00	-4.60	-4.94
	(1.000)	(0.550)	(0.682)
b1900	-0.00	6.68	6.41
	(1.000)	(0.480)	(0.595)
b1920		13.46	13.19

		(0.164)	(0.283)
Constant	84.71***	-9.51	-4.26
	(0.000)	(0.770)	(0.918)
Observations	44	36	32
R-squared	0.98	0.70	0.69

Note: p-values in parentheses. */**/*** denote statistical significance at the 10, 5, and 1% levels, respectively. The constant represents the numeracy of female, non-capital inhabitants born in 1910-19 (in Col. 3 non-indigenous).

Table 5: Regional Distribution of Observations in Colombia, Sorted by Regional Numeracy in 1930-39

Region	ABCC, birth decade 1930s	Number of observations on numeracy in		
		1777	1870	1928
Central	88	236		
Amazonia	91		1016	419
Andina Sur	91		287	
Andina Norte	91	447		
Orinoquia	91		176	
Pacifico Norte	93		184	148
Caribe	93	99	100	
Pacifico Sur	93		597	
Eje Cafetero	95			
Bogotá	98			

Table 6: Regression: Impact of Immigration and Schooling on Numeracy Development

Dependent Variable	(1) ABCC
Immigration Dummy	12.68***
	(0.005)
Schooling	0.46**
	(0.042)
Constant	67.17***
	(0.000)
Observations	32
R-squared	0.42

in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Immigration is measured by an indicator variable that is 1 when a country experiences substantial immigration, defined as over 5,000 arrivals per decade. Data on immigration are from Mitchell (1975) and schooling data are from Benavot and Riddle (1983); for the numeracy estimates, see text. The schooling data allow us to include the following countries and decades (decades during which immigration was substantial are provided in parentheses): Argentina 1870 and 1890-1910 (1860-1910), Brazil 1870-1920 (1880-1920), Colombia 1880-1930, Ecuador 1890-1920, Mexico 1870 and 1890-1920 (1900-1920), Peru 1890-1910, and Uruguay 1900-1930 (1840 and 1900-1930). Results also hold with controls for time-fixed effects, but in this case the schooling estimates are no longer significant.
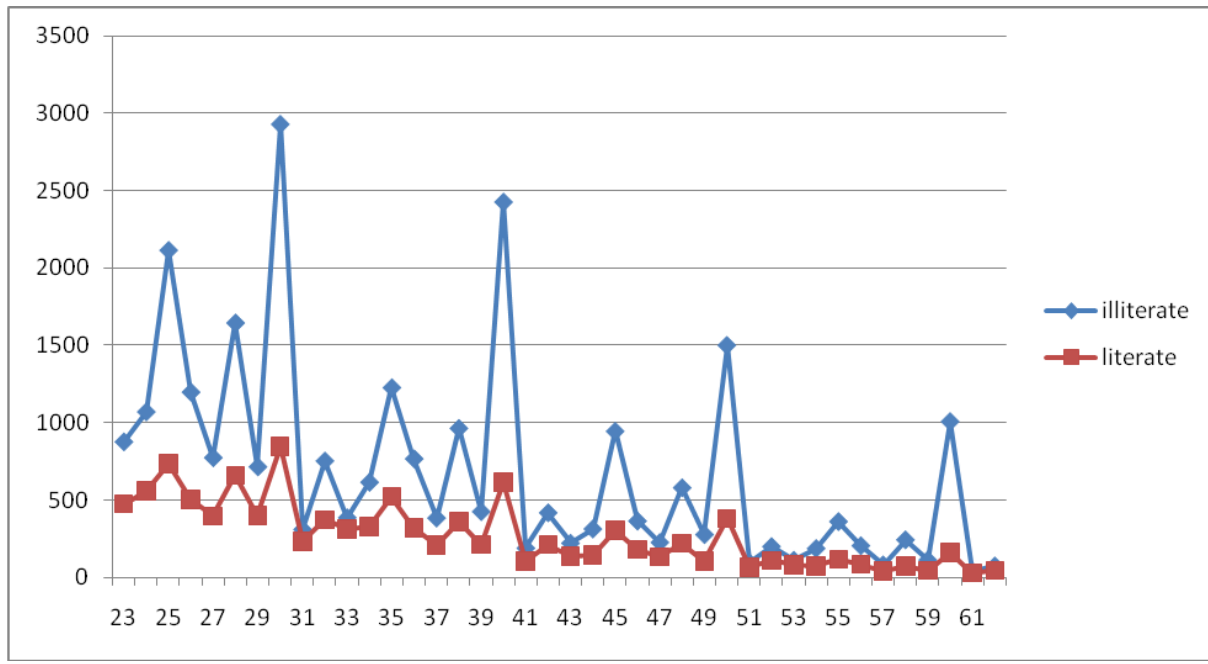
Figures

N

Figure 1: Age statements by literacy status (1869 census of Argentina in 1869)
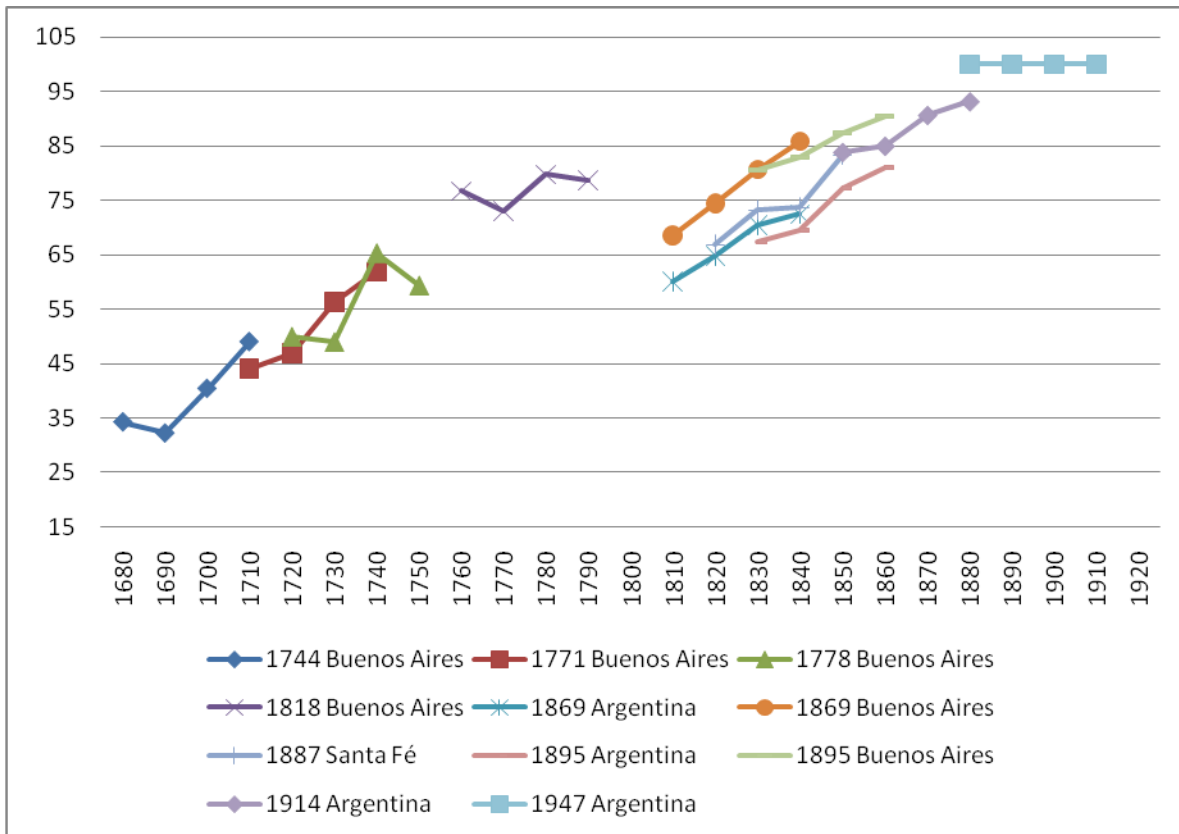
Figure 2: Argentina - ABCC index of basic numeracy (vertical axis) by birth decades

