# GLOBALIZATION AND EDUCATIONAL INEQUALITY DURING THE $18^{\mathrm{TH}}$ TO $20^{\mathrm{TH}}$ CENTURIES: LATIN AMERICA IN GLOBAL COMPARISON * 

JOERG BATEN<br>University of Tuebingen and CESifo ${ }^{\text {a }}$

CHRISTINA MUMME

University of Tuebingen


#### Abstract

This paper explores the inequality of numeracy and education by studying school years and numeracy of the rich and poor, as well as of tall and short individuals. To estimate numeracy, the age-heaping method is used for the $18^{\text {th }}$ to early $20^{\text {th }}$ centuries. Testing the hypothesis that globalization might have increased the inequality of education, we find evidence that $19^{\text {th }}$ century globalization actually increased inequality in Latin America, but $20^{\text {th }}$ century globalization had positive effects by reducing educational inequality in a


[^0]broader sample of developing countries. Moreover, we find strong evidence for Kuznets's inverted U hypothesis, that is, rising educational inequality with GDP per capita in the period until 1913 and the opposite after 1945.

Keywords: human capital, inequality, age heaping, numeracy, globalization
JEL Code: I21, N30, N36, O57

## RESUMEN

En este artículo se estudia la inequidad en la alfabetización matemática -numeracy- y en la educación analizando los años de escolarización y alfabetización matemática de ricos y pobres, ya sean individuos altos o bajos. Para estimar la alfabetización matemática utilizamos el método age heaping para los siglos XVIII y XIX. Contrastando la hipótesis de que la globalización puede haber incrementado la inequidad educacional, encontramos evidencias de que la globalización del siglo XIX aumentó esta inequidad mientras la globalización del siglo XX tuvo, sin embargo, efectos positivos en la reducción de la inequidad educacional en una amplia muestra de países en desarrollo. Además, hemos encontrado evidencias que confirman la hipótesis de la U invertida de Kuznets: una relación positiva entre la inequidad educacional y el PBI per cápita en el período anterior a 1913 y lo opuesto en el período posterior a 1945.

Palabras clave: capital humano, desigualdad, age heaping, globalización

## 1. INTRODUCTION

Inequality is an important factor in today's globalization of the world economy and one of the underlying causes of income inequality - educational inequality - is at the core of the debate. Educational inequality is in many cases difficult to measure because micro surveys, which are not normally comparable across countries and periods, are necessary. This contribution uses the numeracy difference between occupational groups as a measure of inequality for the period from the $18^{\text {th }}$ to the early $20^{\text {th }}$ centuries, and the difference in years of schooling between the taller and shorter half of the female population as a measure for late $20^{\text {th }}$ century educational inequality. For the earlier period we concentrate on Latin America, a region where $20^{\text {th }}$ century income inequality is famously high, and for the latter period we compare the countries of this region with less developed countries (LDCs) and medium-income countries elsewhere.

The relationship between globalization and educational inequality is one of today's major issues. Is it possible that the current globalization will fail, just as the previous globalization tendency in the period 1850-1914 did, because inequality stimulates anti-integration forces? Timmer and Williamson (1998) found that during the $19^{\text {th }}$ century, inequality in new world countries such as the United States, Brazil, Argentina, Canada and Australia provoked anti-immigration policies that led to the disintegration of Atlantic labor markets. Rising inequality could also decrease the legitimation of international integration, so that those groups that normally benefit from it (e.g. the well educated in rich countries and world inhabitants in general) might not give it their full support.

There are other reasons why the study of educational inequality determinants is important. First, inequality is now often considered as a component of the standard of living. Being at the bottom of the income distribution is much harder to bear if the distance to the wealthier part of the economy is large, and educational inequality is a determinant of later income inequality. This also applies to the inequality of schooling (Castello and Domenech 2002; Thomas et al. 2001).

Wood (1997) argued that for the 1980s and 1990s more open trade increased wage inequality in some parts of the world, particularly in Latin America (Wood 1997; UN 1995; Cepal 2004). His studies have focused on the 1980s and 1990s. The question is whether the relationship of globalization and inequality holds before the 1980s. In our study, we are not restricted to only two decades. We are able to analyze the relationship for three centuries and will assess the difference between educational inequality in the «First Era of Globalization Period», as O'Rourke and Williamson (1999) defined it — namely, the 1850-1913 period with the early phases. For the second study period, 1945-1984, we will use the openness indices of Sachs and Warner (1995) as well as trade shares to assess whether openness increased educational inequality. Inequality, though, is a complex phenomenon and many potential determinants should be taken into account as well as globalization. We will therefore control for as many other potential determinants as possible.

Why should openness matter for educational inequality? Most research in this field has focused on income inequality in OECD countries, arguing that imports of goods mainly produced with unskilled labor could decrease the demand for unskilled labor within the rich OECD countries, depressing unskilled wages and increasing inequality. However, factor endowments and relative scarcities in developing countries (LDCs) differ fundamentally (Wood 1994, 1997, 1998).

Our expectation is that openness in fact increases inequality in countries with abundant land, and lowers inequality if unskilled labor, relative to potential trading partners, is the abundant factor. In the absence of unusual complementarities between factor inputs and other counteracting forces, poor countries will increase their exports of unskilled labor-intensive products during
globalization periods because their abundant factors and their comparative advantage are likely to be in this segment. Increasing production with unskilled labor should increase unskilled labor demand and wages, and the opposite should be true for land-rich countries. If labor demand rises (falls), even children of unskilled workers should receive some schooling (or less), although in most cases not enough to move into the upper half of income recipients.

Now, Latin America was clearly a land-rich region between the $18^{\text {th }}$ and early $20^{\text {th }}$ centuries (Prados de la Escosura 2007). Therefore, the expectation would be a rise in inequality during the «First Era of Globalization» (18501913; O'Rourke and Williamson 1999). Does the expectation also hold for the broader sample of developing countries during the 1945-1984 period? It is somewhat less clear as some were already industrializing during the period and land became less important.

A number of scholars have studied the influence of globalization and deglobalization on within-country income inequality worldwide (Lindert and Williamson 2001). Bourguignon and Morrison (2002) found a strong decline of within-country inequality during the de-globalization phase of 1914-1945, whereas within-country inequality rises during globalization phases. Prados de la Escosura (2007) finds increasing Latin American income inequality during this period. In section 5 , we will assess whether educational inequality also increases with growing openness.

