GENDER EQUALITY AND INEQUALITY IN NUMERACY: THE CASE OF LATIN AMERICA AND THE CARIBBEAN, 1880-1949

KERSTIN MANZEL[1]
JÖRG BATEN[2]

RESUMEN

En este artículo se estudia las disparidades de género en numeracy para 28 países de América Latina y el Caribe para el período de 1880 a 1949. Para la consecución de este propósito; el trabajo explora el alcance del denominado "age heaping", el cual suministra valiosa información sobre la tendencia de las personas a redondear su edad y que, por tanto, entrega indirectamente información sobre el nivel educativo de un país en el largo plazo. En particular investigamos la hipótesis de un desarrollo en forma "U", es decir, un aumento de las disparidades de género a un bajo nivel de educación, y asimismo, una disminución de las disparidades a niveles más altos de educación. En este sentido, la parte superior es plenamente confirmada, sin embargo, en la parte inferior encontramos sólo algún grado evidencia, sin perjuicio que esta parte es relativamente pequeña.
A partir de esta nueva línea de investigación, los resultados del estudio evidencian que por un lado, los países no-hispanos del Caribe mostraron altos niveles de numeracy y, por otro, que las diferencias de género en este indicador eran más tenues respecto de los países latinoamericanos.

Palabras Clave: Capital Humano, Age Heaping, Educación, Disparidades de Género, América Latina

Abstract
This article outlines the development of gender disparities in education for 28 Latin American and Caribbean countries for the period from 1880 to 1949, using age heaping techniques. We explore in particular the hypothesis of a U-shaped development of women's education during economic development, i.e., a decrease in gender equality at lower levels of overall education, and increasing gender equality at higher levels. For the downward sloping part, we find some evidence, although this part is relatively small. The upward sloping part is strongly confirmed. We also find that non-Hispanic Caribbean countries had substantially lower gender inequality

---

[1] University of Tuebingen, Department of Economics, Mohlstrasse 36, 72074 Tübingen. eMail: Kerstin.manzel@uni-tuebingen.de
[2] University of Tuebingen and CESifo, Department of Economics, Mohlstrasse 36, 72074 Tübingen. eMail: Joerg.baten@uni-tuebingen.de

rates than Latin American countries. A second major contribution is to document the development of average numerical abilities (of both genders) in these 28 countries.

1. Introduction

Goldin (1995) argued that female labor force participation follows a U-shaped pattern: at low income levels, women work intensively in the production of goods, such as in today's West, Central, and East Africa. Much of the production takes place in the household and consists of agricultural activities (dairy, poultry, rice etc.), or proto-industrial cottage work. As market integration increases and incomes start to rise, two separate effects might occur. Firstly, an income effect might lead women to spend more time with children and to do unpaid household work such as cooking and cleaning, especially if women's participation in the new opportunities of manual manufacturing employment is socially stigmatized. Secondly, the relative productivity of traditional home production might decline, as new production techniques in factories become more efficient. Home produced goods are substituted by factory products or those produced on large modern farms. At this second level of development, women's labor force participation might be lower, as Goldin observes in a cross-section for some of today's middle income countries (in Latin America or Southern Africa, for example). The relative social status of women might also decline. Finally, in the third phase of development, women's relative education strongly improves and opens the way to white-collar employment. This type of employment tends to be less stigmatised than manual manufacturing work. In this third phase or level, female participation rises again, confirmed by cross-sectional evidence from Europe, East Asia, North America, and parts of the Asia/Pacific region.

Mammen and Paxson (2000) confirm Goldin's study by estimating the U-shape for a number of years[3], and compare the absolute and relative education of women.[4] Especially important for our study is their finding that the level of female schooling remains mostly

---

[3] They also included a fixed effects estimation in which the shape of the U is determined by variation within countries over time (1970-1985), rather than cross-sectionally.

[4] Relative to the level of development (measured with GDP per capita), the absolute number of women's schooling years increases more or less monotonically, whereas the gap in years of schooling reaches a maximum around 1000 US-$ with gap of about 2 years, and then remains more or less constant, after a modest increase to around 1 year (i.e. overall, the relative gap declines).

stagnant at very low development levels, i.e. for per capita income levels between 400 and 800 US-$. The gap between male and female education increases in this segment, from about one to two years of missing schooling for women. At an income level of 400 US-$, males have received about two years and females one year of formal schooling; at an income level of 800 US-$, males received three years, whereas women still only went to school for one year. In other words, they showed a declining part of the U-shape of education (i.e., increasing gender inequality) for a cross-section of developing countries 1970-1985 in the very low income range (Mammen and Paxon 2000, Figure 1). [5]

In general, female labor force participation tends to correlate with the social status of women. If female employment prospects are good, families might also invest more in girls' education. The main contribution of our study is to test whether the U-shape can also be observed for the relationship between average education and the ratio between female and male education, measured in terms of numerical abilities, in 28 countries of Latin America and the Caribbean during the 1880s-1940s. In earlier research, the strongest evidence for the U-shape came from cross-sectional evidence from rich, middle, and poor countries up until today, since long-run data on the development of gender equality is quite limited. [6] A new measurement strategy, age heaping, allows us to trace the relative numerical abilities of women in Latin American and Caribbean countries back to the birth cohort of the 1880s. Applying this methodology to assess gender equality in such a large number of countries is a substantial contribution, as it opens new avenues for research in this field. We find some evidence on the downward sloping part of the educational U-shape function, although this part is relatively small. In contrast, we find that most Latin American development during this period can be characterized by the upward sloping part of the educational U.

---

[5] These countries were at that time as poor or slightly poorer than the Latin American countries studied here, whose GDP per capita ranged between 681 US-$ in 1870 to 1,481 US-$ in 1913 (Maddison 2001).
[6] With the favorable exception of Goldin, who assessed the U.S. case and concluded that evidence on the declining part of the U had to be adjusted by underreporting of female production for the market.

Why is gender inequality in education important? There is a fair amount of evidence showing an influence of gender inequality on economic growth. Most studies, however, come to different conclusions. Both negative and positive effects of gender inequality on economic growth are detected, although the latter are less numerous (see, for example, Dollar and Gatti 1999; Klasen 2002; Barro and Lee 1994). However, apart from the question of whether gender inequality has a (negative) effect on economic growth, there are more reasons why equity between sexes may be important for a society and should, therefore, be studied in greater detail. Equity in education is, for instance, a necessary precondition for equity of life chances. Better educated women tend to improve nutrition levels and the prevention of illnesses which reduces childhood mortality (Murthi et al. 1995, Hill and King 1993, p. 18). Intra-household distribution models show that income controlled by mothers has a greater positive effect on the health situation of the whole family than income under male control (Thomas 1990). If now the bargaining power of females relative to men increases with their educational level, this also leads to a positive impact on the health status of children.

However, in order to study these effects more closely, more human capital estimates for both genders are needed. Until now, data on the educational system in Latin America and the Caribbean were scarce. Literacy data were available for all countries after 1950, but before this time, only fragmentary information had been published for the larger countries. One important component of human capital is numeracy, the ability to process numbers. In order to measure numeracy, we estimate in this study the degree of age heaping, i.e. the tendency to round up or down one's age. It provides information about numeracy skills or numerical discipline and can be used as a proxy variable for an important component of the educational level of a population. By applying this method, we make a first attempt to estimate and outline the degree and development of gender disparities for numeracy for the period 1880 to 1949.

The Latin American and Caribbean region is not a homogenous one. There are cultural, social and economic development differences as well as historical ones. All these differences – in some cases striking differences, in other cases only subtle nuances – contribute inter alia to the evolution of gender disparities in education. Consequently, we will adopt a generalist approach and underline differences between countries and subgroups of countries without providing much country-specific detail. Further studies may be required to describe gender inequality in education in individual countries in greater detail.

The study is organized as follows. Section 2 presents a review of the literature concerning gender inequality in Latin America. Section 3 describes the methodology used to approximate gender differentials in numeracy, presents our data and discusses possible selectivity issues. Section 4 focuses on the development of numerical abilities and gender disparities in numeracy. Section 5 presents regression results on the determinants of gender inequality and section 6 concludes.

2. Education and gender related issues in Latin America and the Caribbean

Today, Latin America has a surprisingly low level of gender inequality in education (Figure 1). It is the only region in the developing world, in which girls' secondary attainment equals boys' educational attainment (UN 2005). The ratio from female to male adult literacy rates has around the same value as for Central and Eastern Europe. Nevertheless, there are striking differences in gender inequality within Latin America. Indigenous women in Bolivia still have significantly lower education levels than men from the same ethnic group (De Ferranti et al. 2004, p. 96).

So why did gender inequality in education exist at all in earlier times? Explanations usually focus on two different subjects: parental discrimination on the one hand and labor market discrimination on the other hand (Kingdon 1997). These two factors are linked to each other. Higher returns for male education than for women, in terms of higher wages for males

are common even in today's industrialized countries, although the wage differential is much larger in developing countries. The return for sending a daughter to school is therefore lower, especially if we consider that daughters usually marry and leave home, so that they are not able to care for their parents in old age. In addition to this, in rural areas, long distances to the nearest school make schooling of girls less likely than that of boys. Parents are afraid of sending their daughters to school alone. Therefore, increased availability of schools may have a greater impact on schooling for girls than for boys (Greer 1969).

Emerson and Portela (2003) emphasize a strong persistence of child labor from one generation to the next but parental schooling influences child labor decisions concerning sons and daughters in different ways. While the father's schooling has a stronger impact on sons` school attendance and child labor, the mother's educational level has a stronger impact on daughters` child labor status (Emerson and Portela 2003).

The social inequality of education is another crucial element, which actually survives until today and which might be correlated with gender inequality. During the colonial period, education was restricted to a small elite and was mainly the church's domain (Kowalewski and Saindon 1992). Engerman and Sokoloff (2005, p. 917) note that the franchise criterion of literacy often used in Latin America might have prevented the ruling elites from extending public schooling too quickly, as the poorer strata of voters would have achieved political power. Independence led to changes in the institutions providing education. The new ideas insisted on the modernization of society via education and this also included female education (Miller 2003, pp. 207; Reimers 2006, p. 434). The church resisted this, but its power was diminishing. One official Chilean church journal in the 1870s predicted that secondary schooling for women is "*nothing more than mere brothels financed by the taxpayers*" (cited in Fisher 1974, pp. 189). This view was not uncommon during the second half of the nineteenth century. While the church worried about female secondary education, only very few women received any education. Even though concern for female education increased, it was never

argued that it was desirable for equity reasons. Rather, female education was exclusively intended to create better daughters, wives and mothers (Lavrín 1998, p. 103; Christiansen and Christiansen 2004, p. 47). In order to fulfill these roles and to raise children who would constitute the new society of an independent Latin America, female education could no longer be overlooked.

As one of the earliest countries in Latin America, Chile decreed in the organic law of Primary Education as early as 1860 that in all departments exceeding 2,000 inhabitants at least one school for boys and one for girls had to be established (Schiefelbein and Farrell 1982, p. 229). However such laws were never executed throughout Latin America. Scarce resources and a lack of teachers made female education a difficult task. Discussions about coeducational schools would not rise for decades and few qualified female teachers existed (Kent Besse 1996 p. 133). Domingo Sarmiento, one of the most important leaders in the promotion of education in Argentina, who influenced the educational debate throughout Spanish America, saw this problem and even recruited female school teachers from the United States (Miller 2003, p. 210). Thanks to these and similar measures, female education progressed slowly but surely.