Note: The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1.
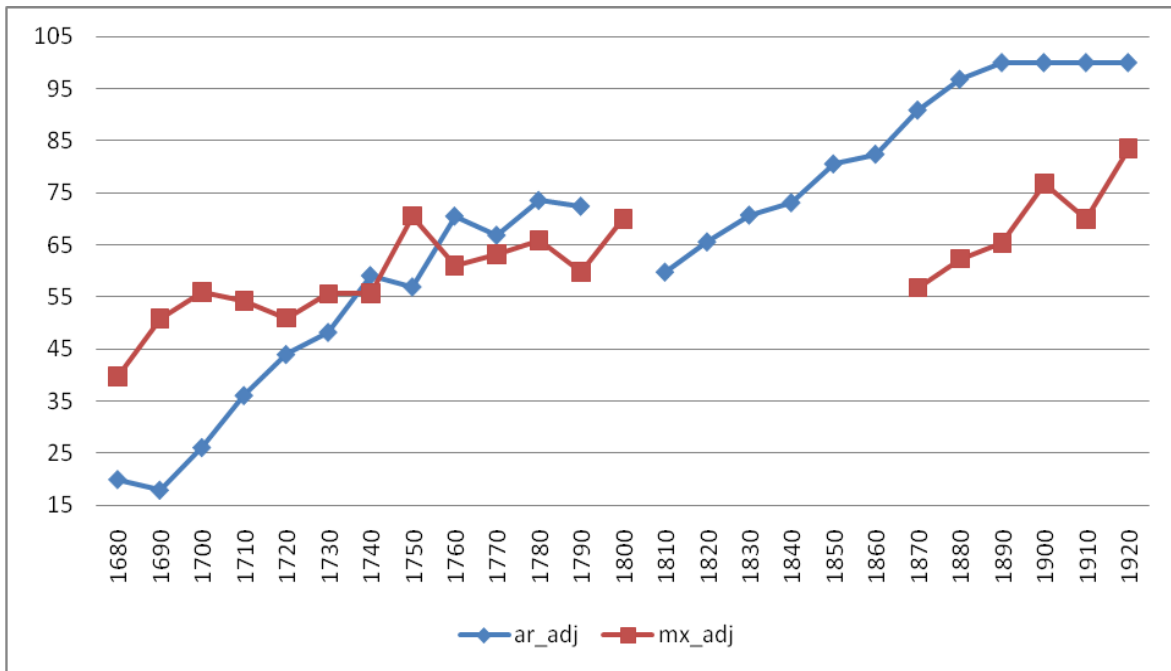
Figure 3: Trends of the ABCC index for Argentina and Mexico, (vertical axis) by birth decades, controlling for capital effect and gender composition

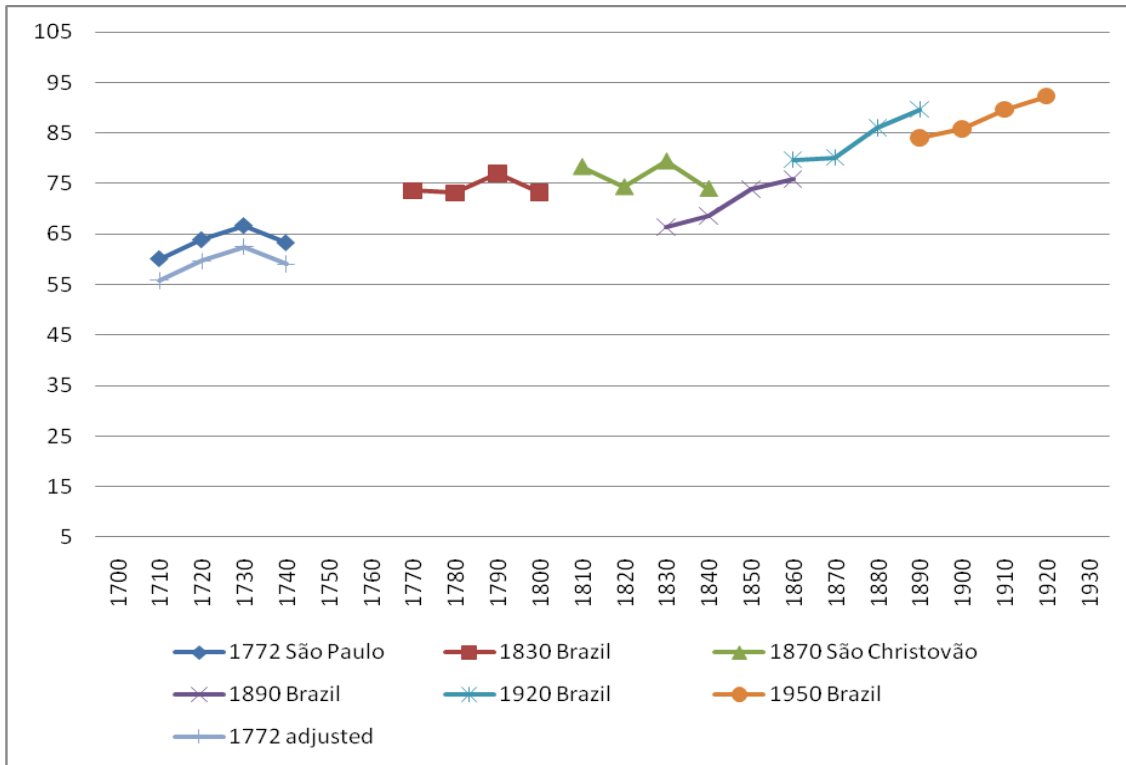Note: The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1.