## 2. YEARS OF SCHOOLING, AGE HEAPING, OCCUPATIONS AND INEQUALITY

We will first discuss the general idea of the age-heaping method in this section, before discussing our measures for inequality of human capital. Age heaping is often used nowadays as a basic numeracy indicator. The share of people who are able to report their exact age rather than a rounded age has been found to be strongly correlated with numerical abilities (Crayen and Baten 2010a). A widely used measure for age heaping is the ABCC index, as suggested by A'Hearn et al. (2009), which divides the number of people who reported non-preferred ages (i.e. ages that are not a multiple of five) by the total number of people ${ }^{1}$. The index ranges from 0 to 100 . If everybody reports the correct age, ABCC has a value of 100 . Here, we restrict the evidence to the age groups 23-32, 33-42, 43-52 and 53-62 years, because ABCCs of younger and older individuals might be biased. Only units that have at least thirty cases per skill and age group are studied. The ABCC index can best be understood by considering an example. If we have 100 people with unskilled occupations of a specific age group reporting age in the census of, say, 1870, we

[^1]would expect twenty of them to report an age ending in zero or five (because two of ten ages end in zero or five). For the remaining eighty individuals, the question is: do they report other ages or do they also choose an age ending in zero or five? If twenty do the latter, then one-quarter (of the eighty remaining persons) report probably a wrong age, and the $A B C C$ is then 75 per cent (one minus one-quarter). If the people with skilled occupations in the same country and birth decade have an ABCC of 85 per cent, then the social difference of numeracy between those occupational groups is ten (85-75 per cent) ${ }^{2}$.

It is important, however, to countercheck whether the census-takers or recruitment officers did explicitly ask for the age (and did not «correct» the reported ages afterward). In the case of the samples studied here, we have good reasons to believe that the people were actually asked for their age, and the number of corrections made afterward was not large. Otherwise, the relatively high level of age heaping that we observed in the data would probably not have occurred ${ }^{3}$.

How close is the relationship between age heaping and other human capital indicators such as literacy and schooling? A'Hearn et al. (2009) used the large US census sample to perform a very detailed analysis of this relationship. They subdivided by race, gender, high and low educational status and other criteria. In each case, they obtained a statistically significant relationship. The fact that the coefficients are relatively stable between samples is also noteworthy, that is, a unit change in age heaping is associated with similar changes in literacy across the various tests. The correlation was both statistically and economically significant for any country studied so far that had substantial age heaping ${ }^{4}$.

Some uncertainty remains about whether age heaping in the sources contains information about the numeracy of the responding individual or, rather, about the diligence of the reporting personnel who wrote down the statements. A potential bias always exists if more than one person is involved in the creation of a historical source. For example, if literacy is measured by analyzing the share of signatures in marriage contracts, there might have been priests who were more or less interested in obtaining real signatures, as opposed to just crosses or other symbols. We find it reinforcing that we estimate generally much more age heaping (and less numeracy) for the lower

[^2]social strata, and among the half of the sample population that had lower anthropometric values. Moreover, the regional differences of age heaping are similar to regional differences in illiteracy.

We conclude that the age-heaping method is now a well-established indicator for numeracy of groups, but the problem regarding how upper and lower group members can be distinguished from each other for historical populations for which we typically have no individual income data remains. Occupations have often been used to classify upper- vs. lower-income group individuals and we will apply this criterion to Latin American data until the $20^{\text {th }}$ century (similarly to Crayen and Baten 2010b). Of course, occupations such as «day laborer» or «agricultural worker» typically yielded a low income, whereas professionals, noblemen, factory owners and skilled craftsmen had higher incomes. As a caveat to this method, it should be noted that some occupations represent a wide income range (e.g. farmers).

For the study of the 1945-1984 period, we also use an alternative, similarly rough proxy to distinguish between social groups, based on human stature as Crayen and Baten (2010b) have suggested. This involves contrasting the number of years of schooling of the taller and shorter 50 per cent of the sample. Almost all anthropometric studies that considered occupational or income groupings found that the well-off strata of society were taller ${ }^{5}$. A second, very interesting aspect of this strategy is that tall individuals are much less likely than short individuals to have suffered from infant protein deficiency syndrome, which reduces learning abilities to a certain extent. The syndrome was widespread during the 1945-1984 period in the poorest countries of the world, when malnutrition was so common that most populations were severely stunted (with adult males being shorter than 170 cm on average). Support for this claim comes from biologists and psychologists who have conducted experiments on the influence of protein malnutrition in childhood and intellectual ability later in life (Paxson and Schady 2007).

One caveat to the proposed anthropometric method is clearly that there is also genetic height variation (especially on the individual level, see Magnusson et al. 2006). Nonetheless, we are confident that most individual variation can be averaged out by means of sufficiently large sample sizes.

Why do we use such a special method to measure inequality of education? Are there not other data sets available that contain some social classification criterion (such as occupation or income), as well as educational measures for the 1945-1985 period? To the best of our knowledge, for such a large number of countries as used here (forty-two countries), consistent data sets of this type do not exist ${ }^{6}$.

[^3]
## 3. DATA SOURCES, SELECTIVITIES AND REPRESENTATIVENESS

Many population counts were carried out in colonial Latin America, aiming at an overview of the population, taxpayers and the military potential. Most early counts were focused on limited regions or cities within a country. Larger censuses were carried out after the mid-18 ${ }^{\text {th }}$ century, covering a higher share of the national population (Table 1; Manzel and Baten 2010). For the post-colonial period, censuses of the republics were carried out mostly after the mid- $19^{\text {th }}$ century, while the early $19^{\text {th }}$ century is clearly less documented. The Latin American countries currently have the best source situation of historical population enumerations among today's developing countries (Platt 1998, p. 7). Our samples cover Argentina, Brazil, Colombia, Ecuador, Mexico, Uruguay and Venezuela ${ }^{7}$ and therefore represent a large part of this world region. All in all, the countries under study today represent around 80 per cent of the Latin American population. An important question is whether our various sources are representative of the whole society during the period under study. This issue has been studied intensively by Manzel and Baten (2010) who used mostly the same sources to study long-term trends. The population enumerations were supposed to have universal coverage in the whole area considered as well as in all social strata. Manzel and Baten have assessed many potential weaknesses of the data, such as social and regional biases. For example, one potential criticism of the padrones of the $18^{\text {th }}$ century is under-enumeration. The government wanted to know the population number and age structure in order to learn about the potential of taxpayers. One could imagine that this stimulated avoidance behavior among the richer part of the population. However, it was not easy for the rich and well-educated strata to avoid being included in the census. We find them in large numbers in our census lists, as is evident from the occupations listed.

Another potential caveat is the problem regarding who really answered the question about age. Is it possible that perhaps only the head of the household answered for the whole house? Manzel and Baten (2010) applied

[^4]TABLE 1
DATA SOURCES FOR THE EARLY PERIOD

Revista de Historia Económica, Journal of Iberian and Latin American Economic History

| Country/region | Year | Number of cases <br> (aged 23-62 years) | Potential bias relative <br> to total population | Source |
| :--- | :---: | :---: | :---: | :---: |
| Buenos Aires, AR | 1744 | 1,146 | Urban, military, <br> including slaves | Military Census in the Documentos para la <br> Historia Argentina (Caillet-Bois 1919) |
| Buenos Aires, AR | 1771 | 4,756 | Urban, including <br> slaves | Archivo Nacional de Argentina, Census <br> 1771 |
| Argentina | 1869 | 43,781 | No | Somoza and Lattes (1967) |
| Santa Fé, AR | 1887 | 808 | Regional | http://www.digitalmicrofilm.com.ar/ <br> censos/geografico.php |
| Argentina | 1895 | 51,715 | No | Somoza and Lattes (1967) |

TABLE 1 (Cont.)