The curriculum focused mainly on religious catechism, reading, writing and arithmetic. Girls were taught less writing and arithmetic, while sewing and household duties were included (Vaughan 1990). In particular, in countries specializing in agriculture, and in rural areas, female education was not seen as an important issue because the relevant knowledge could be learned through informal, oral methods (Stromquist 1992).

Total literacy rates in Spanish America increased from under 10% at independence (ca. 1820) to 15% around 1850, and to 27% in 1900 (Greer 1969). Nevertheless, regional disparities remained large until the end of the 20th century and literacy rates were consistently higher in cities than in urban areas (Mariscal and Sokoloff 2000, Newland 1994). Vaughan (1990) observes for Mexico: "*Whether a region or locality was commercially prosperous […]*

was critical to school expansion". While 45% of the population in northern Mexico around 1910 was literate, in the center and in the south only 27 and 14 percent respectively knew how to read and write (Vaughan 1990). As expected, gender disparities were considerably higher in the Mexican south and center than in the more prosperous northern regions (Greer 1969).

Mariscal and Sokoloff (2000) explore the reasons for the late investment in education in Latin American countries in contrast with the United States and Canada. Although income levels in the Latin New World were relatively high, wealth has traditionally been concentrated in the hands of the elite. The authors' main finding is that social inequality constrained the introduction of tax-financed compulsory primary education. In Argentina and Uruguay, in contrast, the desire to attract European immigrants resulted in a special interest on the part of the state in improving education. Immigrants demanded better public services and their higher educational level had a positive impact on the educational level of the whole country. Nevertheless, with the exception of these countries, Mariscal and Sokoloff (2000) conclude that social inequality and the concentration of political power in few hands impeded the formation of appropriate institutions to promote education during the 19th and early 20th centuries.

Thorp (1998) studies the gender gap in education for Latin American countries and notes that education usually improves first for boys and then for girls.[7] Therefore, the gender gap will rise initially until female education catches up and the gender gap declines.

Did differences exist between continental Latin America and the Caribbean in terms of gender inequality? Ellis (2003) emphasizes that the belief that non-Hispanic Caribbean societies were matrifocal is misleading, because this would obscure the fact that women also suffer discrimination in this region. However, gender inequality in education was substantially lower than in the Hispanic Caribbean or Latin America (Ellis 2003). Slavery had, in some

---

[7] She defines the gender gap as the absolute difference between adult literacy rates of males and females.

cases, a leveling effect between the sexes. Race, color and class always played a more important role in this society than gender distinctions. Rich male and female whites enjoyed similar power over the lower classes. Differences between male and female slaves were minor. Both worked and both had been torn away from their cultural roots. Practicing the language or culture of their homelands was strictly forbidden. In this institutional framework, traditional gender roles could not develop as strongly as elsewhere (Wiltshire-Brodber 1999, pp.136-138). Caribbean women worked outside the home more often than Latin American women and had a greater economic influence within the family.

Today, these countries even show female education advantages (Ellis 2003, pp. 11) and the expression "marginalization of Caribbean men" (Ellis 2003, p. 147) has become famous. Caribbean women contribute significantly to the income of their families. Moreover, males migrated more often in search of employment, leaving wife and children who had to earn their own incomes (Brereton 1999, p. 130). The share of female-headed households in the Caribbean is therefore exceptionally high and these households are not as stigmatized as in Latin American societies. Therefore, we might expect lower gender inequality especially in the non-Hispanic Caribbean.

In sum, Latin American educational development was relatively slow, partly due to social inequality, although education developed more favorably in the southern cone countries. Gender inequality was high in earlier times, especially in continental Latin America and the Hispanic Caribbean. Today gender inequality in education is relatively low, so we might expect a more egalitarian labor force participation in the future (Goldin 1995).

3. Data and Methods

Age heaping has been used a number of times recently to measure education levels (Mokyr 1983, Crayen and Baten 2008a and 2008b, A'Hearn, Baten and Crayen 2009, de Moor and van Zanden 2008, Clark 2007, Manzel 2007, Baten, Crayen and Manzel 2008, see also the

applications in Cinnirella 2008, Mironov 2006, O'Grada 2006). It describes the phenomenon that people tend to round up or down their age, mostly in multiples of five, when asked how old they are. The main reasons for this are lack of knowledge about their real age or lack of numerical discipline. Consequently, estimating the degree of age heaping gives us information about the educational system as well as about institutions in a society.

As early as the 1950s Bachi (1951) and Myers (1954) found a correlation between the degree of age heaping and literacy. Mokyr (1983) was the first to apply age heaping as a proxy variable for the educational level of a population in order to investigate whether there was a brain drain from pre-famine Ireland. Studies find a strong negative correlation between age heaping and literacy or schooling, such as Crayen and Baten (2008b) for the 19th and 20th centuries, A'Hearn, Baten and Crayen (2009) for the 19th century U.S. states and the countries of Europe during the early modern period, Manzel and Baten (2008) for Argentina during the 19th century, and Nagi, Stockwell and Snavley (1973) for African countries of the mid-20th century. To measure the degree of age heaping, various indices can be used. A'Hearn, Baten and Crayen (2009) show that the Whipple Index is most appropriate for this purpose. It determines the tendency of age heaping on the digits 5 and 0 and is calculated by taking the ratio of the sum of people reporting an age ending on multiples of five and the total sum of people in a certain age range. This ratio is then multiplied by 500. Meaningful interpretations of the index vary between 100 and 500.[8] In the case of 100, no age heaping on multiples of five is present, in the case of 500, the age data contain only digits ending in multiples of five (Hobbs 2004).

$$(1) \ Wh = \left(\frac{\sum (Age25 + Age30)}{1/5 * \sum Age23 + Age24 + Age25 + .. + Age32)} \right) \times 100$$

Hence, the Whipple Index (Wh) gives us information about numeracy skills or numerical discipline and can be used as a proxy for an important component of the

---

[8] A Whipple Index of 0 is theoretically possible and would mean an avoidance of ages ending in 5 and 0. However, values below 95-100 are uncommon.

educational level of a population. The calculation of the Whipple Index requires single age data for ten successive years, so that each terminal digit appears once. Mortality will have the effect that fewer people are alive at age 44 than at age 40, and at age 49 than at age 45, which could bias the Whipple Index downwards (Crayen and Baten 2008a). Therefore we choose the age groups 23-32, 33-42 etc. to overcome this problem. We exclude age data for under 23-year olds, because many young males and females married in their early twenties or late teens and had to register as voters, military conscripts etc. On such occasions, they were sometimes subject to age requirements, a condition which gave rise to increased age awareness. Moreover, individuals grow physically during this period, which makes it easier to determine their age with a relatively high accuracy.[9] Age information for over 72-year olds is not included as age statements of older people involve several problems: age exaggeration, survivor bias, higher mortality of males (Del Popolo 2000) and other household members who report the ages of older persons play a more pronounced role than at younger ages.[10]

The Whipple Index is defined inversely, i.e. it represents lack of numeracy rather than numeracy. For an easier interpretation, A'Hearn, Baten and Crayen (2009) suggested another index, the ABCC index.[11] It transforms the Whipple Index and yields an estimate of the share of individuals who correctly report their age:

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

The method of approximating educational levels with age heaping behaviour certainly has its deficiencies in measuring human capital, as misreporting of ages may also have

---

[9] A 17-year-old might round up/down to 18 or 16, but not to 15 or 20. Moreover, children were excluded because of a high likelihood that the parents rather than the child himself answered the question.

[10] Studying population enumerations of eight Latin American countries, Del Popolo finds that the share of population with a stated age of 90 and higher is highly correlated with the Whipple Index for 53 to 82-year olds. Thus, countries with stronger age heaping might have more age exaggerations. A further result of her study is that the age error increases with age. Thus, not only does heaping from 72 to 75 play a role, but also heaping from age 72 to 80, 90 or 100. At what age do these effects become too strong to measure age heaping in a reliable way? We do not know this with certainty. In some countries, the effect becomes visible from the age of 70 onwards, in others only from the age of 80. In order to obtain reliable results, we exclude those older than 72 from our analysis.

[11] The name comes from the initials of the authors' last names plus that of Greg Clark, who suggested this in a comment on their paper.

political or cultural reasons. The degree to which age heaping is influenced by schooling and the effect of other institutional factors is not easy to disentangle, although Crayen and Baten (2008b) assessed this and found that schooling was more important than other factors such as bureaucracy and previous census-taking. We conclude that -- at least in the absence of other indicators – age heaping is a valuable instrument to approximate the development of human capital.

Data and representativeness

We use official census data available from the United Nations Statistical Yearbooks to estimate numeracy levels for 28 Latin American and Caribbean countries from 1880 to 1949 (see the Appendix for a complete list). Our data contain information for all continental Latin American countries with the sole exception of Paraguay.[12] For the Caribbean, we have reliable data for 9 countries, including Puerto Rico as a relatively advanced country and Haïti, the poorest country of the region. Based on the number of inhabitants at each age, we calculated Whipple and ABCC Indices for each country and birth decade. If data were available from more than one census in a country, we obtained two different Whipple Indices for a birth decade. In a few cases, the age heaping estimates differed substantially between the two census years. After studying the institutional environment of census-taking we found that in some cases the situation differed. For example, passports had been introduced, so that people could look up their age. This occurred, for instance, in the case of Haïti. While in the census of 1950 pronounced age heaping patterns were observable, in the census of 1971, age heaping had disappeared completely. In order to obtain reliable estimates, we only included censuses if the population was directly asked for their age and passports or similar documents were not widely available (see the appendix concerning the included censuses). If the

---

[12] We decided to exclude Paraguay from an analysis, as data from the censuses of 1962 and 1972 published in the United Nations Demographic Yearbooks not only differed considerably in their heaping behavior, but also gave completely different total population figures. While the census of 1962 gave a total population of 1.7 Mio for the age range 23-72, the census of 1972 covered only around 875,000 inhabitants in the same age range.

institutional circumstances changed only very modestly, and the estimated age heaping between an early and a late census differed little, we estimated Whipple Indices on the basis of the earlier census and extrapolated for later birth decades with the growth rates of the later census, but indexing the series on the levels of the earlier census (see the Appendix Table A1.). However, in the regression analysis in part 4, we excluded the estimated values to make sure that we did not introduce biased data. Estimated values are only used in the descriptive figures.

How did the census enumerators obtain their information? Is it possible that we are measuring the numerical abilities of the enumerator or his ability to estimate age? These are legitimate concerns when we try to approximate numeracy with census data. We do not deny that this is a severe problem and may in some cases bias our results. Some enumerators might have taken their duties more seriously than others, and we do not have information on whether all household members were asked individually in all cases. We do know that census takers were required to ask each person individually (see, for instance, Ministerio de Economia 1965, p. XII) and information was collected by canvassers (Goyer and Domschke 1983, p. 8). If census takers influenced the results strongly, Whipple Indices should vary considerably from one census to the next or within areas enumerated by different census takers. However, we find that this was not the case. Whipple Indices for the same countries and birth decades in different census years differed only slightly if the institutional framework did not change. Therefore, our analysis is based on the assumption that census taker errors are uniformly distributed across our samples. For the late 18th and early 19th centuries individual census data for Latin American countries are available. In these sources, census enumerators asked each person individually and did not adjust for obviously erroneous age information (Manzel and Baten 2008).