Figure 4: Brazil - ABCC index of basic numeracy (vertical axis) by birth decades

Note: The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1.
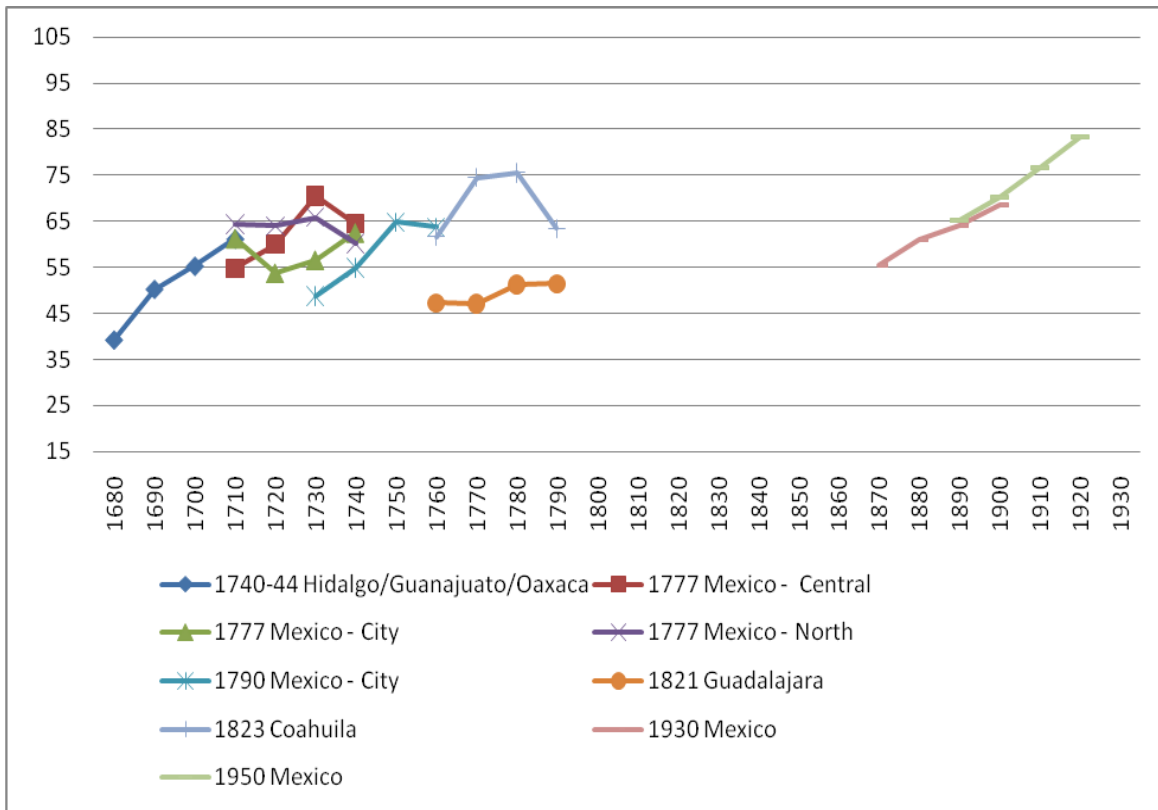
Figure 5: Mexico - ABCC index of basic numeracy (vertical axis) by birth decades

Note: The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1.
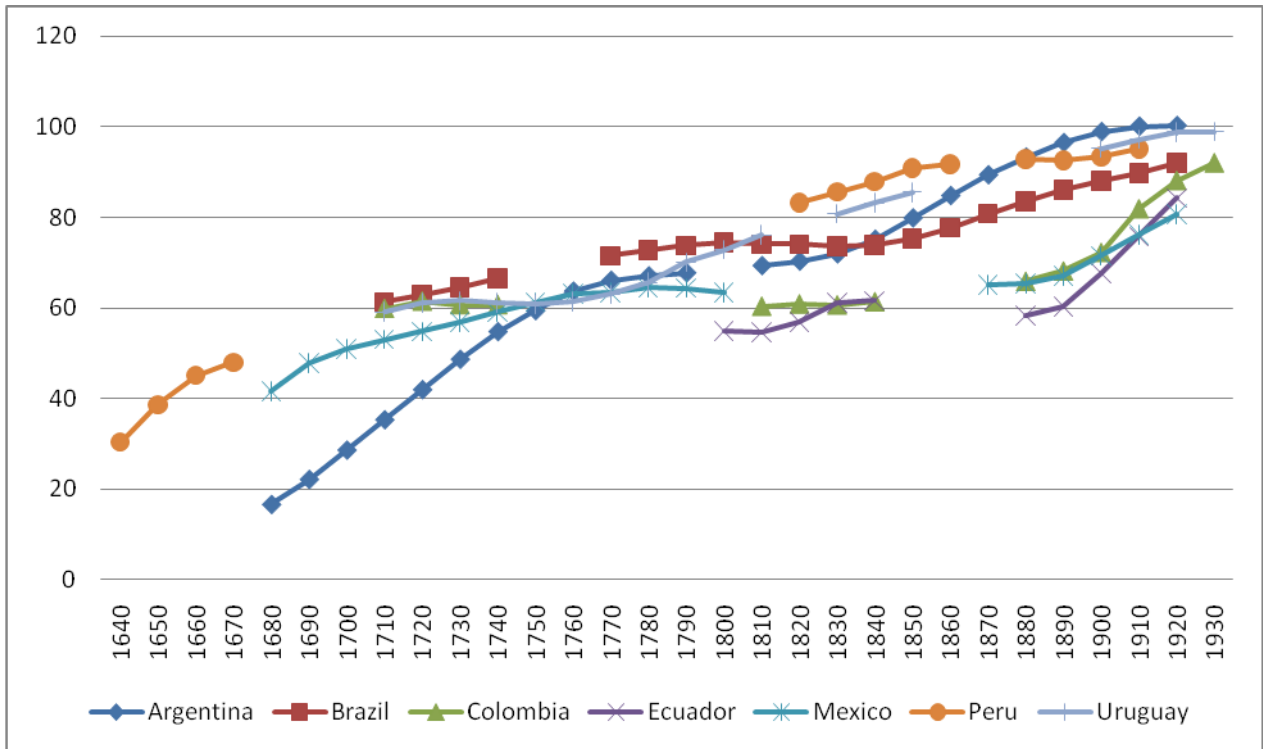
Figure 6: Argentina, Brazil, Colombia, Ecuador, Mexico, Peru (Lima), and Uruguay -- ABCC index of basic numeracy (vertical axis) by birth decades, LOWESS-transformed