| Central and West <br> Mexico | 1777 | 3,998 | Regional | AGI, Mex 2578/9. |
| :--- | :---: | :---: | :---: | :---: |
| Mexico City, MX | 1790 | 3,079 | Capital, only Spanish <br> and mestizo <br> household heads | Instituto Nacional de Estadística, <br> Geografía e Informática: CE <br> Revillagigedo (2003) |
| Coahuila, MX | 1823 | 1,598 | Regional | Grupo Exploradores Coahuiltecos |
| Mexico | 1930 | 7,007 | Various regions, but <br> nationally <br> representative | FSI: CE Guanajuato, Minas de Luz, <br> Mineral de los Llamitos, Ahualuco, <br> Benitez, Tepoztlán, Mezquital, Tetecala, <br> Tlaltizapan: MF 4107114, 4107751, <br> 4107265, 4107065 |
| Soriano/ <br> Maldonado | $1834-1836$ | 588 | Regional | Archivo Nacional de R.O. Uruguay, <br> CE Soriano/Maldonado |
| Montevideo, UY | 1846 | 1,569 | Capital, prisoners | Archivo Nacional de R.O. Uruguay |
| Cumarebo/ <br> Quisque, VE | $1818-1820$ | 1,476 | Regional | AGI: Cuba 759B |

AGI: Archivo General de Indias, AR: Argentina, BR: Brazil, CE: census of; FSI: family search indexing project, MF: microfilm; MX: Mexico, UY: Uruguay, VE: Venezuela.

Sources: adapted from Manzel and Baten (2010), plus other sources as explained in column «Source».
an indirect method by calculating the age-heaping indices for household heads and other members of the household. The expectation was that the head knew his exact age more often than the age of other household members. The difference between the two groups, however, was not very large. Other scattered evidence comes from remarks of the census officials about heads and other household members. In both cases, there were statements such as «she did not know her age», combined with an age statement of a preferred age. This can be interpreted as evidence that other household members were also actually questioned.

Ethnic composition is important for Latin America. Were Indios or slaves of African origin sometimes omitted from the padrones? The direct comparison of population structure by ethnic group given in the literature and the composition of $18^{\text {th }}$ century padrones ${ }^{8}$ revealed that the bias was limited Manzel and Baten (2010; Table 3). In some Mexican censuses, there was some under-representation of American Indios and in one of the Buenos Aires padrones there was some under-representation of African Americans, but in general the samples were quite representative in terms of ethnicity.

Regional bias is another issue that we need to address. Clearly, the early samples in particular were more often concentrated on the population of the capital (Tables 1 and 2). Large cities tend to have higher levels of inequality (Baten 1999), and hence we expect higher inequality values for the $18^{\text {th }}$ century. We will assess this effect with appropriate dummy variables below.

Finally, an important point for Latin America in particular is whether migrants should be included in the individual samples. Here, we are mainly interested in the educational inequality of countries and migrants who contributed to this inequality. Therefore, we decided to include migrants as well ${ }^{9}$.

While the sources for the study of Latin American educational inequality during the $18^{\text {th }}$ to early $20^{\text {th }}$ centuries deserve the most scrutiny, the later $20^{\text {th }}$ century evidence, which we are employing for our second study period, is easier to use. The Macro International Inc. performs surveys of child health and health-related behavior in order to create a solid and representative database for improving child health (among other aims). They recorded years of schooling and heights of women mostly born between 1945 and 1984 in many developing countries. We included only those aged $20-50$ years in many developing countries. As the height of adults is mostly determined in the 3 years after birth, the height of the mother can shed light on the development of status differences in this period after birth. One potential lacuna in our data is the environmental influence on growth at later ages, especially during the adolescent growth spurt. However,

[^5]TABLE 2
ABCC INDICES IN SEVERAL LATIN AMERICAN COUNTRIES BY BIRTH DECADE AND OCCUPATIONAL GROUPS

| Country | Birth decade | Unskilled | Skilled | Difference |
| :---: | :---: | :---: | :---: | :---: |
| Argentina | 1680 | 24 | 41 | 17 |
|  | 1690 | 24 | 43 | 19 |
|  | 1700 | 38 | 47 | 10 |
|  | 1710 | 44 | 58 | 14 |
|  | 1720 | 41 | 56 | 15 |
|  | 1730 | 51 | 59 | 8 |
|  | 1740 | 56 | 64 | 8 |
|  | 1810 | 63 | 77 | 14 |
|  | 1820 | 68 | 80 | 12 |
|  | 1830 | 71 | 84 | 14 |
|  | 1840 | 72 | 84 | 11 |
|  | 1850 | 77 | 89 | 12 |
|  | 1860 | 81 | 90 | 9 |
| Brazil | 1710 | 63 | 76 | 12 |
|  | 1720 | 63 | 63 | -1 |
|  | 1730 | 60 | 76 | 16 |
|  | 1740 | 53 | 67 | 15 |
|  | 1810 | 72 | 92 | 20 |
|  | 1820 | 79 | 88 | 9 |
|  | 1830 | 70 | 83 | 13 |
|  | 1840 | 60 | 82 | 22 |
| Colombia | 1830 | 56 | 74 | 17 |
|  | 1840 | 55 | 65 | 10 |
| Ecuador | 1810 | 58 | 63 | 4 |
|  | 1820 | 62 | 68 | 7 |
|  | 1830 | 64 | 68 | 4 |
|  | 1840 | 60 | 68 | 8 |
| Mexico | 1730 | 56 | 63 | 6 |
|  | 1740 | 66 | 61 | -5 |
|  | 1750 | 70 | 70 | 0 |
|  | 1760 | 70 | 75 | 5 |
|  | 1880 | 61 | 78 | 16 |
|  | 1890 | 62 | 85 | 23 |
|  | 1900 | 72 | 75 | 4 |

TABLE 2 (Cont.)

| Country | Birth decade | Unskilled | Skilled | Difference |
| :--- | :---: | :---: | :---: | :---: |
| Uruguay | 1780 | 55 | 71 | 16 |
|  | 1790 | 62 | 75 | 12 |
|  | 1800 | 79 | 85 | 6 |
|  | 1810 | 83 | 83 | 0 |
| Venezuela | 1780 | 55 | 71 | 16 |
|  | 1790 | 62 | 75 | 12 |

Notes: «Skilled» refers to occupational groups that were skilled, or professionals. «Unskilled» refers to those with unskilled or only semi-skilled occupations.

Baten (2000) finds that this effect is negligible compared to the impact of the first 3 years, as long as individuals have reached their final height when measured. Second, there could be survivor bias effects, but Moradi and Baten (2005) and Guntupalli and Baten (2006) rejected this possibility in detailed studies ${ }^{10}$.

We consider here the difference in years of schooling of the taller 50 per cent compared with the shorter half, and organize the data by individual country and birth decade. It is remarkable that in most cases taller women had more years of schooling (Table 3). Small differences refer to cases such as Ghana, Madagascar or Tanzania, in which the urban centers of education differed from the regions of tallest heights, which were sometimes characterized by specialization in cattle farming (Moradi and Baten 2005). In Latin America, some of the strong educational inequalities by height group are partly determined by the Indio vs. European ancestor difference. It is difficult to disentangle socio-economic differences from nutritional habit differences (and perhaps genetic ones) here. Hence, we will rely on fixed effect regressions below, which control for country-specific characteristics.