The reliability of official statistics, including also census statistics, has often been questioned for Latin American countries. So the crucial question is how reliable these

population enumerations are. In the cases of Costa Rica and Peru, remote areas of the country were not directly enumerated. This will probably bias our numeracy estimates in these two countries upward, as remote areas with higher shares of indigenous people have fewer schools and therefore have lower educational levels. In Brazil parts of the census schedules of 1950 from Minas Gerais, São Paulo and Paraná were lost (Goyer and Domschke 1983, p. 85). In the analysis, we will check the sensitivity of results by including and excluding these cases.

We consider censuses taken from 1947 onwards, a period when growing concern about improving official statistics began to arise. The United Nations, especially the Latin American and Caribbean Demographic Center (CELADE), provided technical assistance on how to carry out effective population enumerations and evaluate demographic studies (Goyer and Domschke 1983, p. 278). One or more test censuses were conducted in countries with little experience of census taking in order to train census takers as well as to improve the questionnaires (Goyer and Domschke 1983, p. 37 – 350). Will this have an effect on the accuracy of age statements? Probably not, as the question regarding one's age is a relatively simple one and we would not expect people to know their age more accurately only because they were asked two or three times for their age in a period of three years.[13]

In sum, the factors "underenumeration" and "pilot census" presented above may potentially influence our results. Therefore, we will check the robustness of our results by excluding countries where underenumeration occurred as well as those who carried out one or more pilot censuses to improve the census quality.

The development of numeracy in Latin America and the Caribbean

In Appendix Table A.1, we present the numeracy levels and gender disparities for the birth decades of the 1880s to 1940s and Table 1 reports summary statistics. The Whipple and

---

[13] Only a very long history of census taking (6 to 7 censuses) may positively influence numerical abilities and discipline in a society, as Crayen and Baten (2008) found out.

ABCC Indices indicate a wide range of numeracy levels. Differences between countries are striking: While as early as 1880, Argentina only showed minor age heaping tendencies (Argentina had a numeracy level of 97.5, the same value as Portugal in 1940, see Crayen and Baten 2008b), Ecuador had numeracy rates of only 52 percent in the 1880s. Huge differences remain between these countries until the middle of the 20th century (Figure 2, a-d). The leading countries in numeracy levels in Latin America are Argentina and Uruguay. Non-Hispanic Caribbean countries also feature very well. An exceptionally rapid improvement in numeracy levels took place in Ecuador, Puerto Rico and Bolivia. In Ecuador, numeracy levels improved from 52 percent in 1890 to 84 percent in the 1920s and in Bolivia, the share of people who reported exact ages increased from 55 percent in the 1880s to 81 percent in the 1920s. And, very interestingly, Guatemala and Haiti started with values that were not exceptionally low in the 1880s, but obviously suffered major educational development problems in the subsequent period. Therefore, they ended up at relatively low levels of numeracy. Numeracy levels increased during the period under study in all three parts of this region at similar pace (Figure 3). However, the non-Hispanic Caribbean countries started at better levels, compared to the Hispanic Caribbean or Latin American countries.[14]

According to their level of numeracy, we can group our countries in 4 categories: countries with a relatively low level of heaping and those with a moderate, high and an extremely high level of age heaping (Figure 4). The countries shaded in dark grey on this map are those with a low numeracy level in 1900. The Southern Cone, Brazil, Costa Rica, Surinam and Guyana had higher numeracy levels than the Central American countries and Mexico around 1900 (the smaller Caribbean islands are not visible here).

Which factors may contribute to the decline in age heaping? The determinants of improving numeracy are generally an expansion of education, via both formal schooling and informal education at home. This happened quite early especially in the countries with a high

---

[14] See Appendix for the classification of the countries.

share of European immigration. In Argentina and Uruguay, the desire to attract European immigrants resulted in a special interest on the part of the state in improving education. Immigrants demanded better public services and their higher educational level had a positive impact on the educational level of the whole Argentinean population (Mariscal and Sokoloff 2000). Towards the end of the period under consideration, numeracy levels improved considerably in almost all countries. However, in the Domincan Republic, Guatemala and Nicaragua, numeracy was still at a lower level than in the other countries. Uruguay, Guadeloupe, Trinidad and Tobago, Martinique, Suriname and the Leeward Islands featured the highest numeracy levels towards the end of the period.[15]

In sum, numerical abilities in the non-Hispanic Caribbean were consistently higher than in Latin America and the Hispanic-Caribbean countries of our sample (Figure 2 and 3). By 1940, differences were reduced but still existed. The educational system of the French, Dutch and British might have led to a high educational level. The regression analysis in part 4 will address these differences further taking alternative variables into account.

The development of gender equality in Latin America and the Caribbean

To measure educational equality between the sexes, we define a measure of "gender equality" (GE) as

$$(3) \ GE = -\left(\frac{(whf - whm)}{whm} \right) \times 100$$

where whf and whm are the Whipple Indices of females and males, respectively. Thus, the higher our measure of gender equality, the lower the share of women rounding up or down their age in comparison to men rounding up or down in a certain country. A positive (negative) gender equality index implies a female (male) numeracy advantage. Most of the time, the index will be negative. We formulate this as gender equality in order to make it more

---

[15] Please note that data are not available for all countries for the birth decade of 1940.

easily comparable with the literature on female labor force participation rates (Goldin 1995, Mammen and Paxson 2000). Of course, this does not imply that our countries were characterized by gender "equality" between 1880 and 1949.

Applying our measure of gender equality in numeracy, we find substantial variation between countries (Figure 5). As we would have expected, our gender equality index is in most cases negative, indicating that women were less numerate than men during this period (indicated by darker grey tones). Argentina, Uruguay, Suriname and Guyana had relatively good equality levels, whereas some of the Central American countries were characterized by stronger inequality between the genders.

Latin American countries had typically lower gender equality indices than Caribbean countries during the whole period under consideration (Figure 6). In both subregions, the overall trend is characterized by increasing gender equality in numeracy which went hand in hand with economic development. Among the Caribbean countries, even in Haïti -- the poorest and least numerate country in the region -- there was no evidence for gender inequality in numeracy. Colombia, Guatemala and the Dominican Republic had relatively low levels of gender equality as well as low levels of overall numeracy (Appendix Table A.1). In the Dominican Republic, the effect of a long and repressive dictatorship might have strengthened the patriarchic gender relationships. In Argentina we can find no evidence for gender inequality in numeracy. It is the only country which, as early as in 1880, showed neither evidence for age heaping nor for gender disparities. Until the 1940s, gender equality increased substantially in all Latin American countries. Colombia and Guatemala still had gender equality indices below 10, but this implies a decline to a third of their initial inequality level.

4. The U hypothesis and other potential determinants of gender equality

Until now, we have mostly described the development of numeracy and gender equality in Latin America and the Caribbean. In the following section, we will assess the question whether the temporal development of gender equality in numeracy follows a U-shaped pattern. In the introduction it was explained that female labor force participation has been identified by a U-shape in modern cross-sections by Goldin (1995) and Mammen and Paxson (2000). They argued that at initial levels of development, stigmata against women's work in the factories or the fixed cost of working outside the home might have reduced women's participation. Only as non-stigmatized white-collar employment became available and female education and hence wages rose, did the share of female participation increase again (Goldin stresses mainly the former, Mammen and Paxson the latter factor).

We are interested here in the question whether the first declining and then increasing female labor force participation is also reflected in the equality of education. The economic mechanism might be that higher relative female income expectations motivate parents to send their daughters to school initially. In a second phase, the relative female education might have declined, or at least stagnated, due to factors such as: (a) a replacement of home production with factory production, (b) income effects enabling work at home (c) stigmata effects of married females working in factories. Finally, in the third phase, female equality of numeracy increases again. We actually see this development in some of the poorer countries in our data set, such as Ecuador and Colombia (Figure 7, panel A and B). In the other countries, this development cannot be observed. However, there is cross-sectional evidence that some of the poorest countries – such as Haiti – actually had quite low gender inequality. On average, the development of gender equality does show a small initial decline in educational equality in the Hispanic Caribbean, where the equality index declined from –20 to –22 between the 1880s and 1890s, and remained low in the 1900s (Figure 6). In Latin America, there was a

stagnation between the 1880s and 1890s, whereas in the non-Hispanic Caribbean equality grew.[16]

To test this relationship systematically, we carry out a panel data analysis with gender equality in numeracy as a dependent variable (unbalanced panel, Table 2). This allows us to examine the time series and cross-sectional evidence for the U-shape hypothesis at the same time. Apart from the U-shape hypothesis, we need to examine a number of other factors. Firstly, female participation in elections could be a potentially important variable. However, in the case of our countries, between 1880 and 1949, only a few provided electoral rights to women and these rights came quite late. Hence, this dummy variable might pick up some of the time trend. As a second additional variable, democracy favors an equal distribution of resources, hence perhaps distributing more education also to women. We measure democracy with the POLITY2 variable, which is basically an index of a wide variety of different participation and franchise indicators, ranging from –9 to +10 in our sample. Democracy values were higher in the 1890s and 1900s, whereas during the interwar period, a number of countries became more autocratic than before (for example, Argentina). Given that at the same time, education increases and gender inequality is gradually reduced, there might be some trend correlation leading to opposite signs. Thirdly, some studies suggest that gender inequality might be high among indigenous tribes. Rosemary Thorp (1998, pp. 38) finds pronounced gender differences in countries with large indigenous populations like Bolivia, Peru and Guatemala. However, Bustillo (1993) argues that the indigenous culture had egalitarian ideas towards women. Hence, indigenous population in a society might influence gender equality in either way. Moreover, the language barrier could also play a role here. King and Bellew (1990) emphasize that the language thought and spoken in school was Spanish. As large parts of the population spoke exclusively indigenous languages, especially

---

[16] Whether this development was a short trend or a temporary shock, perhaps during the first era of globalisation, will have to be explored in further studies.

in rural areas, they were excluded from the schooling system, leading to lower education levels of indigenous people.

To test whether higher gender inequality is present in countries with a higher share of indigenous population, we include a dummy variable for these countries. Information on the share of the indigenous population in Latin America and the Caribbean in a historical perspective is not available. Therefore, we will include a dummy variable that takes the value of one if a country has a share of 60 percent or higher of indigenous or part-indigenous population today.

We use a variety of different models to assess the U-shape hypothesis and the other variables (Table 2). We start with a fixed effects model, which basically shows that female equality declines at low levels of general education (the linear term of ABCC is negative), but it starts to increase at higher levels (the squared term of ABCC is positive). How steep the declining or increasing portion of this U-shape might be depends on the size of the coefficients. Plotting the predicted ("fitted") values of gender equality against the ABCC Index indicates that the declining part of the U is small and the decline is modest (Figure 8). In contrast, the upward sloping part is large. This shape does not depend on the squared functional form that we have chosen for the ABCC variable. If we use a LOWESS estimator which does not impose a special functional form, the results are quite similar (Figure 8).[17]

In order to test the robustness of the fixed effects model, we also used a least square dummy variable model with and without dummy variables for birth decades, and find that the results are robust. By including the time dummy variables, the coefficients for the U-shape are still significant, at least at the 10% level of significance. The upward sloping part of the U is also economically significant, as can be seen in Figure 8. In contrast, our political variables are insignificant (general democracy) or only sometimes significant (female voting rights). The democracy variable has the opposite sign once (Table 3, Col. 2), i.e. democracy reduces

---

[17] Lowess (locally weighted regression scatterplot smoothing) uses locally-weighted polynomial regression techniques (see Cleveland 1979).

gender equality. This might be caused by the fact that democratic values actually declined in the interwar years in some countries, whereas gender equality was growing. Moreover, the number of cases of this control regression in Column 3 of Table 3 might be too small to obtain reliable results. Female voting rights were introduced only towards the end of the period. Therefore we cannot fully disentangle its effect from a general trend of increasing gender equality. When time dummies are included, it becomes insignificant due to multicollinearity effects. The share of the indigenous population is insignificant in two of three regressions. In one model it has a significantly negative influence on gender equality, that is, a higher share of indigenous population is associated with a lower level of gender equality (Column 3). A similar inconclusiveness of this factor is also given in Table 3. Hence, we can conclude that countries with a high share of indigenous population might show lower gender equality in education, but the effect is not robust.