Note : Only the values for Lima are included in the Peruvian data. For Argentina and Mexico, estimates are based on regressions (controlling for capital effect and male share). LOWESS bandwidth is 0.5. The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1. Non-representative samples, such as that of São Christovão, for Brazil, are excluded.
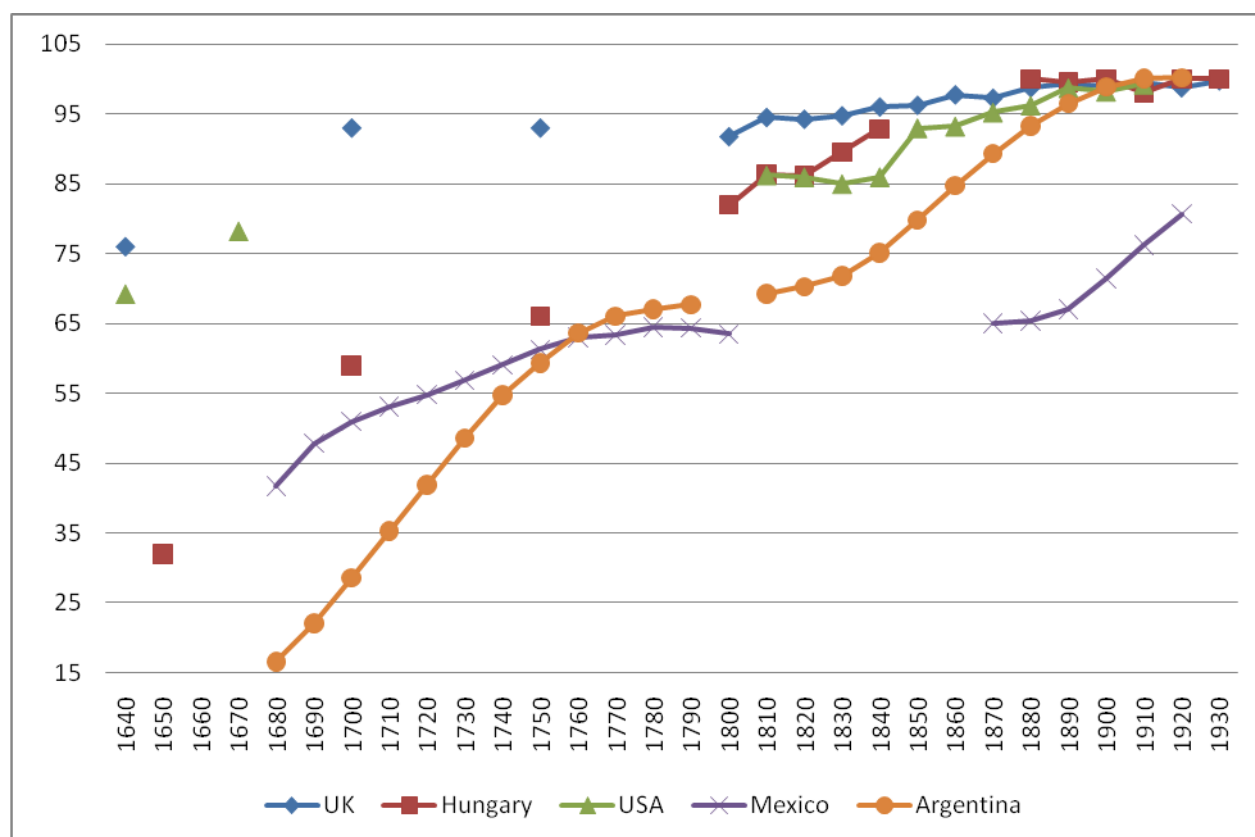
Figure 7: ABCC numeracy indexes for Argentina, Mexico, the UK, Hungary, and the US (vertical axis) by birth decades

Notes: For Argentina and Mexico, estimates are based on regressions (controlling for capital effect and male share). The value "UK 1640" actually refers to the ABCC of around 1600; the earliest US value is for Massachusetts 1620s and 1630s. Hungary during the early 19th century refers to the Hungarian part of the Habsburg monarchy. Sources for Hungary and the UK are from A'Hearn, Baten, and Crayen (2009). For Hungary, the US, and the UK during the 19th century, see Crayen and Baten (2009). Some of the Hungarian values come from Baten and Benyus (2009). The US values are based on the Integrated Public Use Micro Samples (IPUMS); see Ruggles et al. (2004). The decades refer to birth decades (1800-1809, etc.). We thank Rita Benyus, Brian A'Hearn, Dorothee Crayen, and the IPUMS team for providing data.

Figure 8: Numeracy of immigrants and natives in Brazil and Argentina

Note: Numeracy of Argentine Data taken from the Argentine censuses of 1869, 1895, and 1914 and also from Brazilian data on native numeracy; see text. Brazilian immigrants are documented in Stolz, Baten, and Botelho (2010).
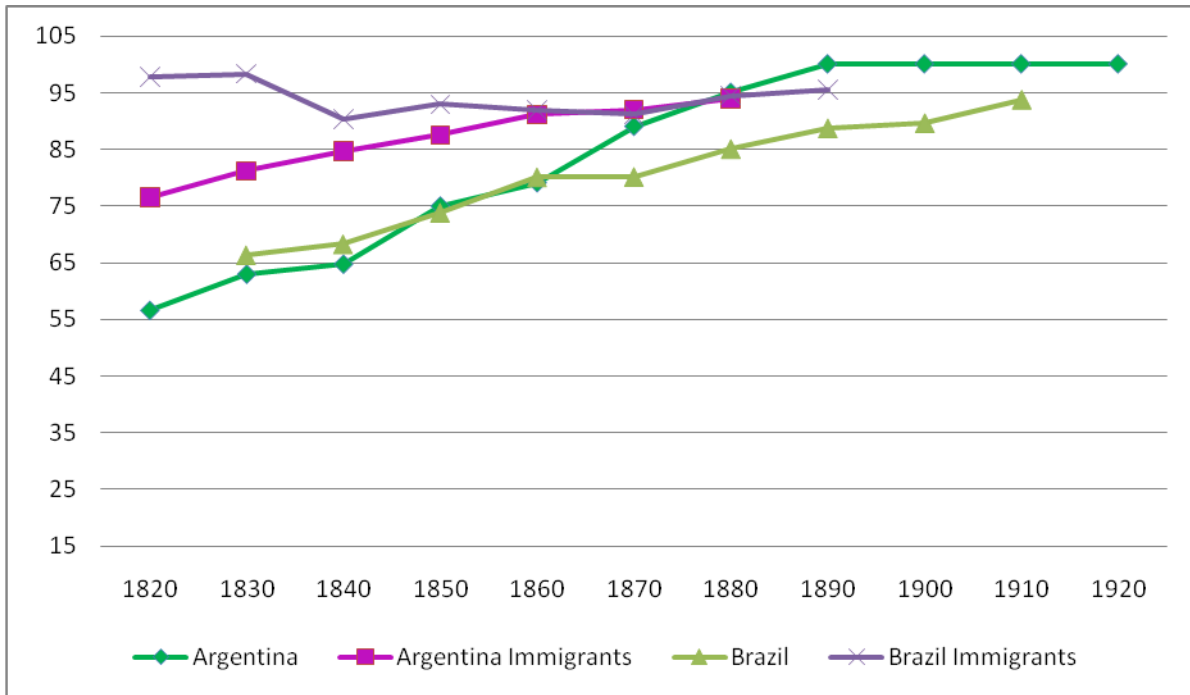
Appendix A: Data Sources (to be provided on an Internet page)

Sources for Argentina

There are a large number of primary sources that provide information on the development of numeracy in Argentina, the earliest being the 1744 Buenos Aires military census, reproduced in the *Documentos para la Historia Argentina* (Caillet-Bois 1919). For Buenos Aires we also have 1771 census data. We use information on 11,140 individuals in this data set (children and the elderly are excluded from our calculations). Data from the military census of 1818 fill the gap between 1771 and 1869. In addition, we incorporate into our analysis the samples of the first two national population censuses, those of 1869 and 1895, collected by Somoza and Lattes (1967), which contain extensive information on a representative nationwide sample of the population. These data provide not only insights into regional differences but information on urban and rural areas as well. The time series is complemented with aggregated data from the 1914 and 1947 censuses.