## 4. DEVELOPMENT OF EDUCATIONAL INEQUALITY

We would expect levels of educational inequality in the various countries of Latin America to be quite different because the institutional and economic

[^6]TABLE 3
DIFFERENCES OF SCHOOL YEARS BY HEIGHT (BIRTH YEARS 1945-1984)

| Country | Difference of school years | School years (tall) | School years (short) | Height (tall) | Height (short) |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Bangladesh | 0.9 | 3.3 | 2.5 | 1547 | 1460 |
| Burkina Faso | 0.4 | 1.2 | 0.7 | 1663 | 1568 |
| Benin | 0.8 | 2.1 | 1.4 | 1636 | 1537 |
| Bolivia | 1.9 | 7.1 | 5.2 | 1557 | 1467 |
| Brazil | 1.4 | 6.3 | 4.9 | 1607 | 1508 |
| Central African Republic | 0.7 | 2.2 | 1.5 | 1639 | 1533 |
| Ivory Coast | 0.5 | 2.5 | 2.0 | 1641 | 1545 |
| Cameroon | 0.5 | 5.5 | 4.9 | 1651 | 1553 |
| Colombia | 1.3 | 7.5 | 6.2 | 1592 | 1496 |
| Dominican Republic | 0.9 | 7.3 | 6.5 | 1614 | 1516 |
| Egypt | 1.2 | 5.5 | 4.3 | 1621 | 1532 |
| Ethiopia | 0.3 | 1.7 | 1.4 | 1619 | 1521 |
| Gabon | 0.7 | 6.2 | 5.6 | 1631 | 1533 |
| Ghana | 0.3 | 4.9 | 4.6 | 1639 | 1541 |
| Guinea | 0.4 | 1.2 | 0.8 | 1638 | 1541 |
| Guatemala | 1.6 | 2.9 | 1.3 | 1518 | 1423 |
| Haiti | 0.9 | 3.5 | 2.6 | 1632 | 1532 |
| India | 0.9 | 4.2 | 3.3 | 1561 | 1470 |

TABLE 3 (Cont.)

| Country | Difference of school years | School years (tall) | School years (short) | Height (tall) | Height (short) |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Kenya | 0.7 | 6.2 | 5.5 | 1646 | 1544 |
| Kyrgyztan | 0.6 | 11.3 | 10.7 | 1628 | 1538 |
| Comoros | 0.7 | 2.6 | 1.9 | 1592 | 1504 |
| Kazakhstan | 0.5 | 11.2 | 10.7 | 1637 | 1541 |
| Morocco | 1.1 | 3.5 | 2.4 | 1630 | 1539 |
| Madagascar | 0.1 | 3.4 | 3.3 | 1578 | 1487 |
| Mali | 0.5 | 1.2 | 0.8 | 1664 | 1568 |
| Malawi | 0.9 | 3.9 | 3.0 | 1607 | 1515 |
| Mozambique | 0.8 | 2.6 | 1.8 | 1609 | 1512 |
| Namibia | 0.7 | 5.4 | 4.7 | 1660 | 1561 |
| Niger | 0.4 | 0.9 | 0.5 | 1654 | 1560 |
| Nigeria | 1.3 | 5.0 | 3.7 | 1644 | 1529 |
| Nicaragua | 1.5 | 6.0 | 4.6 | 1586 | 1492 |
| Peru | 1.9 | 8.0 | 6.0 | 1549 | 1461 |
| Ruanda | 1.0 | 4.2 | 3.2 | 1631 | 1530 |
| Senegal | 0.4 | 1.3 | 0.9 | 1670 | 1574 |
| Chad | 0.5 | 1.3 | 0.7 | 1678 | 1578 |
| Togo | 0.4 | 1.6 | 1.2 | 1638 | 1542 |
| Turkey | 0.8 | 4.7 | 3.9 | 1600 | 1511 |

TABLE 3 (Cont.)

| Tanzania | 0.2 | 3.5 | 3.2 | 1606 | 1508 |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Uganda | 0.5 | 4.6 | 4.0 | 1635 | 1534 |
| Uzbekistan | 0.4 | 11.0 | 10.6 | 1649 | 1553 |
| Zambia | 1.0 | 5.4 | 4.4 | 1628 | 1531 |
| Zimbabwe | 0.8 | 7.0 | 6.1 | 1649 | 1552 |

Notes: «Tall» is defined here as the tallest $50 \%$; «short» as the shortest $50 \%$. The difference in school years is the number of school years of the taller minus the shorter $50 \%$. Female height is reported in millimetres.
structures were so varied. For example, Mesoamerica and the Andes had large shares of indigenous or mestizo populations, who received less schooling and other public goods compared with the middle and upper strata of European origin. In these regions, the inequality heritage of land distribution in favor of the Spanish conquerors and later European immigrants might have been strongest (Lambert 1968, p. 581)) Similarly, strong differences might have prevailed in the countries that kept slavery until the late $19^{\text {th }}$ century and whose population component of African origin was disadvantaged. In contrast, the population of the Southern Cone was more homogeneous in ethnicity-related aspects, because the Indio population share was smaller and slavery was abolished earlier. However, ethnicity, slavery and colonial heritage were not the only factors at work. O'Rourke and Williamson (1999) argued convincingly that the Southern Cone countries had strongly increasing inequality during the late $19^{\text {th }}$ century globalization movement.

Going further back in time, how might inequality have differed between colonial times and post-independence Latin America? Unfortunately, today we have little evidence for the pre-independence inequality history of Latin America. Williamson (2009) and Dobado and Garcia (2009) have recently raised some doubts about the early colonial heritage hypothesis (of continuously high inequality). Dobado and Garcia argued that real wages were quite high in some parts of Bourbon Latin America, whereas average income was lower than in Europe. Hence, inequality might actually have been lower than in Europe. The question is, of course, whether this wage evidence is representative and can inform us about the situation of other poorer strata, which did not earn wages (such as the population majority of peasants). Williamson (2009) considered the fact that especially the low population density of the $17^{\text {th }}$ and early $18^{\text {th }}$ centuries might have generated relatively low inequality, compared with Europe. In times of labor scarcity, wages tend to be higher and even the nutrition and general treatment of slaves and indigenous bound labor might be slightly less terrible. Recent study on Uruguay suggests that in the Southern Cone during the $18^{\text {th }}$ century, inequality might also not have been very pronounced (Vicario 2010).

Economic policy differences certainly also mattered for country-specific differences. For example, the famous Mexican dictatorship era of Porfirio Díaz (1877-1911) increased average income and education, but had a reputation for bringing about sharp increases in inequality (Tutino 2001, pp. 700701). In sum, during the $19^{\text {th }}$ century the variation of inequality between countries was probably large. Pre-independence inequality history is largely unexplored, but some authors have recently argued for a modest inequality level relative to Europe.