In general, the relationship between general numeracy and gender equality might be bi-directional, not necessarily causal in only one direction. As we mentioned in the introduction, gender inequality may also hinder development (reverse causality). However, these growth effects might be more long-run in nature. Studying the development from one decade to the next over just seven decades, the effect from the general numeracy level on gender inequality might be stronger than the opposite direction of causality. In order to study this question, we provide some Two Stage Least Square estimates, using instrumental variables in the following section.

What might be good instrumental variables for the potentially endogenous effect of general numeracy? Given our discussion above, the fact of having been a Spanish or Portuguese colony before independence could be a good instrument, as the English, French and Dutch brought a different educational culture to their colonies, which might have had an

impact on gender equality via this factor.[18] As for the remaining part of the former Spanish and Portuguese colonies, we observed some differences in educational level to be caused by European immigration. Hence, we generated a dummy variable which is one for the cases of Argentina, Uruguay, and Chile. For the border case of Brazil, we generated a separate instrumental variable which is one for the three southern cone countries plus Brazil.

Good instrumental variables should be correlated with the potentially endogenous variable while not influencing the dependent variable, except via the potentially endogenous variable. The former is clearly true for both instruments, since both are correlated with general numeracy. The correlation of the former Spanish or Portuguese colonies variable for numeracy is –0.51. For the former Spanish or Portuguese colonies, the correlation of both immigration dummies is 0.28, both significant at the 1% level. In a joint regression of numeracy on both variables, both are significant. The latter criterion is also the case here: the effect of general numeracy on gender equality follows the same causal channel. Hence we conclude that both instruments are justified. In columns 4 and 5, two versions of IV regressions are displayed. They confirm the U-shape of the relationship studied here.

To test the robustness of our results further, we carried out the same regressions without the countries where pilot censuses were conducted (Table 3, Column 2) and in a second step without the countries where underenumeration occurred (Table 3, Column 3). The coefficients have the same sign and magnitude. Also, the significance level does not change considerably. The shape of the U might even be steeper using the more restricted samples.[19] Thus, our results remain robust to the inclusion of further variables or slight specification changes, as well as to a changing composition of the data set.

---

[18] For example, Browne (2003) and Ellis (2003, p.17-21) argued that female slaves in the French and British Caribbean occupied important economic roles. Therefore gender distinction had less tradition in these countries. In fact, the descriptive analysis above has shown that most non-Spanish-colonized Caribbean countries exhibited lower gender inequality, and that these disparities had been low, at least, since the end of the 19th century. In the Hispanic Caribbean, in contrast, women were more constrained to household work, and formal education had been less important for them.

[19] The information regarding which countries were excluded for the robustness check regressions can be found in the Appendix.

6. Conclusion

In this paper, we tested a hypothesis of a U-shaped development of gender equality relative to general numeracy, which was derived from a U-shaped development of female labor force participation relative to the general level of development (Goldin 1995, Mammen and Paxson 2000). The basic economic mechanism was that if relative female labor market prospects were better, the families might have invested slightly more in the education of their female offspring, and vice versa. This investment might have declined in the first phase of development studied here, but the decline turned out to be modest empirically. In contrast, an increase in gender equality (i.e. the upward sloping part of the U-shape) during the period under study can be strongly confirmed.

In order to test this hypothesis, we developed a method to quantify gender equality in numeracy by using the age heaping method. The advantage of this new method is the improved availability of age statistics or census data in order to estimate the education level. Especially for time periods and countries where data is scarce, this method promises new insights into topics that could not be explored yet, due to the lack of adequate statistics.

We compiled a new data set for 28 Latin American and Caribbean countries measuring gender disparities in numeracy, covering the decades from 1880 to 1940. We also find evidence for pronounced differences within the region: While South American countries like Argentina and Uruguay already had relatively high gender equality in the late 19th century; Central American countries had traditionally lower levels. Non-Hispanic Caribbean countries performed better in terms of gender equality, as well as in overall numeracy. Gender equality increased considerably during this period, leading to equal numeracy levels of males and females at the end of the period in most countries of our sample.

A second major contribution of this study is to document the development of average numerical abilities (of both genders) in these 28 countries. Some of the countries had already

solved their innumeracy problem by the 1880s (Argentina, Guyana). Others experienced strong numeracy growth from initially low levels, but still did not reach 100 percent age numeracy by the birth decade of the 1940s, such as Mexico and Bolivia. And, interestingly, Guatemala and Haiti started with values that were not exceptionally low in the 1880s, but had obviously strong educational development problems in the subsequent period. Therefore, they ended up at relatively low levels of numeracy. In conclusion, this study is a first step to estimate the average numeracy level in these 28 countries, as well as to assess gender equality of numeracy and its determinants.

REFERENCES

A'HEARN, B., BATEN, J. and CRAYEN, D. (2009). "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital", *Journal of Economic History (forthcoming)*.

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

BARRO, R. and LEE, J.-W. (1994). "Sources of economic growth", Carnegie-Rochester Conference Series on Public Policy.

BATEN, J., CRAYEN, D. and MANZEL, K. (2008): "Zahlenfaehigkeit und Zahlendisziplin in Nord- und Westdeutschland, 16.-18. Jahrhundert", *Jahrbuch für Wirtschaftsgeschichte,* 2008-2, pp. 217-229.

BRERETON, B. (1999). "General Problems and Issues in Studying the History of Women." in P. Mohammed, and C. Shepherd (Eds.), *Gender in Caribbean Development: Papers Presented at the Inaugural Seminar of the University of the West Indies, Women and Development Studies Project*. Barbados, Jamaica, Trinidad and Tobago Canoe Press University of the West Indies.

BROWNE, K. (2003). "Female Entrepreneurship in the Caribbean: A Multisite, Pilot Investigation of Gender and Work". In G. CLARK (Ed), *Gender at Work in Economic Life*. Monographs in Economic Anthropology Series, Walnut Creek, CA: Alta Mira Press, pp. 63-94.

BUSTILLO, I. (1993). "Latin America and the Caribbean", in E. M. KING and A. A. HILL (Eds), *Women's Education in Developing Countries: Barriers, Benefit, and Policies*. Baltimore, MD: The Johns Hopkins University Press, pp. 175-210.

CHRISTIANSEN, T. K. and CHRISTIANSEN, T. (2004). *Disobedience, slander, seduction, and assault: women and men in Cajamarca, Peru, 1862-1900*. University of Texas Press.

CIA (1999): *The World Factbook 1999*, Washington DC: CIA.

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

CRAYEN, D. and BATEN, J. (2008a). "Numeracy, Inequality, Age Heaping, and Economic Growth: New Estimation Strategies for Western Europe and the U.S. (17th - 19th centuries)", *University of Tuebingen Working Paper*.

CRAYEN, D. and BATEN, J. (2008b): "Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth", *CESifo Working Paper Series*, 2218.

DE FERRANTI, D., PERRY, G. E. , FERREIRA, F. AND WALTON, M. (2004). *Inequality in Latin America and the Caribbean: Breaking with History?*. Washington DC, The World Bank.

DE MOOR, T. AND J.-L. VAN ZANDEN (2008): 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 5-4, pp. 55-86*

DEL POPOLO, F. (2000). "Los problemas en la declaración de la edad de la población adulta mayor en los censos", *Serie Población y Desarrollo*, Santiago de Chile, CEPAL.

DOLLAR, G. and R. GATTI (1999). "Gender Inequality, Income and Growth: Are Good Times Good for Women?", *Policy Research Report on Gender and Development*, Washington DC, 1 , The World Bank.

ELLIS, P. (2003): *Women, Gender and Development in the Caribbean: Reflections and Projections*, London, Zed Books.

ENGERMAN, S. L. AND SOKOLOFF, K.L. (2005), "The Evolution of Suffrage Institutions in the New World," *Journal of Economic History* 65 (4), pp. 891-919.

EMERSON, P. M AND PORTELA SOUZA, A. (2003). "Is There a Child Labor Trap? Intergenerational Persistence of Child Labor in Brazil," *Economic Development and Cultural Change* 51(2), pp. 375-98.

FISHER, N. (1974). "Chile: The Evolution of an Educational System", *PhD. Thesis,* Nottingham, University of Nottingham.

GOLDIN, C. (1995): "The U-Shaped Female Labor Force Function in Economic Development and Economic History", in T. P. Schultz (Ed), *Investment in Women's Human Capital*. Chicago, The University of Chicago Press.

GOYER, D. S. and DOMSCHE, E. (1983). *The handbook of national population censuses: Latin America and the Caribbean, North America and Oceania.* Westport, Conn., Greenwood.

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

GREER, T. V. (1969). "An Analysis of Mexican Literacy", *Journal of Inter-American Studies*, 3, pp. 466-476.

HILL, A. and KING, E. M. (1993). "Women's education in developing countries: An overview." in E. M. KING (Ed): *Women's Education in Developing Countries: Barriers, Benefit, and Policies*. Baltimore, MD, The Johns Hopkins University Press, pp. 1-50.

HOBBS, F. (2004). "Age and Sex Composition", in J. S. SIEGEL and D. A. SWANSON (Eds), *The Methods and Materials of Demography.* Amsterdam, Elsevier.

KENT BESSE, S. (1996). *Restructuring Patriarchy – The Modernization of Gender Inequality in Brazil, 1914-1940.* Chapel Hill, University of North Carolina Press.

KING, E. M. and BELLEW, R. (1990). "Gain In The Education Of Peruvian Women, 1940 To 1980", *Policy Research Working Paper Series*, Washington DC, The World Bank.

KINGDON, G. G. (1997). "Does the Labour Market Explain Lower Female Schooling in India?," *STICERD - Development Economics Papers*, 01, Suntory and Toyota International Centres for Economics and Related Disciplines, London School of Economics.

KLASEN, S. (2002). "Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development", *The World Bank Economic Review*, 16(3), pp. 345-373.

KOWALEWSKI and SAIDON, J. J. (1992). "The Spread of Literacy in a Latin American Peasant Society: Oaxaca, Mexico, 1890 to 1980", *Comparative Studies in Society and History*, 1, pp. 110-140.

LAVRÍN, A. (1998). *Women, Feminism, and Social Change in Argentina, Chile, and Uruguay, 1890-1940.* Lincoln, University of Nebraska Press.

MADDISON, A. (2001). *Historical Statistics for the World Economy: 1-2003 AD*, Paris, OECD Publishing

MANZEL, K. (2007). "Long run development of human capital and the impact of the hungry forties in Spain", *University of Tuebingen Working Paper*.