Sources for Brazil

Our earliest evidence for Brazil, dating from 1772, consists of a single sample of a very early census of the São Paulo district of Sorocaba, representing the birth decades 1700-1750 (N=6,279). While this census does not include slaves, it does include *agregados,* akin to live-in servants in pre-industrial rural northern Europe. *Agregados* were free but mostly unpaid, performing household tasks in exchange for food and shelter (Stolz, Baten, and Botelho 2010). We also include a complete aggregated 1830 census evidence for Paraná, Minas Gerais, and São Paulo, which includes information also on slaves and thus can be deemed representative of late-18th -century Brazil, or at least the southern region thereof. Further evidence comes from the 1890 and 1920 censuses. The 1920 data are limited in that only the provincial capitals are reported in sufficient age-specific detail. Last but not least, we have the nationwide aggregated 1950 census.

Sources for Colombia

The Colombian census data derive from the National Archive in Bogotá, the earliest being

population counts from the year 1777 from several central and Northern provinces.[79] The census records of Cartagena de las Indias have been studied extensively by Calvo and Meisel (2005). Information on four districts of this important merchant city has survived. The 1870 census' information on Cauca, Magdalena, Chocó, Quindio, and other departments and thus hints offers hints about the development of basic numerical abilities in a wide variety of provinces. Finally, the 1928 census of Putumayo, Vaupés, and Chocó provides individual age data.

Sources for Ecuador

Our Ecuador data, mostly for the year 1871, derive from the *Archivo Nacional de Quito*. The complete enumeration, extending from the 1860s to 1893 and comprising 71,545 observations, covers the provinces of Manabi, Azuay, and Pichincha. The census offers complete breakdown of the enumerated population. To account for the birth decades of the late 19th and early 20th centuries, we include the published ABCC values taken from the 1950 nationwide census (Manzel and Baten 2009).

Sources for Mexico

The *Archivo de las Indias* in Seville houses a considerable wealth of primary sources for Mexico. Our sample includes population censuses from 1740 to 1743 for Hidalgo, Guadalajara, and Oaxaca. We also had access to age data from 1777 for Mexico City, Durango, Chihuahua, Baja California, Oaxaca, Puebla, and Veracruz. The *Censo de Revillagigedo,* our most important source for all of northern Latin America, and carried out in Mexico between 1790 and 1794, was the first census equipped with a standard format for listing individuals by name, age, sex, and family status (Werner 2001). Those census forms that have survived provide information on 15 Mexico City districts, from which we have drawn a sample of 4,212 individuals.

The Guadalajara Census, a joint project of researchers from Mexico and the US, aims to preserve and to provide public access to census data for Guadalajara, a province in the Western Pacific area of Mexico. Another source provides data from the 1823 census of Coahuila, a province

---

[79] We thank Adolfo Meisel for providing additional data 1777 Cartagena de las Indias sample.

bordering Texas. For the 20th century, we have data from 1930 on several other provinces and aggregated data from the 1950 nationwide census.

Sources for Peru

Our earliest source of Peruvian age data is a 1770 Lima population census considered one of the most important of the colonial period (Cook 1985). Because its main purpose was to determine the number of men available for military service in case of foreign aggression, the data are limited to the age, profession, and race of male household heads. Thus we have information on no more than about 3,000 of Lima's 37,000 inhabitants (Perez Canto 1985, 185), and none at all on the female portion of Lima's population: a situation that virtually guarantees a numeracy overestimation -- all the more nearly certain because Peru's indigenous population was underrepresented in the sample since it was underrepresented in Lima, relative to the proportion that it constituted of other cities' populations (Mabry 2002, 58). Our analysis of numeracy in Peru is therefore primarily an analysis of the educational development of Lima's elite. However, we also have access to the 1940 nationwide census data, which shed light on the birth cohorts from the 1880s onwards, permitting us to compare numeracy rates in Lima with that in Peru generally. In addition, we have access to a Lima prison sample that extends from 1866 to 1909. Although one assumes, understandably, that this prison sample overrepresents Lima's underclass, in fact the age heaping to be found in the portion of this sample spanning the 1880s is quite similar to that of Lima's general population (Baten, Pelger, and Twrdek 2008).

Sources for Uruguay

The National Archive in Montevideo contains a variety of data sets valuable for the study of the evolution of numeracy in Uruguay: prison records (1847-1868), the 1832 census of Soriano, and the 1836 census of Maldonado. Finally, there is also the published nationwide census of 1963.

Data Sources

Argentina:
Military Census 1744: Reproduced in Caillet-Bois, R. R. (1919); **Census 1771**: Archivo Nacional Buenos Aires, Sala IX, Documentos de Gobierno, Censos y padrones; **Census Alcalde Matheo 1778; Census of Santa Fé 1887:** http://www.digitalmicrofilm.com.ar /censos/geografico.php; **Census of Buenos Aires 1818**: Archivo Nacional Buenos Aires, Sala IX, Documentos de Gobierno, Censos y padrones, **Census of Argentina 1869**: Somoza and Lattes (1967); **Census of Argentina 1895:** Somoza and Lattes (1967), **Census of Argentina 1914**: Archivo Nacional Buenos Aires, Sala IX, Documentos de Gobierno, Censos y padrones, **Census of Argentina 1947**: UN (1955): Demographic Yearbook, p. 311.

Brazil:
Census São Paulo, Sorocaba 1772: Arquivo Hist«orico Ultramarino, Lisboa, Portugal, AHU_ACL_CU_,Cód.1270, **Census of Paraná, Minas Gerais, São Paulo 1830: # Census of Brazil 1890**: Biblioteca do IBGE, [http://biblioteca.ibge.gov.br/]; **Census of Brazil 1920**: http://biblioteca.ibge.gov.br/ **Census of Brazil 1950**: UN (1955), Demographic Yearbook, p. 313.

Colombia:
Census of Cartagena 1777: Archivo Nacional de Bogotá, Microfilm 23, Volume 8 (1), No. 9 and 58, Meisel (2005); **Census of Media Granada 1777:** Archivo Nacional de Bogotá, Microfilm 23, Volume 8 (1), No. 11; **Census of Magdalena 1777:** Archivo Nacional de Bogotá, Microfilm 23, Volume 8 (1), No. 15-16; **Census of Mogotes 1777:** Archivo Nacional de Bogotá, Microfilm 23, Volume 8 (1), No. 38; **Census of San Juan Girón 1777:** Archivo Nacional de Bogotá, Microfilm 23, Volume 8 (1), No. 53; **Census of Sativa 1777:** Archivo Nacional de Bogotá, Microfilm 23, Volume 8 (1), No. 61; **Census of Bolivar 1777:** Archivo Nacional de Bogotá, Microfilm 23, Volume 8 (1), No. 62; **Census of Cauca 1870:** Archivo Nacional de Bogotá, Microfilm 2, No. 4; **Census of Magdalena 1870:** Archivo Nacional de Bogotá, No. 6; **Census of Quibdo 1870:** Archivo Nacional de Bogotá, Microfilm 2, No. 15; **Census of Quindio 1870:** Archivo Nacional de Bogotá, Microfilm 2, No. 19; **Census of Putumayo 1928:** Archivo Nacional de Bogotá; **Census of Vaupés 1928:** Archivo Nacional de Bogotá; **Census of Chocó 1928:** Archivo Nacional de Bogotá; Census of Colombia 1963: UN (1972): Demographic Yearbook. For the regional breakdown in Table 5 we used the IPUMS International Sample on the Colombian Census of 1963; see Ruggles et al. (2004).