In the following, we first compare numeracy of the upper and lower occupational strata in Latin American countries for the early period. For Argentina, we have some data for Buenos Aires for the birth decades until 1740 and some representative national data after this date. The ABCC index
increased from 24 to 56 for the lower-income groups in Buenos Aires from the late $17^{\text {th }}$ to the mid- $18^{\text {th }}$ centuries (Table 2). The upper-income groups started with a level above 40 per cent age numeracy in the 1680s, but grew to just 64 per cent in the 1740s. The fact that the early evidence on Argentina covers only Buenos Aires is certainly a caveat. The gap for the $19^{\text {th }}$ century was large and relatively constant, declining only slightly from a 14 per cent difference in the 1820 s to a 9 per cent difference in the 1860s.

Second, we have long-term data on Mexico covering sufficient observations to study both social groups between the 1730 s and 1760 s, and between the 1880 s and 1900s. The latter evidence is nationally representative, while the former relates to some Mexican regions, including Mexico City. In short, social differences in age numeracy were small and in one case even negative in $18^{\text {th }}$ century Mexico. This is consistent with the observation of Tanck de Estrada (1999) that the Bourbon reforms of the $18^{\text {th }}$ century had some positive impact even on school building in Indio villages, that is, for the poorest Mexicans. However, the late $19^{\text {th }}$ century saw a highly stratified society with large differences. This fits with available evidence on the Porfiriato.

We have more scattered data for a number of other Latin American countries (Table 2). In Brazil, the highest educational inequality levels were reached in the 1840s, which are the latest values we have ${ }^{11}$. In Uruguay, inequality was declining, but this might have been caused by the fact that the later birth cohorts were based on data from prisons. Venezuela and Colombia had quite high inequality, whereas in poor Ecuador even the skilled groups had low levels of numeracy. Summing up, we can document a number of Latin American countries, but the cases for which we have sufficient numbers of observations are somewhat distributed over the different centuries. Most striking in Table 2 is the fact that, of the forty cases studied, only two have negative values for the difference between the skilled and unskilled groups.

Now we move to the birth cohorts of the mid-to-late $20^{\text {th }}$ century. We studied the difference of school years, subtracting the figure for the taller half from that of the shorter half (Figure 1). The years of schooling are a better measure for the mid-to-late $20^{\text {th }}$ century than age heaping, which had already disappeared in many countries by then. It should be noted that there are many African countries in the sample, a few Latin American ones, but fewer countries from other regions. One important result from these samples is that the taller half of the population (those who probably came from more advantaged family backgrounds) always had higher or equal school year values, which tended to be lower among the shorter half of the population ${ }^{12}$. Latin America

[^7]FIGURE 1
SCHOOL-YEAR DIFFERENCE (VERTICAL AXIS) IN SEVERAL WORLD REGIONS BY BIRTH YEARS


Notes: Difference in years of schooling, value of tallest 50 per cent minus the shortest 50 per cent (women). Years refer to the beginning year of a 5 -year birth cohort. Central Asia includes only the former Soviet Republics.
had the highest difference (which means the largest educational inequality), whereas Soviet Central Asia had the lowest value (Figure 1).

## 5. OPENNESS AND OTHER POTENTIAL DETERMINANTS OF INEQUALITY

We first describe the explanatory variables in Table 4 that we include in the regressions and then discuss the results.

### 5.1. Openness

How can we measure «openness»? Given the importance of this variable, much study has been done in this field. Most economists agree that simple trade shares of GDP are insufficient by themselves to capture the degree of openness of an economy. If two neighboring free-trade countries have exactly the same factor endowments, it is possible that their trade is relatively low in spite of their openness simply because production is so similar. On the other hand, two highly protected countries can experience high trade shares if their endowments are sufficiently different. One alternative measure is the openness index of Sachs

| Estimation method | FE | FE | FE | FE | FE |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Birth decades | $\mathbf{1 6 8 0 s - 1 9 0 0 s}$ | $\mathbf{1 6 8 0 s - 1 9 0 0 s}$ | $\mathbf{1 9 4 0 s - 1 9 8 0 s}$ | $\mathbf{1 9 4 0 s - 1 9 8 0 s}$ | $\mathbf{1 9 4 0 s - 1 9 8 0 s}$ |
| World regions | Latin America | Latin America | LDCs | LDCs | LDCs |
| Openness concept | Era of Glob. | Era of Glob. | Sachs and Warner <br> $(1995)$ | Sachs and Warner <br> $(1995)$ | Penn WT |
|  |  |  | $-0.35^{* * *}$ | $-0.41^{* *}$ | 0.01 |
| Openness | $7.63^{* *}$ | $(0.041)$ | $(0.042)$ | $(0.004)$ | $(0.040)$ |

TABLE 4 (Cont.)

| Estimation method | FE | FE | FE | FE | FE |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Birth decades | $\mathbf{1 6 8 0 s} \mathbf{- 1 9 0 0 s}$ | $\mathbf{1 6 8 0 s}-1900 \mathbf{s}$ | $\mathbf{1 9 4 0 s}-\mathbf{1 9 8 0 s}$ | $\mathbf{1 9 4 0 s}-\mathbf{1 9 8 0 s}$ | $\mathbf{1 9 4 0 s} \mathbf{- 1 9 8 0 s}$ |
| Constant | $15.62 * *$ | -0.78 | $1.22 * * *$ | 1.54 | $2.06 * *$ |
|  | $(0.030)$ | $(0.93)$ | $(0.001)$ | $(0.20)$ | $(0.023)$ |
| Observations | 40 | 40 | 174 | 63 | 145 |
| $R^{2}$ | 0.18 | 0.36 | 0.35 | 0.62 | 0.33 |

FE: fixed effects, LDCs: less developed countries, Glob.: globalization, WT: World Tables, p.c.: per capita.
Sources: Maddison (2001). Where GDP was lacking, it was linearly interpolated. Ecuador was assumed to have had the average GDP per capita of Peru and Brazil. In columns 1 and 2, civil war data are collected from Clodfelter (2002), the share of mature was calculated from the age distributions in the censuses, see Table 1, and the same applies to the education average (using age heaping). For the explanatory variables in columns 3-5, see the Appendix (see footnote 1).

Notes: $P$-values are given in parentheses. *, ** and *** refer to significance levels of $1 \%, 5 \%$ and $10 \%$, respectively. Dependent variable in columns 1 and 2 : difference in numeracy, skilled vs. unskilled occupations. Dependent variable in columns 3-5: difference of school years, tallest $50 \%$ vs. shortest $50 \%$ (calculated as the difference between the two groups). As usual with fixed effects regressions, we reported the « $R^{2}$ within». The models were also estimated including a prison dummy, but there was almost no difference in the other coefficients. We also checked regressions with a capital city dummy, with almost no change in the other coefficients. In columns 3-5, we included time-fixed effects. GDP per capita is expressed in units of $\$ 1000$ (Geary Khamis $\$$ ).
and Warner (1995) for seventy-nine countries. These authors consider high tariffs, important tariff barriers, plus state monopolies of major commodity exports, a high black market premium for national currencies and a socialist economic system. This variable is coded as a binary variable. Rodriguez and Rodrik (1999) criticized the fact that two factors in particular, the state monopoly and currency black market premium, might measure other economic characteristics rather than just a lack of openness. The currency value distortion also indicates other macroeconomic problems. However, if there is no perfect measure of openness, it is a promising strategy to use both this one and the trade share. In spite of our conceptual skepticism against the trade share of GDP as a measure of «openness», we will also test this variable below.