MANZEL, K. AND BATEN, J. (2008): "The Development of Numeracy in Colonial and Post-Colonial Latin America (1640-1949)", *University of Tuebingen Working Paper*.

MARISCAL, E. AND SOKOLOFF, K. (2000). "Schooling, suffrage and Inequality in the Americas, 1800-1945", in Haber, S. (Ed.), *Political Institutions and Economic Growth in Latin America*. Stanford, Hoover Institution Press.

MILLER; FRANCISCA (2003). "Women and Education in Latin America", in J.C. CHASTEEN and J. A. WOOD (Eds), *Problems in Modern Latin American History. Sources and Interpretations*. Wilmington, Rowman and Littlefield, pp. 205-210

MINISTERIO DE ECONOMIA (1965). *Tercer Censo Nacional de Población 1961*, San Salvador, Dirección Gereral de Estadística y Censos.

MIRONOV, B.N. (2006) Novaya Istoricheskaya Demografia Imperskoy Rossii: Analycheski obsor Corremennoy Istoriografii (Part 1). Vestnik Sankt-Peterburgskovo Universiteta 2006, Series 2, Istoria, Issue 4, December, pp. 62-90.

MOKYR, J. (1983). *"Why Ireland Starved: A Quantitative and Analytical History of the Irish Economy, 1800-1850"*, London, George Allen & Unwin.

MOMSEN, J. (1987): "The Feminisation of Agriculture in the Caribbean", in J. H. Momsen and J. G. Townsend (Eds), *Geography of Gender in the Third World*. New York, State University of New York Press.

MURTHI, M., GUIO, A.-C. and DRÈZE, J. (1995). "Mortality, fertility, and gender bias in India: A district-level analysis", *Population and Development Review*, 21.

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

NAGI, M. H., STOCKWELL, E. G. and SNAVLEY, L. M. (1973). "Digit Preference and Avoidance in the Age Statistics of Some Recent African Censuses: Some Patterns and Correlates", *International Statistical Review*, 41 (2), pp. 165-174.

NEWLAND, C. (1994). "The Estado Docente and Its Expansion: Spanish American Elementary Education, 1900-1950", *Journal of Latin American Studies*, 2, pp. 449-467.

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

REIMERS, F. (2006): "Education and Social Progress", in V. BULMER-THOMAS, J. H. COATSWORTH and R. CORTÉS CONDE (Eds), *The Cambridge Economic History of Latin America*, Cambridge, Cambridge University Press.

SCHIEFELBEIN, E. AND FARRELL, J. P. (1982). "Women, Schooling, and Work in Chile: Evidence from a Longitudinal Study.", in G. P. Kelly and C. M. Elliott (Eds.), *Women's Education in the Third World: Comparative Perspectives*. Albany, State University of New York Press, pp. 228-243.

STROMQUIST, N. P. (1992). "Women and Literacy: Promises and Constraints", *Annals of the American Academy of Political and Social Science*, 520, pp. 54-65.

THOMAS, D. (1990) "Intra-household Resource Allocation: An Inferential Approach," *Journal of Human Resources*, 25, pp. 635-664.

THORP, R. (1998). *Progress, Poverty and Exclusion: An Economic History of Latin America in the 20th Century*, Baltimore, John Hopkins University Press.

UNESCO (2004): *EFA Global Monitoring Report 2005: Education for All, the Quality Imperative*. Paris, UNESCO.

UNITED NATIONS (2005). "*The Millennium Development Goals Report, 2005*". New York: United Nations.

UNITED NATIONS (various issues), *Demographic Yearbook*, New York, UN.

VAUGHAN, M. K. (1990). "Primary Education and Literacy in Nineteenth-Century Mexico: Research Trends, 1968-1988", *Latin American Research Review*, 25 (1), pp. 31-66.

WILTSHIRE-BROADBER (1999). "Gender, Race and Class in the Caribbean". in P. Mohammed and C. Shepherd (Eds.), *Gender in Caribbean Development: Papers Presented at the Inaugural Seminar of the University of the West Indies, Women and Development Studies Project*, Barbados, Jamaica, Trinidad and Tobago, Canoe Press University of the West Indies.

Figures and Tables
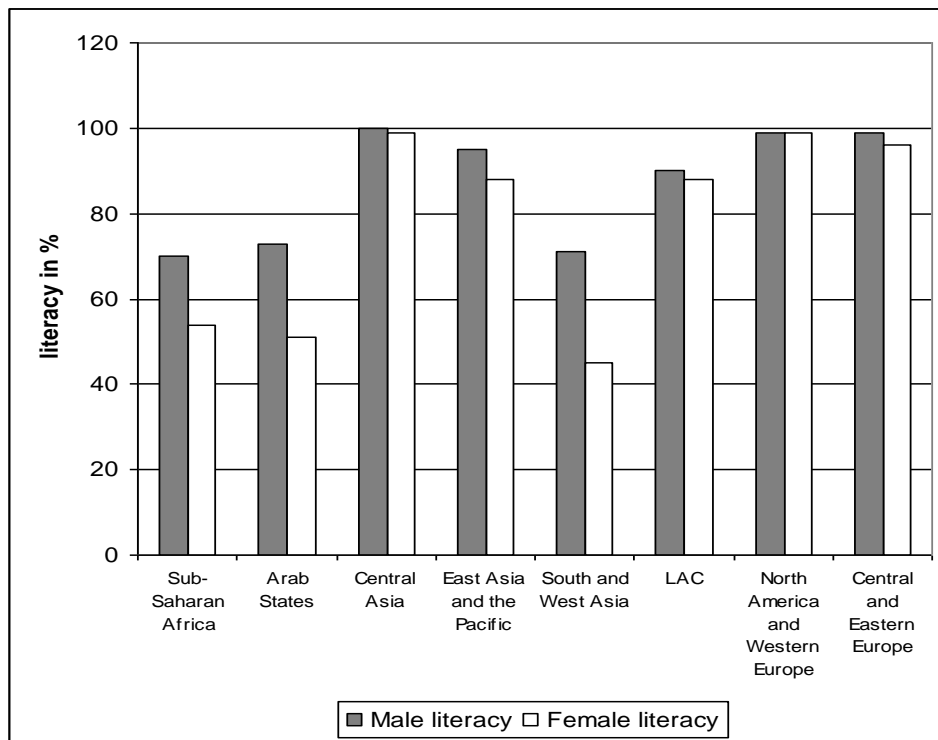
Figure 1: Male and Female Literacy Rates, 2000-2004. Source: UNESCO (2004), table 3.9
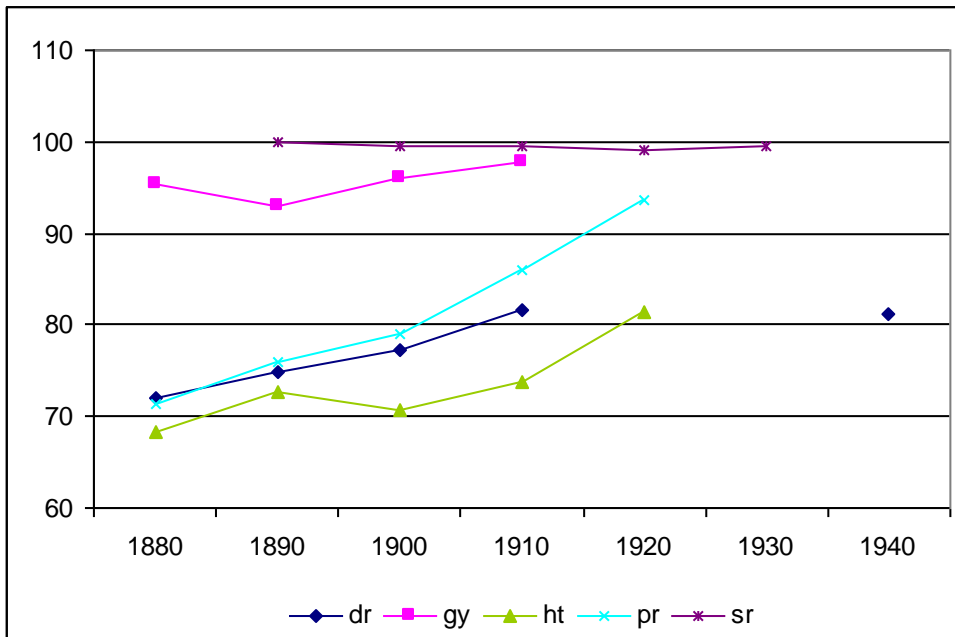
Figure 2a: Numeracy (ABCC, in %) in the Dominican Republic, Puerto Rico, Haiti, Surinam, and Guyana

Note: Estimated values are included here. Abbreviations: see Appendix section "Census Data"
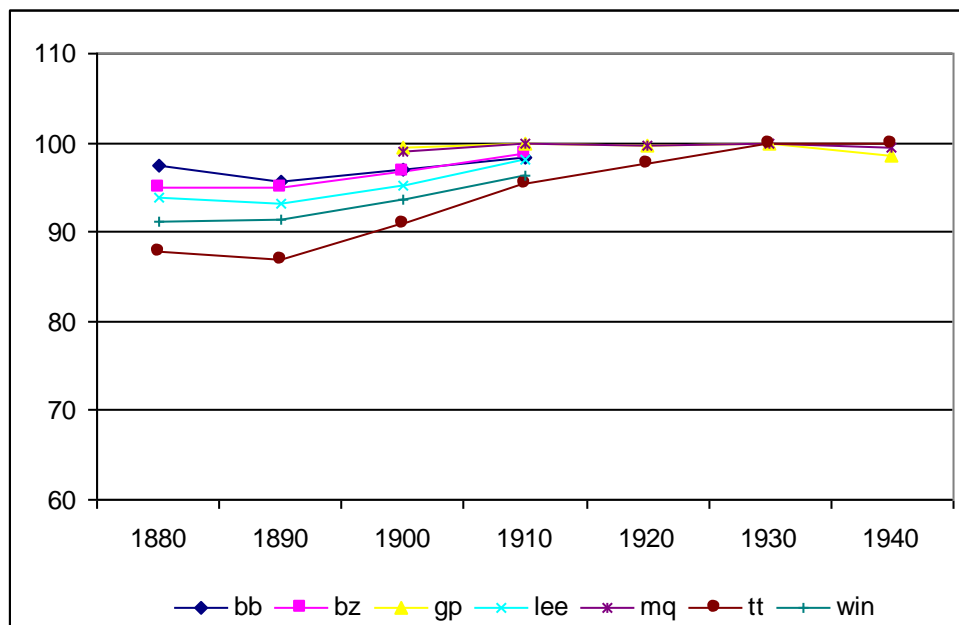
Figure 2b: Numeracy (ABCC, in %) in Non-Hispanic Caribbean except Haiti, Surinam, and Guyana