Ecuador
Census for all Ecuador excluding the Amazonas region and the Galapagos Archipelago. It was taken between **1861 and 1893**. Archivo Nacional de Quito. **1950** Census of Ecuador. See Manzel and Baten 2009.

Mexico
[Mexico 1740-44] Census of Ixmiquilpan 1740: Archivo General de Indias, Ind, 107; **Census of Pozos 1743:** Archivo General de Indias, Ind, 107; **Census of southern central Mexico 1743:** [Place unreadable, Platt (1998): Tlazazalca, Michoacán, Tetela del Rio or Guerrero], Archivo General de las Indias, Ind, 108; **Census of Chichihualtepec 1743:** Archivo General de las Indias, Ind, 108;
[Centro 1777] Census of Ciudad de los Angeles 1777: Archivo General de las Indias, Mex, 2578; **Census of Piaxtla 1777:** Archivo General de las Indias, Mex, 2578; **Census of Totoltepec 1777:** Archivo General de las Indias, Mex, 2579; **Census of Nopalucan 1777:** Archivo General de las Indias, Mex, 2579; **Census of Quanquecholan 1777:** Archivo General de las Indias, Mex, 2579;
[Norte 1777] Census of San Gregorio 1777: Archivo General de las Indias; **Census of Real de**

Minas de Nuestra Senora del Rosario 1777: Archivo General de Indias, Gua, 103 and Gua, 250;**Census of los Remedios 1777:** Archivo General de las Indias, Ind, 1526; **Census of San José de Animas 1777:** Archivo General de las Indias; **Census of San José de Pimas 1777:** Archivo General de las Indias; **Census of Suchil 1777:** Archivo General de las Indias; **Census of Penol 1777:** Archivo General de las Indias; **Census of Tonanchi 1778:** Archivo General de las Indias; **[Mexico City 1777] Census of Mexico City 1777:** Archivo General de las Indias;
Census of Revillagigedo 1790: Instituto Nacional de Estadística, Geografía e Informática (2003); **Census of Guadalajara 1821:** Guadaljara Census project [http://www.fsu.edu/]; **Census of the Municipality of Monclova 1822-23**; **Censo del Pueblo de San Francisco de Tlaxcala 1823**; **Censo de la Hacienda de Castaños y Bajan 1822-23**; **Censo de la Hacienda de Alamo 1823**; **Censo de la Hacienda de Encinas 1823**; **Censo de la Hacienda de San Vicente el Alto 1823**; **Censo de la Hacienda de Santa Ana 1823**; **Censo de la Hacienda de San Juan Bautista 1823**; **Censo de la Hacienda de San José 1823**; **Censo de la Hacienda de San Ignacio del Paso Tapado 1823:** Grupo Exploradores Coahuiltecos [http://mx.geocities.com/camino_real_ mva/]; **Census of Guanajuato 1930:** FSI, Microfilm 4107114; **Census of Minas de Luz 1930:** FSI, Microfilm 4107114; **Census of Mineral de los Llamitos 1930:** FSI, Microfilm 4107114; **Census of Ahualuco 1930:** FSI, Microfilm 4107751; **Census of Coyuca de Benitez 1930:** FSI, Microfilm 4107141; **Census of Tepoztlán 1930:** FSI, Microfilm 4107265; **Census of Mezquital 1930:** FSI, Microfilm 4107065; **Census of Tetecala 1930:** FSI, Microfilm 4107265; **Census of Tlaltizapan 1930:** FSI, Microfilm 4107265; **Census of Tetecala 1930:** FSI, Microfilm 4107265; **Census of Tlaltizapan 1930:** FSI, Microfilm 4107265; **Census of Mexico 1950:** UN (1955), Demographic Yearbook, p. 304.

Peru
Census of Lima 1700: Reproduced in Cook (1985); **Prison sample Lima 1871:** Manzel and Twrdek (2009); **Census of Peru 1940:** Parro (1942).

Uruguay
Padrón of Aldecoa 1772: Reproduced in Apolant (1975), Volume III; **Census of Soriano 1834:** Archivo Nacional de Montevideo; **Census of Maldonado 1836:** Archivo Nacional de Montevideo; **Prison sample 1846:** Archivo Nacional de Montevideo, **Prison sample 1868:** Archivo Nacional de Montevideo; **Census of Uruguay 1963:** UN (1972): Demographic Yearbook, p. 214.

United Kingdom
Before 1800: A'Hearn, Baten, and Crayen (2009). After 1800: Crayen and Baten (2009), Baten and Benyus (2009).

United States
Census of Westchester County: Wells (1975); **Census of 1850, 1870 and 1900:** A'Hearn, Baten, and Crayen (2009), based on Integrated Public Use Micro Samples (IPUMS). See Ruggles et al., Integrated Public Use.

Appendix B: Whipple and ABCC indexes (to be provided on an Internet page)

Data Source	Birth decade	Observations	Whipple	ABCC
1744 Buenos Aires	1680	303	363	34
1744 Buenos Aires	1690	501	371	32
1744 Buenos Aires	1700	941	338	40
1744 Buenos Aires	1710	1,434	304	49
1771 Buenos Aires	1710	949	324	44
1771 Buenos Aires	1720	1,832	313	47
1771 Buenos Aires	1730	3,358	275	56
1771 Buenos Aires	1740	5,001	253	62
1778 Buenos Aires	1720	105	300	50
1778 Buenos Aires	1730	153	304	49
1778 Buenos Aires	1740	243	239	65
1778 Buenos Aires	1750	394	262	59
1818 Buenos Aires	1760	88	193	77
1818 Buenos Aires	1770	190	208	73
1818 Buenos Aires	1780	254	181	80
1818 Buenos Aires	1790	358	186	79
1869 Argentina	1810	3,221	260	60
1869 Argentina	1820	6,598	241	65
1869 Argentina	1830	11,142	218	70
1869 Argentina	1840	17,815	210	72
1869 Buenos Aires	1810	352	226	69
1869 Buenos Aires	1820	826	202	74
1869 Buenos Aires	1830	1,493	177	81
1869 Buenos Aires	1840	2,334	157	86
1887 Santa Fé	1820	114	232	67
1887 Santa Fé	1830	234	207	73
1887 Santa Fé	1840	280	205	74
1887 Santa Fé	1850	474	167	83
1895 Argentina	1830	4,001	230	67
1895 Argentina	1840	7,989	222	69
1895 Argentina	1850	13,875	191	77
1895 Argentina	1860	18,032	176	81
1895 Buenos Aires	1830	682	177	81
1895 Buenos Aires	1840	1,415	169	83
1895 Buenos Aires	1850	2,548	150	87
1895 Buenos Aires	1860	3,173	139	90
1914 Argentina	1850	340,213	165	84
1914 Argentina	1860	574,992	160	85
1914 Argentina	1870	922,034	137	91
1914 Argentina	1880	1,449,605	127	93
1947 Argentina	1880	1,140,200	96	100
1947 Argentina	1890	1,697,562	100	100
1947 Argentina	1900	2,286,936	99	100
1947 Argentina	1910	2,605,241	98	100
1772 São Paulo	1740	968	247	63