### 5.2. Kuznets Curve Effects

Kuznets (1955) found that inequality first rises and then declines with economic development. He explained his inverted U curve by labor-market disequilibria: since technological progress initially favors the rewards for some specialized skills, demand for unskilled labor decreases and its wage falls. Therefore, inequality rises initially. The diffusion of skills and economic policies serve as egalitarian forces that reduce inequality. We therefore add Kuznets variables by adding real GDP per worker in linear and quadratic forms and expect a positive coefficient of the former and a negative effect for the latter ${ }^{13}$.

### 5.3. Democracy

Li et al. (1998) also emphasized the importance of political freedom for income equality. If dictatorship provides privileges to certain groups in society, this might lead to higher inequality. The research project «POLITY IV», at the University of Maryland, created comprehensive surveys on the democratic or autocratic behavior of governments in recent history, approximating democracy with a numerical score.

### 5.4. Demographic Effects (Mature)

Did competition reduce the wages of baby boomers? Demographic effects could have an influence on inequality. According to the normal life cycle effect of income, people receive their highest income in their 40 s and 50 s.

[^8]Teenagers and young adults between age 15 and 40 years earn less on average, and beyond the age of 60 years income starts to decline again. If «fat» cohorts (e.g. the «baby boom» generation of the 1960s) enter the labor market, we would expect a rise in inequality because the supply of «young» labor is very large, whereas the share of the richer «mature» age group is relatively smaller. Higgins and Williamson (1999) found a robust influence of both cohort sizes in the mature age groups. We include cohort size effects by taking the share of the mature population (aged 40-59 years), relative to the total population of the age groups 15-69 years (working age), using the same specification as Higgins and Williamson (1999).

### 5.5. Speed of Structural Change

How much did agricultural productivity lag? Agricultural productivity, and therefore agricultural incomes, might lag behind industry and services and this could lead to rising inequality (Baten and Fraunholz 2004).

### 5.6. Civil War

Civil war is one of the strongest determinants of welfare and educational development in developing countries. Civil war has a very destructive effect on average schooling levels, but it is less clear whether this terrible military nightmare increases or decreases the inequality of schooling. In some cases, the better-off population might be able to flee to quiet parts of the country and their children might continue attending school. On the other hand, the destruction of expected human capital returns in the future might particularly affect those strata that otherwise would have invested a lot in the schooling of their children. Hence, it is an empirical question whether this variable increased or decreased educational inequality ${ }^{14}$.

### 5.7. Results for the Early Period

All regression models are estimated as fixed effects in order to control for unobservable characteristics, such as cultural or geographic factors. In the regressions for the early period, we employed a dummy variable for the «First Era of Globalization» (1850-1913) as an indicator for openness. For the $18^{\text {th }}$ to early $20^{\text {th }}$ centuries, there is insufficient evidence to reconstruct the trade share or political protectionism, except perhaps for the last few decades of our study period (and for this period alone the number of

[^9]observations would be too small). We find the regressions to have a positive coefficient of the first era of globalization dummy variable (Table 4).

The absolute level of numeracy might reduce the inequalities in this period slightly (only significant in column 2 ). The «baby boom effect» of the «mature» variable is not visible in this period. Civil war did not have strong and significant effects on early inequality.

In the second regression, we also included GDP per capita as well as its squared term. The value of the former was large and positive while the latter had a large negative value. The former term is larger than the latter and has a greater effect on the observed values ${ }^{15}$. Hence, for this early period the Kuznets curve was on the rise.

### 5.8. Results for the $\mathbf{1 9 4 5 - 1 9 8 4}$ Period

We compare three regression models for the later period (Table 4, columns $3-5)$. Openness actually reduced educational inequality in this period. The coefficients of Sachs and Warner (1995) openness are statistically significant, although the values of the coefficients are not very large. In column 5 of Table 4, we also included an alternative measure of openness, the trade share as reported in the Penn World Tables. However, this alternative measure is not significant (after adjustment for population size, it remains insignificant).

In contrast, GDP per capita and its squared term are significant, except in column 4, which records a much smaller number of cases. The coefficient of the non-squared GDP term is much larger than the squared term. This implies that educational inequality declines with increasing income during this period. At very high levels, the decline stops. The results are also quite robust over different specifications.

The level of average education increases the gap in less-developed countries, which is quite the contrary to what we might have expected. In contrast, civil war mostly reduces educational inequality - it appears that the richer strata do not send their children to school during a civil war either. Finally, there are no obvious effects of democracy, productivity lags or the «baby boom» effect (i.e. the «mature» variable). The explanatory power of these models is in general quite large.

Is openness endogenous here? The question is whether the lower educational inequality of children born in a specific birth decade would cause more openness. One could imagine that - based on the general StolperSamuelson view - labor-abundant countries with high inequality would open their economies during a globalization period to profit from more demand and therefore higher wages in that sector or vice versa. This might at least be the case if unskilled workers have sufficient political power. On the

[^10]other hand, especially among richer countries, there might be an economic or psychological effect of higher inequality leading to less openness. Baltzer and Baten (2008) tested these hypotheses and provided evidence that low inequality in Latin American countries in the mid-to-late $20^{\text {th }}$ century did not lead to more openness. Hence, we tentatively conclude that endogeneity is not a major problem here.

## 6. CONCLUSION

We explored inequality of numeracy and education by studying school years and numeracy of the rich and poor, as well as of tall and short individuals. To estimate numeracy, the age-heaping method was used. In this study, we mobilized a large body of new evidence on inequality, going back to the $18^{\text {th }}$ century and covering a number of Latin American countries, namely Argentina, Brazil, Colombia, Ecuador, Mexico, Uruguay and Venezuela. Looking at the time trend of educational inequality, Mexico displays only modest numeracy advantage for the skilled groups in the $18^{\text {th }}$ century, but the gaps between the upper and lower strata increased strongly until the $19^{\text {th }}$ century. Similarly, Argentina suffered substantial educational inequality during the $19^{\text {th }}$ century. In a regression analysis, the «First Era of Globalization» was mostly confirmed as having higher inequality than earlier periods.

We studied many developing countries in the period from the 1940s to the 1980s, looking at the schooling difference between the taller half of the population and the shorter half. One remarkable finding was that the taller half always had more years of schooling. This applied to forty-two different countries without exception. Latin America had the greatest educational inequality in this period, which is certainly one of the reasons for its high income inequality today.

Testing the hypothesis that globalization might have increased inequality of education, we found evidence that $20^{\text {th }}$ century globalization had positive effects by reducing educational inequality. Moreover, we found strong evidence for Kuznets's inverted U hypothesis, which was on the rise during the $18^{\text {th }}$ and $19^{\text {th }}$ centuries in Latin America and tended to fall in the second half of the $20^{\text {th }}$ century in the developing world.