Note: Estimated values are included here. Abbreviations: see Appendix section "Census Data".
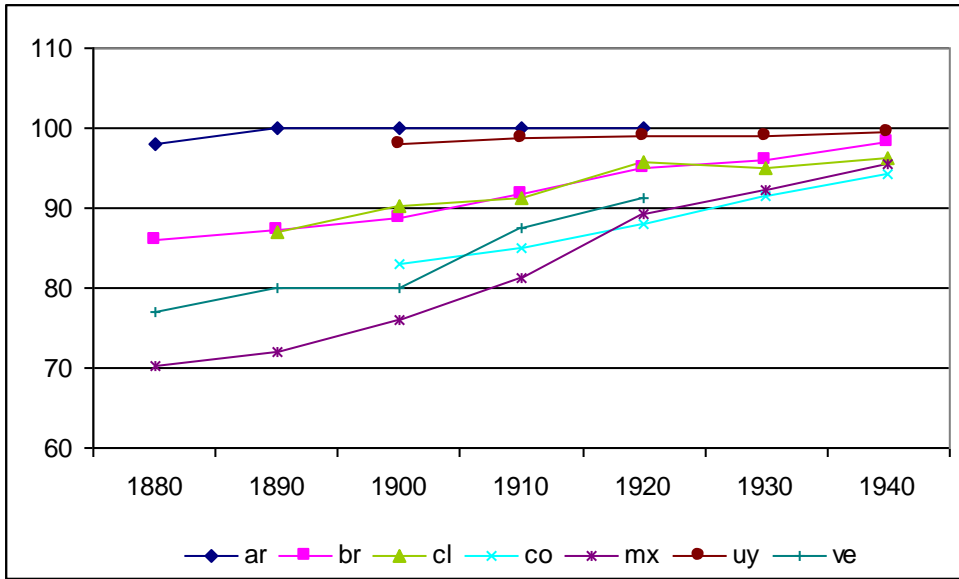
Figure 2c: Numeracy (ABCC, in %) in Latin America, group I
Note: Estimated values are included here. Abbreviations: see Appendix section "Census Data".
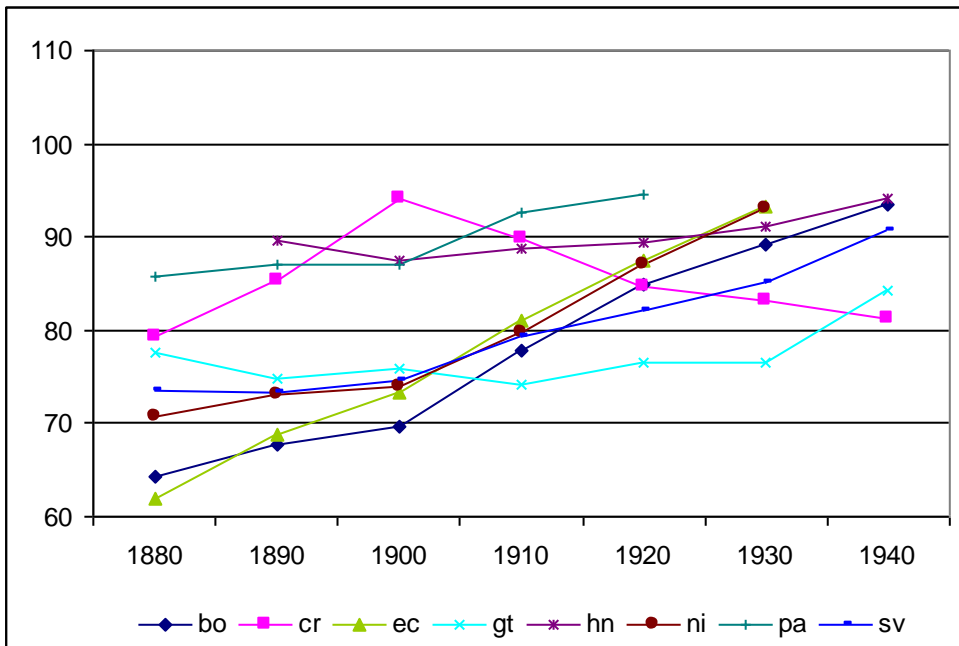
Figure 2d: Numeracy (ABCC, in %) in Latin America, group II
Note: Estimated values are included here. Abbreviations: see Appendix section "Census Data".
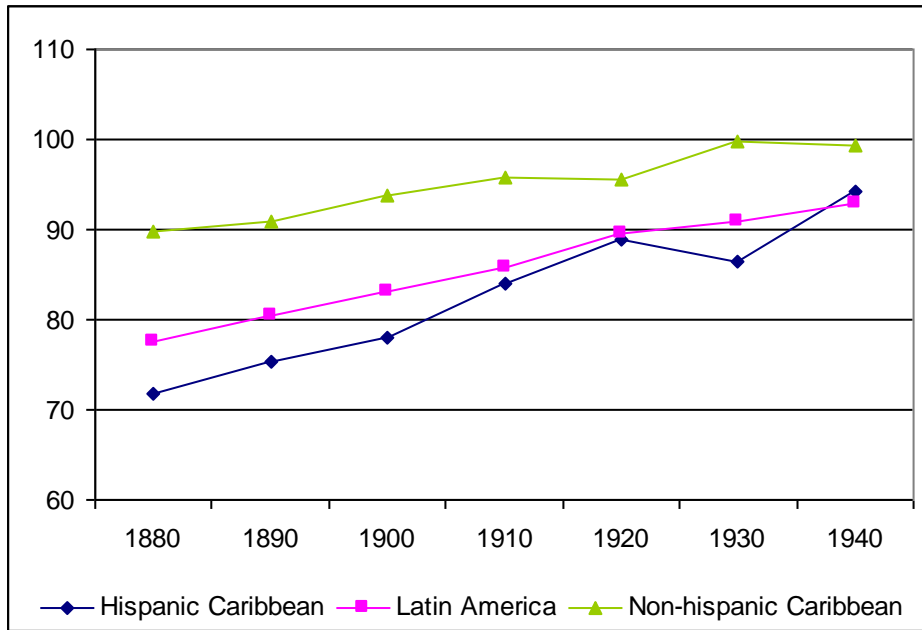
Figure 3: The Development of Numeracy in LAC
Note: estimations included, see Appendix.

Figure 4: Numeracy in Latin American and Caribbean Countries, 1900
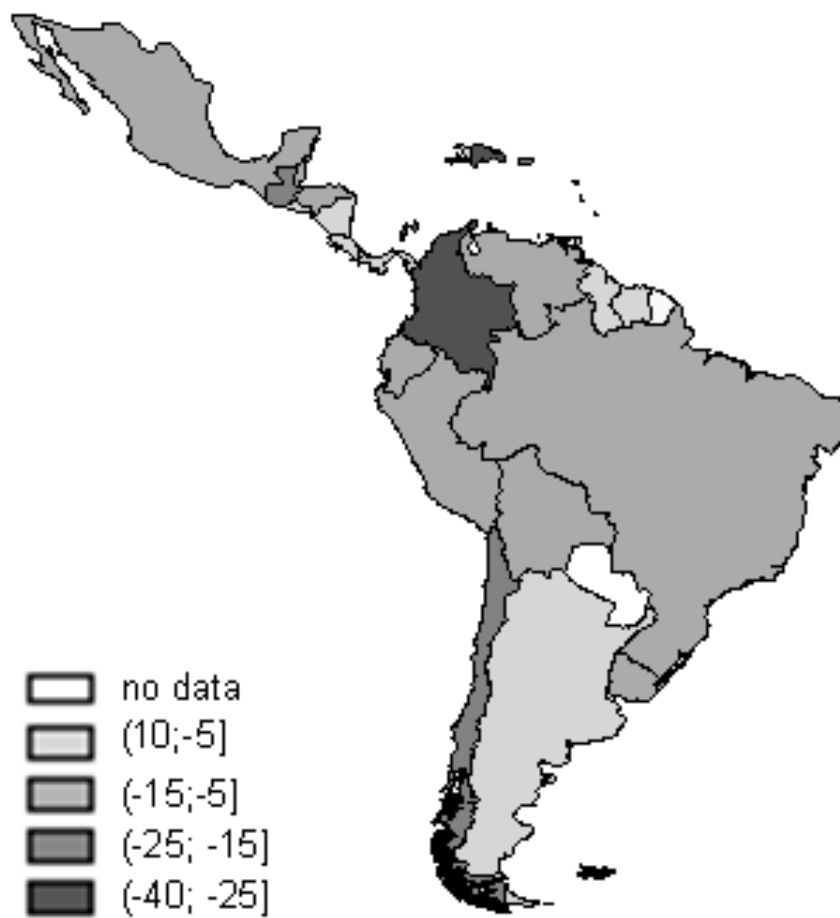
Figure 5: Gender Equality Indices in Latin American Countries, 1900
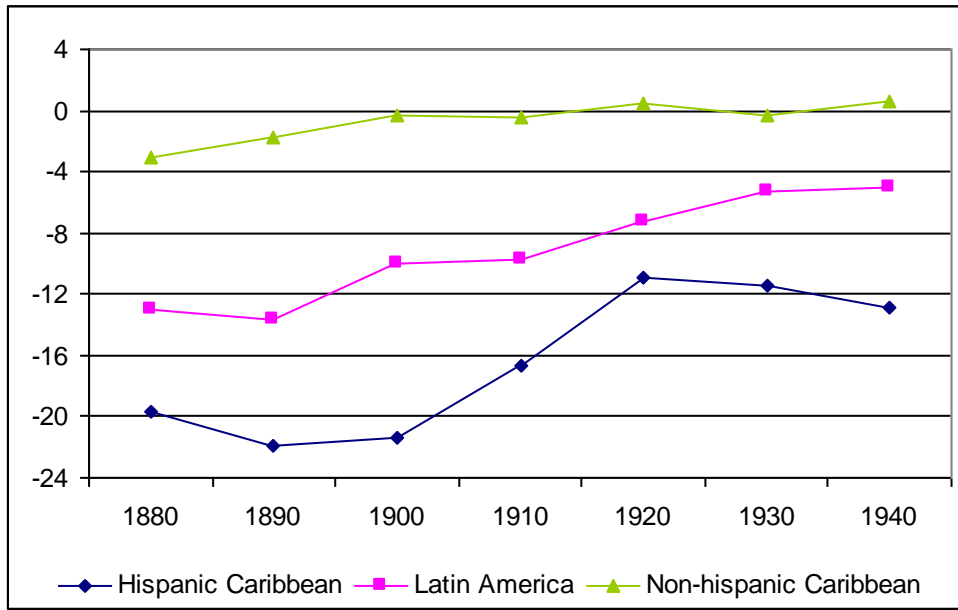
Figure 6: Gender Equality in LAC
Note: Estimated cases are included, see appendix Table A.1