1772 São Paulo	1730	639	233	67
1772 São Paulo	1720	409	244	64
1772 São Paulo	1710	331	260	60
1830 Brazil	1800	120,000	208	73
1830 Brazil	1790	73,325	193	77
1830 Brazil	1780	46,458	211	73
1830 Brazil	1770	27,812	212	74
1870 São Christovão	1810	59	186	78
1870 São Christovão	1820	143	203	74
1870 São Christovão	1830	239	182	79
1870 São Christovão	1840	250	204	74
1890 Brazil	1830	586,793	235	66
1890 Brazil	1840	1,021,027	226	69
1890 Brazil	1850	1,605,498	205	74
1890 Brazil	1860	2,323,131	197	76
1920 Brazil	1860	102,312	181	80
1920 Brazil	1870	180,316	179	80
1920 Brazil	1880	279,862	156	86
1920 Brazil	1890	447,566	141	90
1950 Brazil	1890	2,221,106	164	84
1950 Brazil	1900	3,844,441	157	86
1950 Brazil	1910	5,774,083	141	90
1950 Brazil	1920	8,143,411	131	92
1777 Cartagena	1710	269	257	61
1777 Cartagena	1720	476	251	62
1777 Cartagena	1730	667	207	73
1777 Cartagena	1740	1,019	275	56
1777 Colombia	1710	168	265	59
1777 Colombia	1720	254	289	53
1777 Colombia	1730	398	241	65
1777 Colombia	1740	734	275	56
1870 Colombia	1810	254	260	60
1870 Colombia	1820	399	259	60
1870 Colombia	1830	557	251	62
1870 Colombia	1840	1,177	250	62
1928 Colombia	1880	85	288	53
1928 Colombia	1890	203	268	58
1928 Colombia	1900	279	223	69
1963 Colombia	1900	727,666	175	81
1963 Colombia	1910	1,178,083	160	85
1963 Colombia	1920	1,796,228	142	89
1963 Colombia	1930	2,356,068	132	92
1871 Ecuador	1800	1,531	280	55
1871 Ecuador	1810	3,053	287	53
1871 Ecuador	1820	4,800	275	56
1871 Ecuador	1830	8,049	254	61
1871 Ecuador	1840	11,718	238	66
1950 Ecuador	1880		290	53

1950 Ecuador	1890		256	61
1950 Ecuador	1900		234	67
1950 Ecuador	1910		195	76
1950 Ecuador	1920		163	84
1740-44 Hidalgo/Guanajuato/Oaxaca	1680	118	343	39
1740-44 Hidalgo/Guanajuato/Oaxaca	1690	189	299	50
1740-44 Hidalgo/Guanajuato/Oaxaca	1700	348	279	55
1740-44 Hidalgo/Guanajuato/Oaxaca	1710	573	255	61
1777 Mexico - Central	1710	550	281	55
1777 Mexico - Central	1720	777	260	60
1777 Mexico - Central	1730	1,437	218	70
1777 Mexico - Central	1740	1,615	242	64
1777 Mexico City	1710	53	255	61
1777 Mexico City	1720	86	285	54
1777 Mexico City	1730	170	274	57
1777 Mexico City	1740	299	250	63
1777 Mexico - North	1710	70	243	64
1777 Mexico - North	1720	125	244	64
1777 Mexico - North	1730	184	236	66
1777 Mexico - North	1740	326	259	60
1790 Mexico City	1730	341	305	49
1790 Mexico City	1740	755	280	55
1790 Mexico City	1750	1,271	241	65
1790 Mexico City	1760	1,845	245	64
1821 Guadalajara	1760	1,438	311	47
1821 Guadalajara	1770	3,017	312	47
1821 Guadalajara	1780	4,975	295	51
1821 Guadalajara	1790	7,195	294	51
1823 Coahuila	1760	150	253	62
1823 Coahuila	1770	307	202	75
1823 Coahuila	1780	483	198	76
1823 Coahuila	1790	658	246	63
1930 Mexico	1870	514	278	55
1930 Mexico	1880	908	256	61
1930 Mexico	1890	1,432	243	64
1930 Mexico	1900	2,113	226	68
1950 Mexico	1890	1,147,619	239	65
1950 Mexico	1900	2,028,193	220	70
1950 Mexico	1910	2,855,705	194	77
1950 Mexico	1920	3,902,717	167	83
1700 Lima	1640	274	374	31
1700 Lima	1650	515	354	36
1700 Lima	1660	896	298	51
1700 Lima	1670	1,112	299	50
1866-1909 Lima prison	1820		163	84
1866-1909 Lima prison	1830		161	85
1866-1909 Lima prison	1840		161	85
1866-1909 Lima prison	1850		128	93

|				
1866-1909 Lima prison	1860		115	96
1940 Peru	1880	287,100	208	73
1940 Peru	1890	441,660	194	76
1940 Peru	1900	698,569	172	82
1940 Peru	1910	942,837	148	88
1940 Lima	1880	36,162	139	90
1940 Lima	1890	62,904	132	92
1940 Lima	1900	104,355	127	93
1940 Lima	1910	149,334	119	95
1772 Montevideo	1710	119	261	60
1772 Montevideo	1720	151	265	59
1772 Montevideo	1730	371	236	66
1772 Montevideo	1740	721	232	67
1791 Montevideo	1740	52	260	60
1791 Montevideo	1750	112	281	55
1791 Montevideo	1760	132	248	63
1834-36 Soriano/Maldonado	1770	110	286	53
1834-36 Soriano/Maldonado	1780	245	218	70
1834-36 Soriano/Maldonado	1790	365	222	70
1834-36 Soriano/Maldonado	1800	446	257	61
1846 Montevideo prison	1790	130	231	67
1846 Montevideo prison	1800	470	145	89
1846 Montevideo prison	1810	965	172	82
1868 Montevideo prison	1830	130	185	79
1868 Montevideo prison	1840	345	178	80
1868 Montevideo prison	1850	793	154	86
1963 Uruguay	1900	234,244	104	99
1963 Uruguay	1910	297,306	109	98
1963 Uruguay	1920	373,428	105	99
1963 Uruguay	1930	385,341	106	99
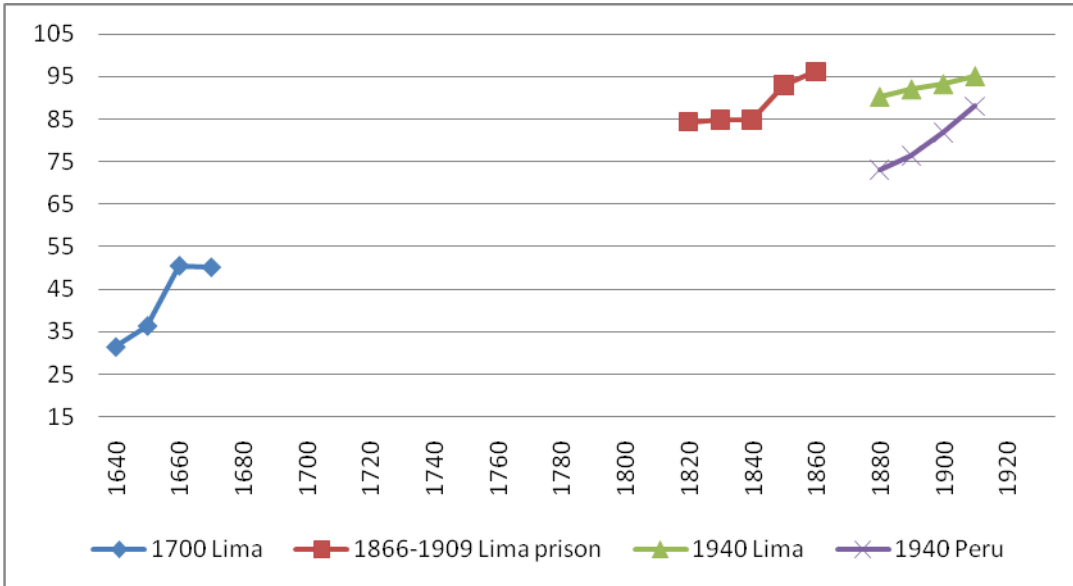
Peru - ABCC index of basic numeracy (vertical axis) by birth decades