## REFERENCES

A’Hearn, B.; Crayen, D., and Baten, J. (2009): «Quantifying Quantitative Literacy: Age heaping and the History of Human Capital». Journal of Economic History 68 (3), pp. 783-808.
Baltzer, M., and Baten, J. (2008): «Heights, Inequality, and Trade in the Latin American Periphery, 1950-2004». Economics and Human Biology 6 (2), pp. 191-203.
Baten, J. (1999): Ernährung und wirtschaftliche Entwicklung in Bayern, 1730-1880. Stuttgart: Steiner.

Baten, J. (2000): «Height and Real Wages: An International Comparison». Jahrbuch fuer Wirtschaftsgeschichte 2000 (1), pp. 17-32.
Baten, J., and Fraunholz, U. (2004): «Did Partial Globalization Increase Inequality? The Case of the Latin American Periphery, 1950-2000». CESifo Economic Studies 50 (1), pp. 45-84.
Baten, J., and Komlos, J. (2004): «Looking Backward and Looking Forward: Anthropometric Research and the Development of Social Science History». Social Science History 28 (2), pp. 1-24.
Baten, J.; Pelger, I., and Twrdek, L. (2009): «The Anthropometric History of Brazil, Lima (Peru), and Argentina during the $19^{\text {th }}-20^{\text {th }}$ Centuries». Economics and Human Biology 7 (4), pp. 319-333.
Caillet-Bois, R. R. (1919): Documentos Para la Historia Argentina, vol. 11: Territorio y Poblacion. Padrones de las Ciudades y Campanas de Buenos Aires y Montevideo. Buenos Aires: Compania Sud-Americana de Billetes de Banco.
Castello, A., and Domenech, R. (2002): «Human Capital Inequality and Economic Growth: Some New Evidence». Economic Journal 112, pp. C187-C200.
CEPAL (2004): Social Panorama of Latin America. Santiago: United Nations Publication.
Clodfelter, M. (2002): Warfare and Armed Conflicts. A Statistical Reference to Casualty and Other Figures, 1500-2000. Jefferson, North Carolina and London: McFarland\& Company, Inc.
Crayen, D., and Baten, J. (2010a): «Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth». Explorations in Economic History 47 (1), pp. 82-99.
Crayen, D., and Baten, J. (2010b): «New Evidence and New Methods to Measure Human Capital Inequality before and during the Industrial Revolution: France and the U.S. in the Seventeenth to Nineteenth Centuries». Economic History Review 53 (2), pp. 452-478.
Dobado, R., and Garcia, H. (2009): «Neither So Low Nor So Short! Wages and Heights in Eighteenth and Early Nineteenth Centuries Colonial Hispanic America». Working Paper of the Universidad Complutense, Instituto Complutense de Estudios Internacionales Madrid 14-09.
Frankema, E. (2008): «The Historical Evolution of Inequality in Latin America: A Comparative Analysis, 1870-2000». Groningen: Rijksuniversiteit, Ph.D. dissertation.
Guntupalli, A., and Baten J. (2006): «The Development and Inequality of Heights in North, West and East India, 1915-44». Explorations in Economic History 43 (4), pp. 578-608.
Higgins, M., and Williamson, J. G. (1999): «Explaining Inequality the World Round: Cohort Size, Kuznets Curves, and Openness». NBER Working Paper no. 7224.
Kuznets, S. (1955): «Economic Growth and Income Inequality». American Economic Review 45 (1), pp. 1-28.
Lambert, D. C. (1968): «La Reparticion Des Los Ingresos y Las Desigualdades Sociales en la America Latina». Revista Mexicana de Sociologia 30 (3), pp. 571-592.
Li, H.; Souire, L., and Zou, H. (1998): «Explaining International and Intertemporal Variations in Income Inequality». Economic Journal 108, pp. 26-43.
Lindert, P., and Williamson, J. G. (2001): «Does Globalization Make the World More Unequal?». NBER Working Paper no. 8228.
Maddison, A. (2001): The World Economy: A Millennial Perspective. Paris: OECD.
Magnusson, P.; Rasmussen, F., and Gyllenstein, U. (2006): «Height at Age 18 Years Is a Strong Predictor of Attained Education Later in Life: Cohort Study of Over 950,000 Swedish Men». International Journal of Epidemiology 35, pp. 658-663.

Manzel, K., and Baten, J. (2009): «Gender Equality and Inequality in Numeracy - the Case of Latin America and the Caribbean, 1880-1949». Revista de Historia Económica - Journal of Latin American and Iberian Economic History 27 (1), pp. 37-73.
Manzel, K., and Baten, J. (2010): «Convergence and Divergence of Numeracy: The Development of Age Heaping in Latin America, 17th to 20th Century». Working Paper. University of Tübingen.
Moradi, A. (2002): «Height and Health of Women in Sub Saharan Africa and South Asia 1950-80». Conference Proceedings of the International Economic History Association Conference, Buenos Aires (on CDROM).
Moradi, A., and Baten, J. (2005): «Inequality in Sub-Saharan Africa: New Data and New Insights from Anthropometric Estimates». World Development 33 (8), pp. 1233-1265.
Morrison, C., and Murtin, F. (2007): «Education Inequalities and the Kuznets Curves: a Global Perspective since 1870». Paris School of Economics Working Paper no. 2007-12.
O'Rourke, K., and Williamson, J. G. (1999): Globalization and History. Cambridge: Cambridge University Press.
Paxson, C., and Schady, N. (2007): «Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health and Parenting». Journal of Human Resources 42 (1), pp. 49-84.
Platt, L. D. (1998): Census Records for Latin America and the Hispanic United States. Baltimore, M.D.: Genealogical Publishing Com.
Prados de la Escosura, L. (2007): «Inequality and Poverty in Latin America: a Long-Run Exploration», in T. J. Hatton, K. H. O'Rourke and A. M. Taylor (eds), New Comparative Economic History. Cambridge, Mass: MIT Press, pp. 291-315.
Prados de la Escosura, L. (2008): «Inequality, Poverty and the Kuznets Curve in Spain, 1850-2000». European Review of Economic History 12, pp. 287-324.
Rodriguez, F., and Rodrik, D. (1999): «Trade Policy and Economic Growth: A Sceptic's Guide to Cross-National Evidence». NBER Working Paper no. 7081.
Sachs, J., and Warner, A. (1995): «Economic Reform and the Process of Global Integration». Brookings Papers on Economic Activity 1995 (1), pp. 1-118.
Somoza, J., and Lattes, A. (1967): «Muestras de los dos Primeros Censos Nacionales de Población, 1869 y 1895». Documento de Trabajo no. 46, Instituto T. Di Tella, CIS, Buenos Aires.
Steckel, R., and Floud, R. (eds) (1997): Health and Welfare During Industrialization. Chicago and London: Chicago University Press.
Stolz, Y.; Baten, J., and Botelho, T. (2010): «Growth Effects of Brazilian Immigration». Working Paper Tuebingen/Belo Horizonte.
Tanck de Estrada, D. (1999): Pueblos de Indios y Educación en el México Colonial, 1750-1821. Mexico City: El Colegio de México.
Thomas, V.; Wang, Y., and Fan, X. (2001): «Measuring Education Inequality: Gini Coefficients of Education, 1960-2000». World Bank Research Paper no. 2525.
Timmer, A., and Williamson, J. G. (1998): «Immigration Policy Prior to the 1930s: Labor Markets, Policy Interactions, and Globalization Backlash». Population and Development Review 24, pp. 739-771.
Tutino, J. (2001): «Rural Economy and Society: 1821-1910», in M. S. Werner (ed.), Concise Encyclopedia of Mexico. Chicago, I.L.: Fitzroy Dearborn Publishers, pp. 697-705.
United Nations (1995): Economic Survey of Latin America and the Caribbean 1994-1995. Santiago: United Nations.