Panel A. Ecuador
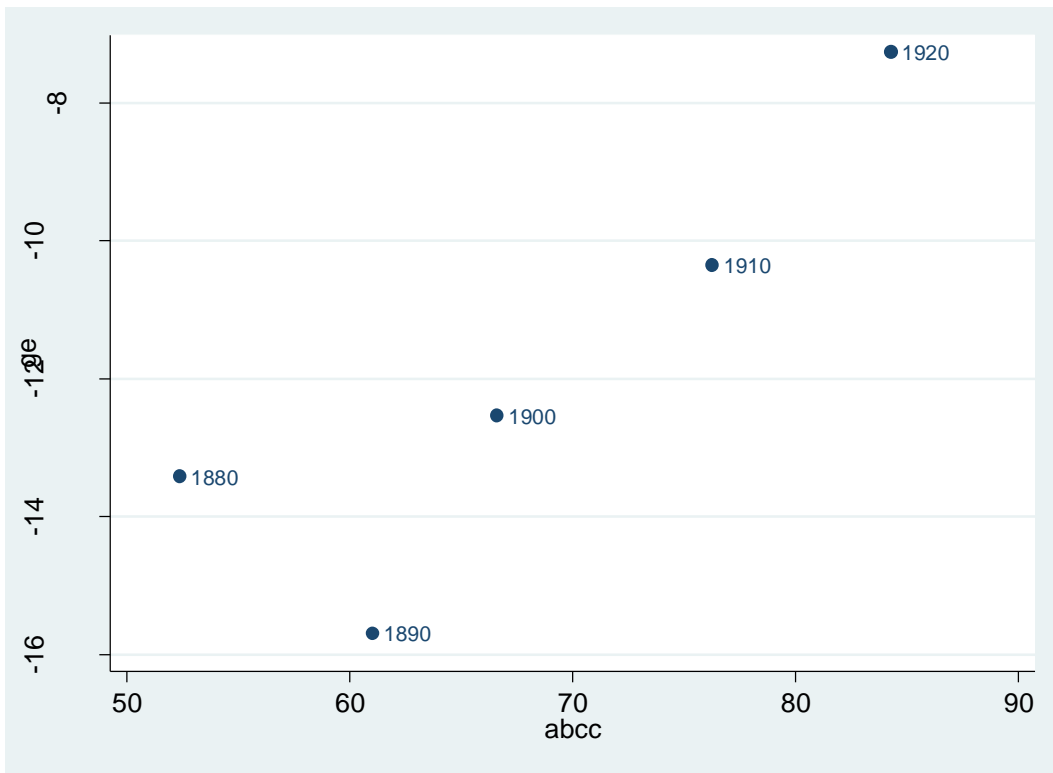
Panel B. Colombia
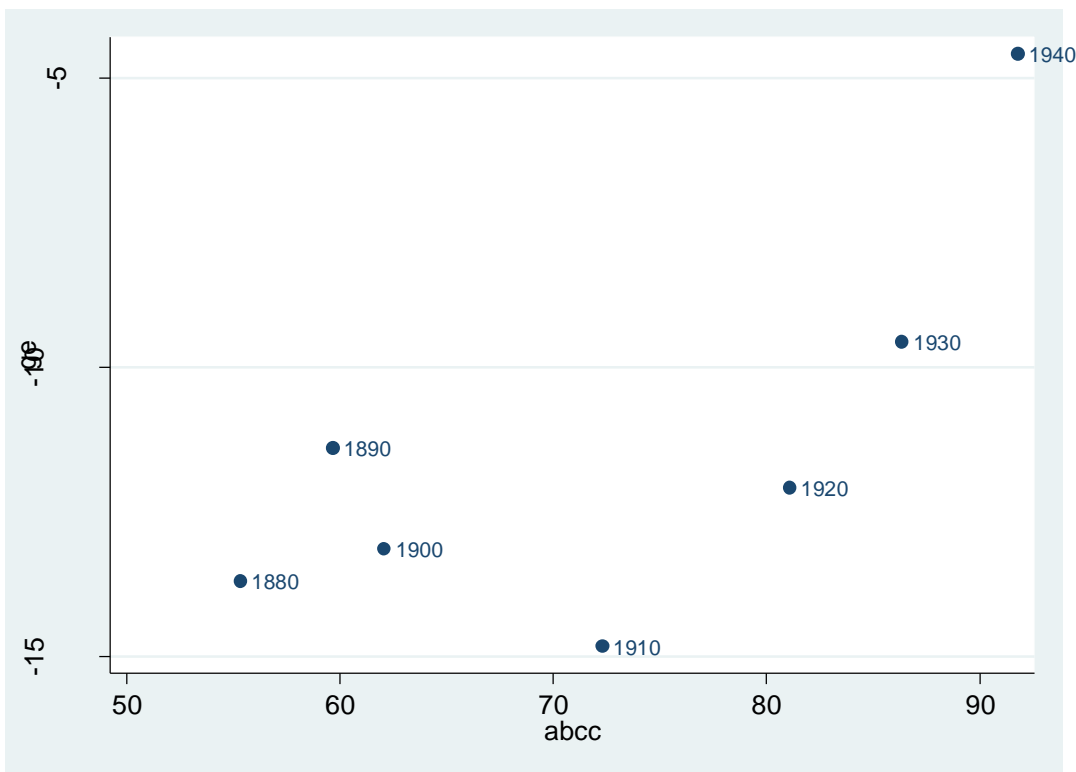
Figure 7: Gender equality and numeracy in Ecuador and Colombia
Note: for Ecuador, the 1930s is estimated based on the 1962 census.
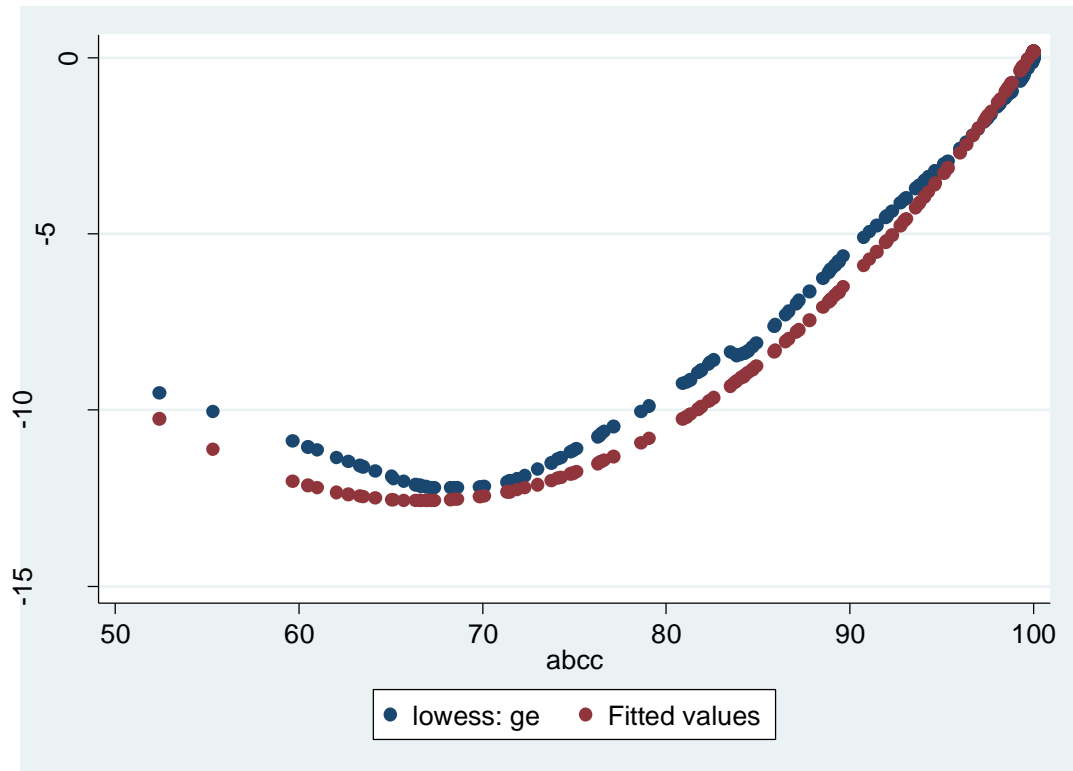
Figure 8: U-shaped pattern of gender equality in numeracy
Note: Estimated values are not used here.

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
ABCC	134	84.57	12.73	52.38	100
ABCC squared	134	7313.47	2078.71	2743.53	10000
Gender equality	134	-7.35	8.74	-49.65	7.41
Europ. immigration dummy incl. Brazil	134	.15	.36	0	1
Europ. immigration dummy	134	.11	.32	0	1
Spanish or Portuguese colony	134	.61	.49	0	1
Democracy	76	-.17	4.77	-9	10
Female voting rights	112	.04	.21	0	1
Indigenous > 60%	134	.16	.37	0	1
Under-enumeration	134	.19	.40	0	1
Pilot census	134	.33	.47	0	1

Notes: Estimated values are not included here. For sources, see Appendix .

Table 2: Regressions of Gender Equality in Latin America, 1880-1949

	(1)	(2)	(3)	(4)	(5)
Estimation technique	FE	LSDV	LSDV	IV	IV
If IV, which instr.?				Col	Col
				Im_BR	Im
ABCC	-1.10*	-1.16*	-1.17*	-0.75***	-0.72***
	(0.061)	(0.08)	(0.06)	(0.00)	(0.00)
ABCC sq.	0.01**	0.01**	0.01**	0.01***	0.01***
	(0.025)	(0.03)	(0.04)	(0.00)	(0.00)
Democracy	-0.03	-0.04	-0.15		
	(0.86)	(0.83)	(0.37)		
Female voting right	3.21*	3.20*	-0.03		
	(0.078)	(0.09)	(0.99)		
Indigenous	-4.81	-3.08	-18.51***		
	(0.14)	(0.40)	(0.00)		
Country dummies included?	(Fixed Effects) Y	Y	Y	N	N
Time dummies included?	N	N	Y	N	N
Constant	22.59	24.01	37.58*		
	(0.31)	(0.29)	(0.09)		
Observations	73	73	69	133	133
Number of countries	16	16	15	28	28
R-squared	0.25	0.84	0.89	0.27	0.29

Notes: Robust p values in parentheses. Only non-estimated observations included.
"Col" stands for former Spanish or Portuguese colony., "Im." for the main 3 immigration countries Argentina, Uruguay, Chile; "Im_BR" for the 4 immigration countries Argentina, Uruguay, Chile, and Brazil. Constants are omitted in Col.umn 4 and 5.

Table 3: Three Regressions excluding Cases of Underenumeration and Pilot Censuses

	(1)	(2)	(3)
Cases included:	All	No underenum.	No pilot census
ABCC	-1.68**	-3.50**	-2.39*
	(0.038)	(0.032)	(0.096)
ABCC sq.	0.01**	0.02**	0.02*
	(0.030)	(0.028)	(0.087)
Democracy	0.10	-0.49**	0.20
	(0.57)	(0.037)	(0.60)
Female voting right	3.24	3.08*	0.99
	(0.18)	(0.090)	(0.62)
Indigenous	-6.23*	-6.46	-4.96
	(0.051)	(0.19)	(0.22)
Constant	52.05*	121.05**	80.05
	(0.082)	(0.048)	(0.14)
Observations	73	47	38
R-squared	0.39	0.46	0.46

Robust p values in parentheses. Only non-estimated observations included.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix: Data Sources and Estimation Decisions

<u>**Census Data**</u>

1. Latin America
Argentina (ar): Census 1947 for birth decades 1880s-1920s; **Bolivia (bo):** Census 1950 for birth decades 1880s-1920s, birth decades 1930s-1940s estimated using growth rates of census 1976; **Brazil (br):** Census of 1950 for birth decades 1880s-1920s, birth decades 1930s-1940s estimated using growth rates of census 1970; **Chile (cl):** Census 1960 for birth decades 1890s-1930s, 1940s estimated using growth rates of 1970; **Colombia (co):** Census of 1964 for birth decades 1900s-1940s; **Costa Rica (cr):** Census of 1927 for birth decades 1880s-1890s, Census of 1950 for birth decades 1900s-1940s; **Ecuador (ec):** Census of 1950 for birth decades 1880s-1920s, 1930s and 1940s estimated using growth rates from census of 1962; **El Salvador (sv):** Census of 1950 for birth decades 1880s-1910s, 1920-40s estimated using growth rates from census 1971; **Guatemala (gt):** Census of 1950 for birth decades 1880s-1900s, 1910s-40s estimated using growth rates from census of 1973; **Honduras (hn):** Census of 1961 for birth decades 1880s-1930s, 1940s estimated using growth rates from census of 1974; **Mexico (mx):** Census of 1950 for birth decades 1880s-1920s, 1930s-40s estimated using growth rates from census of 1970; **Nicaragua (ni):** Census of 1950 for birth decades 1880s-1920s, 1930s estimated using growth rates from census of 1963 for birth decade 1930s; **Panama (pa):** Census of 1950 for birth decades 1880s-1920s, 1930s estimated using growth rates from census of 1960; **Peru (pe):** Census of 1940 for birth decades 1880s-1910s; **Uruguay (uy):** Census of 1975 for birth decades 1900s-1940s; **Venezuela (ve):** Census of 1950 for birth decades 1880s-1920s, 1930s/40s estimated using growth rates from census of 1961 for birth decades 1930s-1940s.