Note: The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1.
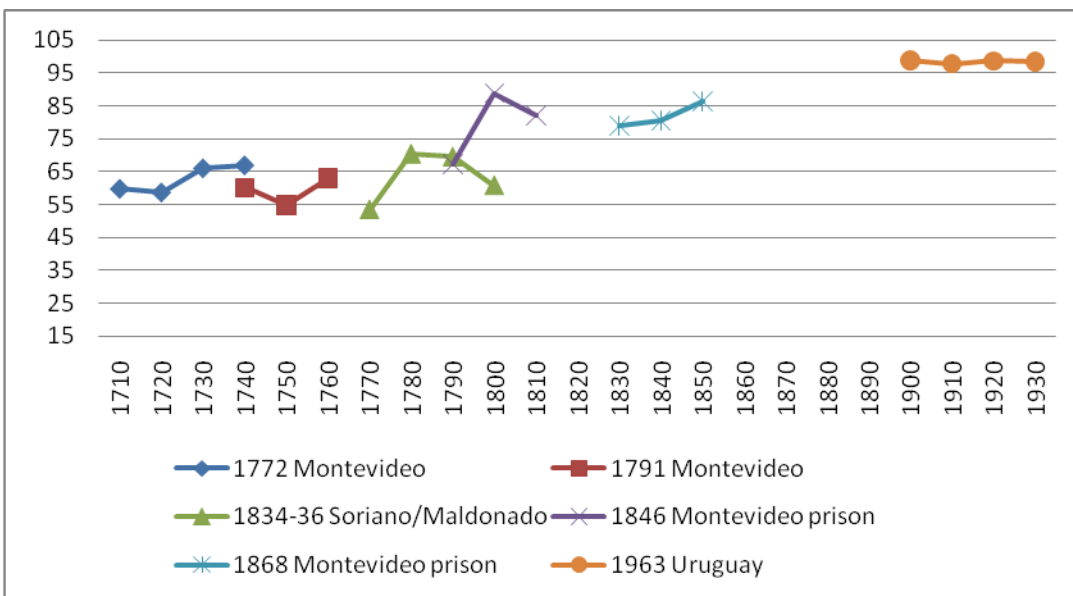
Uruguay - ABCC index of basic numeracy (vertical axis) by birth decades

Note: The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1.
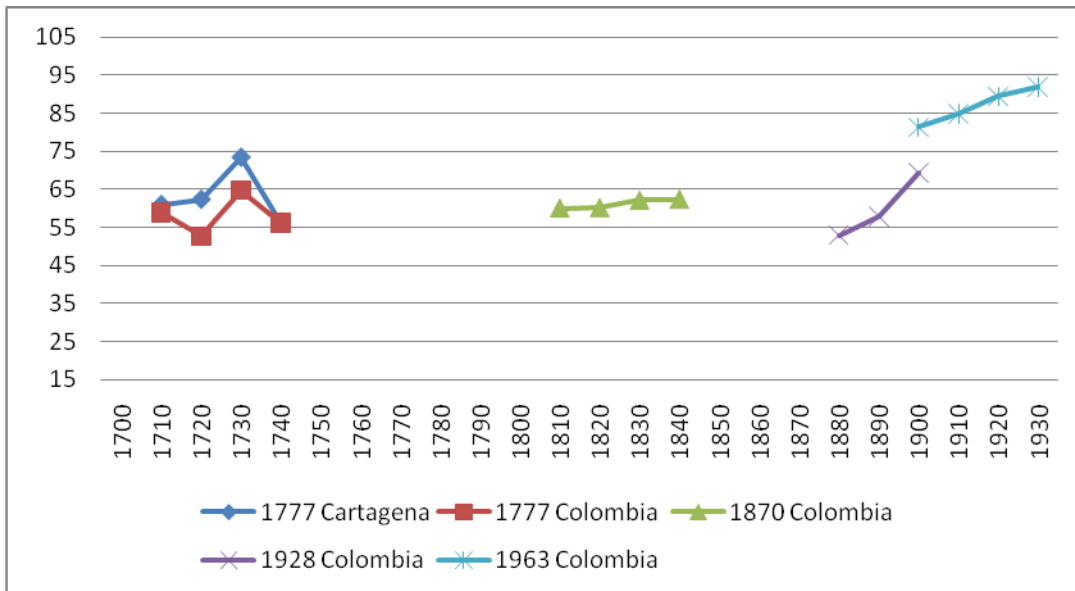
Colombia - ABCC index of basic numeracy (vertical axis) by birth decades

Note: The decades refer to birth decades (1800-1809, etc.). Sources: See text and Table 1.

Internet Appendix C: 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.[80] 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.[81]

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

---

[80] For more detailed surveys on the age-heaping methodology see A'Hearn, Baten, and Crayen (2009).

[81] 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.

number of indices, the Whipple index being considered the most reliable.[82] 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.[83]

$$(1) \quad Wh = \frac{\sum(Age_{25,30,35\dots70})}{\frac{1}{5}\sum(Age_{23,24,25\dots72})} \times 100$$

For an easier interpretation, A'Hearn, Baten, and Crayen (2009) suggested another index, which we call the ABCC index.[84] It is a simple linear transformation of the Whipple index and yields an estimate of the share of individuals who correctly report their age

$$(2) \quad ABCC = \left(1 - \frac{Wh-100}{400}\right) \times 100 \text{ if } Wh \geq 100; \text{ else } ABCC = 0.$$

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 2009). A'Hearn, Baten, and Crayen (2009) 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.[85] 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.

---

[82] A'Hearn, Baten, and Crayen (2009) 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.

[83] 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.

[84] 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.

[85] See A'Hearn, Baten, and Crayen (2009). Appendix available from the authors.

In addition, A'Hearn, Baten, and Crayen 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 (2009) 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 have 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 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.[86] 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.[87]

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.

---

[86] 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 2009).

[87] 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 (2009) to increase the Whipple value (minus 100) by 24% before calculating the ABCC value.