Van Zanden, J. L. (2009): «The Skill Premium and the Great Divergence». European Review of Economic History 13, pp. 121-153.
Vicario, C. (2010): «Estructura Social y Ocupacional en el Sur Atlántico. Montevideo 1760-1860». Université. de la Republica de Uruguay, Master thesis.
Williamson, J. (2009): «History without Evidence: Latin American Inequality since 1491». Inequality Conference Madrid, May 2009.
Wood, A. (1994): North-South Trade, Employment and Inequality. Oxford: Clarendon Press.
Wood, A. (1997): «Openness and Wage Inequality in Developing Countries: The Latin American Challenge to East Asian Conventional Wisdom». World Bank Economic Review 11, pp. 33-57.
Wood, A. (1998): «Globalisation and the Rise in Labour Market Inequalities». Economic Journal 108, pp. 1463-1482.


[^0]:    * Received 21 December 2009. Accepted 12 May 2010. We are indebted to the participants of the conference «A Comparative Approach to Inequality and Development: Latin America and Europe» in Madrid 2009 and seminar participants at Tuebingen for scientific discussions and advice. We thank the ESF-GlobalEuroNet and EU HIPOD projects for financial support. For important comments and data, we thank an anonymous referee, Lety Arroyo Abad, Pablo Astorga, Luis Bértola, Tarcísio Botelho, Stephen Broadberry, Albert Carreras, Rafael Dobado, Ewout Frankema, Thomas Ginter, Anton Glatz, Bishnu Gupta, Marianne Hock, Christina Jedermann, Dacil Juif, Kerstin Manzel, Adolfo Meisel, Alexandra Mylius, Sarah Nalle, Ricardo Paixao, Javier Puche Gil, Heike Schmutz, Bernd Stein, Yvonne Stolz, Linda Twrdek, Gerrit Ulrichs, Jan Luiten van Zanden and Carolina Vicario. Appendix: http://www.wiwi.uni-tuebingen.de/cms/fileadmin/Uploads/ Schulung/Schulung5/Paper/baten_mumme_appendix.pdf
    ${ }^{\text {a }}$ Department of Economics, Mohlstrasse 36, 72074 Tuebingen. joerg.baten@uni-tuebingen.de, $\underline{\text { mumme@gmx.net }}$

[^1]:    ${ }^{1}$ It is called ABCC after the authors' initials and Gregory Clark who gave comments on this index.

[^2]:    ${ }^{2}$ This holds only if age distributions are relatively smooth. Crayen and Baten (2010a) studied the influence of famines, epidemics, wars and civil wars and found that the effect was randomly distributed and in the vast majority of cases not influential for individual age groups. The method also assumes that ages ending in zero and five are the most clearly preferred ages. This is least clear for the age group 2332 years because heaping also frequently takes place in multiples of two. Crayen and Baten (2010a) suggest reducing the ABCC for this age group, a recommendation that we also follow.
    ${ }^{3}$ Even if the precise birthday (often related to a saint's day or a holiday) is known to the individual, it might well be the case that the exact number of years since birth means little to an individual, even though the annual event is celebrated again and again.
    ${ }^{4}$ On the regions of Argentina see, for example, Manzel and Baten (2009).

[^3]:    ${ }^{5}$ For recent collections of anthropometric studies, see Steckel and Floud (1997) and Baten and Komlos (2004).
    ${ }^{6}$ Another interesting measure of educational inequality was proposed by Frankema (2008) who uses a «comparative grade enrolment distribution» to determine educational inequality. His idea is

[^4]:    (F'note continued)
    that the higher the secondary school completion shares of the attainment distribution, the larger the educational «middle class», which might imply less educational inequality. Frankema (2008) finds that grade distribution in Latin America is skewed toward lower grades during the mid-to-late $20^{\text {th }}$ century with almost 43 per cent of the pupils leaving school without passing the first grade and more than 70 per cent dropping out of school with less than 4 years of school attendance. So, despite the fact that Latin America reached almost full primary school enrollment rates, levels of school completion were very low. Unfortunately, this alternative measure is not available for the early period studied here and it is also not available by birth cohort for the later period. Yet another measure of human capital inequality uses skill premia, as studied for long-term periods by van Zanden (2009). He found, for example, that less developed countries such as Indonesia and India had quite high skill premia.
    ${ }^{7}$ While borders changed during the colonial and post-colonial period, we always refer to today's borders as far as possible.

[^5]:    ${ }^{8}$ During the $19^{\text {th }}$ century, the New Republican governments forbade statements about ethnicity.
    ${ }^{9}$ See notes to Table 3.

[^6]:    ${ }^{10}$ The data set refers mostly to mothers. Moradi (2002) explored the potential difference between mothers and non-mothers. He finds a very moderate selectivity among young mothers. Mothers aged 20-25 years were slightly less educated than the reference population. By employing usual height elasticities for education levels, Moradi estimates about 1 mm shorter height of mothers, compared with the reference population of all women aged 20-25 years. There was no significant selectivity among older women. This result suggests that selectivity of mothers might not be a major problem.

[^7]:    ${ }^{11}$ We thank Yvonne Stolz, who plans to study the Brazilian case in more detail, for providing the $18^{\text {th }}$ century evidence.
    ${ }^{12}$ See also the Appendix referenced in footnote 1. In a separate Appendix (available from the authors), we show that the inequality of literacy between the taller and shorter half correlates with the inequality of numeracy measured in the same way.

[^8]:    ${ }^{13}$ Kuznets effects have recently been studied by Morrison and Murtin (2007) for educational inequality. They construct a within-country indicator on the basis of primary, secondary and tertiary enrollment rates and confirm the existence of an educational Kuznets curve by studying this indicator for the 1870-2000 period. Kuznets curves have also been a traditional field of study for income inequality research. For example, Prados de la Escosura (2008) recently found a Kuznets curve for Spain, 1850-2000.

[^9]:    ${ }^{14}$ Descriptive statistics of the explanatory variables are available in the Appendix, see footnote 1. Some variables are slightly skewed, but in most cases the skewness is only modest. Given that some of the values of the dependent variables are negative, we decided not to use a logarithmic transformation.

[^10]:    ${ }^{15}$ Predicted values are available from the authors.