2. Hispanic-Caribbean
Dominican Republic (dr): Census of 1950 for birth decades 1880s-1910s, 1920s-40s estimated using growth rates from census 1976; **Puerto Rico (pr):** Census of 1950 for birth decades 1880s-1920s

3. Non-Hispanic Caribbean
Barbados (bb): Census 1946 for birth decades 1880s-1910s; **Belize (bz):** Census 1946 for birth decades 1880s-1910s; **Guadeloupe (gp):** Census of 1946 for birth decades 1900s-1940s; **Guayana (gy):** Census of 1946 for birth decades 1880s-1910s; **Haïti (ht):** Census of 1950 for birth decades 1880s-1920s; **Leeward Islands (lee):** Census of 1946 for birth decades 1880s-1910s; **Martinique (mq):** Census of 1967 for birth decades 1900s-1940s; **Surinam (sr):** Census of 1964 for birth decades 1890s-1930s; **Trinidad and Tobago (tt):** Census of 1950 for birth decades 1880s-1910s, census of 1970 for birth decades 1920s-1940s; **Windward Islands (win):** census of 1946 for birth decades 1880s-1910s.

<u>**Data sources:**</u>
Census Data: UN (various issues), *Demographic Yearbook*, New York, UN. For Costa Rica 1927: Centro Centroamericano de Población: http://ccp.ucr.ac.cr/bvp/censos/1927/index.htm, approached on August 25, 2008.
Indigenous population: http://www.integrando.org.ar/datosdeinteres/indigenasenamerica.htm
Women's suffrage: http://www.ipu.org/wmn-e/suffrage.htm, approached on August 25, 2008. We took only the unrestricted cases.
Democracy: Polity2 index, downloaded from http://www.systemicpeace.org/ polity/polity4.htm approached on August 25, 2008.

Countries with pilot censuses:

Chile, Costa Rica, Dominican Republic, Ecuador, Guatemala, Mexico, Nicaragua, Panama, Venezuela.

Countries with underenumeration:

Brazil, Costa Rica, Ecuador, Honduras, Peru

Appendix Table A.1: The data set

Abbreviations: see list of Census Data above.

co	censusyear	bdec	whall	whm	whf	ge	abcc	
ar	1947	1880	110	110	110	0.22	97	
ar	1947	1890	100	100	100	0.00	100	
ar	1947	1900	100	100	100	0.35	100	
ar	1947	1910	100	100	100	0.00	100	
ar	1947	1920	100	100	100	0.00	100	
bb	1946	1880	113	115	112	2.73	97	
bb	1946	1890	121	116	125	-7.28	95	
bb	1946	1900	115	113	116	-2.78	96	
bb	1946	1910	108	107	108	-0.61	98	
bz	1946	1880	126	123	129	-4.77	94	
bz	1946	1890	125	124	126	-1.58	94	
bz	1946	1900	116	111	120	-8.32	96	
bz	1946	1910	106	106	106	0.01	98	
bo	1950	1880	279	261	296	-13.69	55	
bo	1950	1890	261	247	275	-11.39	60	
bo	1950	1900	252	235	266	-13.14	62	
bo	1950	1910	211	195	224	-14.82	72	
bo	1950	1920	176	165	185	-12.08	81	
bo	estimated	1930	155	147	161	-9.56	86	
bo	estimated	1940	133	129	135	-4.58	92	
br	1950	1880	170	158	183	-15.60	82	
br	1950	1890	164	156	172	-9.91	84	
br	1950	1900	157	151	163	-8.20	86	
br	1950	1910	141	138	145	-4.70	90	
br	1950	1920	125	123	127	-3.19	94	
br	estimated	1930	120	120	120	-0.03	95	
br	estimated	1940	109	110	111	-1.13	98	
cl	1960	1890	165	130	195	-49.65	84	
cl	1960	1900	149	137	160	-16.52	88	
cl	1960	1910	143	134	150	-11.90	89	
cl	1960	1920	121	113	128	-13.86	95	
cl	1960	1930	125	123	126	-2.11	94	
cl	estimated	1940	119	117	121	-3.19	95	
co	1964	1900	186	164	206	-26.00	79	
co	1964	1910	175	158	191	-21.01	81	
co	1964	1920	160	148	172	-16.36	85	
co	1964	1930	142	135	150	-10.88	89	
co	1964	1940	128	120	138	-14.62	93	
cr	1927	1880	203	193	214	-10.51	74	

cr	1927	1890	173	171	175	-2.06	82
cr	1950	1900	129	127	131	-3.55	93
cr	1950	1910	151	147	155	-5.84	87
cr	1950	1920	176	172	180	-4.65	81
cr	1950	1930	184	180	188	-4.50	79
cr	1950	1940	194	185	203	-9.88	77
dr	1950	1880	240	210	270	-28.34	65
dr	1950	1890	226	201	256	-27.20	68
dr	1950	1900	214	191	241	-26.57	71
dr	1950	1910	192	171	214	-25.38	77
dr	estimated	1920	180	166	195	-17.62	80
dr	estimated	1930	167	157	175	-11.43	83
dr	estimated	1940	129	119	135	-12.90	93
ec	1950	1880	290	271	307	-13.42	52
ec	1950	1890	256	237	274	-15.69	61
ec	1950	1900	234	220	247	-12.53	67
ec	1950	1910	195	185	204	-10.35	76
ec	1950	1920	163	157	168	-7.27	84
ec	estimated	1930	134	122	145	-18.27	92
sv	1950	1880	232	220	244	-11.25	67
sv	1950	1890	234	227	240	-5.54	67
sv	1950	1900	227	223	231	-3.74	68
sv	1950	1910	204	202	205	-1.62	74
sv	estimated	1920	189	190	189	0.54	78
sv	estimated	1930	174	177	174	1.61	81
sv	estimated	1940	146	148	144	2.78	88
gp	1967	1900	102	106	100	5.45	99
gp	1967	1910	100	100	100	0.00	100
gp	1967	1920	101	100	105	-5.44	100
gp	1967	1930	100	100	100	0.20	100
gp	1967	1940	107	107	107	0.11	98
gt	1950	1880	212	185	243	-31.15	72
gt	1950	1890	225	199	254	-28.12	69
gt	1950	1900	220	197	243	-23.49	70
gt	estimated	1910	229	207	251	-21.13	68
gt	estimated	1920	218	201	236	-17.39	71
gt	estimated	1930	218	207	231	-11.69	71
gt	estimated	1940	178	170	189	-11.69	80
gy	1946	1880	124	121	126	-3.40	94
gy	1946	1890	136	135	137	-1.58	91
gy	1946	1900	120	119	121	-1.76	95
gy	1946	1910	111	110	111	-1.30	97
ht	1950	1880	258	254	261	-2.58	60
ht	1950	1890	237	236	238	-1.09	66
ht	1950	1900	247	252	241	4.65	63
ht	1950	1910	231	240	223	7.41	67
ht	1950	1920	193	201	187	7.01	77
hn	1961	1890	152	143	160	-12.44	87
hn	1961	1900	163	157	169	-7.91	84
hn	1961	1910	156	151	161	-6.87	86
hn	1961	1920	153	152	155	-2.19	87

hn	1961	1930	144	143	145	-1.85	89
hn	estimated	1940	130	130	131	-0.73	93
lee	1946	1880	131	125	134	-7.27	92
lee	1946	1890	134	131	136	-3.54	91
lee	1946	1900	124	122	125	-2.37	94
lee	1946	1910	109	108	110	-1.85	98
mq	1967	1900	105	104	105	-0.32	99
mq	1967	1910	100	100	100	0.29	100
mq	1967	1920	102	104	101	3.04	99
mq	1967	1930	100	100	100	-0.13	100
mq	1967	1940	102	104	101	2.26	99
mx	1950	1880	249	231	266	-15.11	63
mx	1950	1890	239	221	257	-15.90	65
mx	1950	1900	220	206	233	-13.26	70
mx	1950	1910	194	181	206	-13.52	77
mx	1950	1920	154	145	162	-11.50	87
mx	estimated	1930	139	134	143	-6.65	90
mx	estimated	1940	122	119	123	-3.05	95
ni	1950	1880	246	234	257	-9.75	63
ni	1950	1890	235	227	242	-6.36	66
ni	1950	1900	231	226	235	-3.83	67
ni	1950	1910	201	197	204	-3.56	75
ni	1950	1920	165	163	167	-2.50	84
ni	estimated	1930	135	137	137	-0.06	91
pa	1950	1880	171	167	175	-4.66	82
pa	1950	1890	165	166	163	1.76	84
pa	1950	1900	165	167	162	2.75	84
pa	1950	1910	137	134	140	-3.82	91
pa	1950	1920	128	125	130	-4.33	93
pa	estimated	1930	124	123	124	-1.14	94
pe	1940	1880	172	192	222	-15.50	82
pe	1940	1890	194	178	210	-18.01	76
pe	1940	1900	208	162	183	-12.84	73
pe	1940	1910	232	138	157	-13.54	67
pr	1950	1880	243	231	256	-11.11	64
pr	1950	1890	221	205	239	-16.55	70
pr	1950	1900	205	190	221	-16.30	74
pr	1950	1910	170	163	176	-8.01	83
pr	1950	1920	132	129	135	-4.13	92
sr	1964	1890	100	105	100	4.69	100
sr	1964	1900	103	104	101	2.71	99
sr	1964	1910	102	100	105	-5.15	99
sr	1964	1920	105	104	106	-1.41	99
sr	1964	1930	102	102	103	-0.66	99
tt	1946	1880	161	160	162	-1.02	85
tt	1946	1890	166	164	168	-2.60	84
tt	1946	1900	146	144	148	-2.61	89
tt	1946	1910	123	120	126	-5.07	94
tt	1970	1920	112	112	112	-0.59	97
tt	1970	1930	100	100	100	-0.48	100
tt	1970	1940	100	100	101	-0.70	100

uy	1975	1900	110	107	113	-5.33	98
uy	1975	1910	106	101	110	-9.01	99
uy	1975	1920	105	102	108	-5.91	99
uy	1975	1930	105	105	106	-1.76	99
uy	1975	1940	103	102	104	-2.06	99
ve	1950	1880	215	198	228	-15.16	71
ve	1950	1890	200	183	216	-17.90	75
ve	1950	1900	200	188	213	-13.20	75
ve	1950	1910	162	152	174	-14.60	84
ve	1950	1920	144	138	150	-8.40	89
ve	estimated	1930	134	129	138	-7.24	92
ve	estimated	1940	125	120	128	-7.24	94
win	1946	1880	145	140	147	-4.98	89
win	1946	1890	142	142	143	-0.77	89
win	1946	1900	132	133	131	2.02	92
win	1946	1910	119	120	117	2.19	95