Determinants of Long-Run Human Capital Formation in the Iberian World

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von

Frau Dácil Juif

aus Las Palmas de Gran Canaria

Tübingen

2014

Tag der mündlichen Prüfung:	10.07.2014
Dekan:	Professor Dr. rer. soc. Josef Schmid
1. Gutachter:	Professor Dr. Jörg Baten
2. Gutachter:	Professor Dr. Bernhard Boockmann

Acknowledgements

The support of several people to whom I would like to express my gratitude was essential for this thesis. First of all, I would like to thank my supervisor Prof. Dr. Jörg Baten for giving me the opportunity to write this dissertation and to be part of a wonderful research team. He encouraged me to present my research in numerous international conferences, to attend summer schools and to visit various archives, which was very fruitful for my work. He also gave me the opportunity to gain practice in teaching, a very valuable experience for me personally. I appreciated his constant professional feedback on my work and thank him for sharing his knowledge with me, I am very grateful to him. As well I want to thank Prof. Dr. Bernhard Boockmann for his willingness to review and evaluate this thesis as co-examiner.

The research group in Tübingen provided me with ideas, criticism and important feedback for my academic work. In particular, I want to thank Christina Mumme, Franziska Tollnek, Dr. Ralph Hippe, Carolina Vicario, Rima Ghanem, Nicholas Meinzer, Dr. Matthias Blum, Valeria Prayon, Dr. Yvonne Stolz, Dr. Mojgan Stegl, Dr. Linda Twrdek and Dr. Dominic Behle. They made my time at the chair of economic history in Tübingen extremely valuable.

Part of this thesis was produced within the HI-POD project funded by the European Commission's 7th Research Framework Programme and within the Clio Infra Project funded by the Netherlands Organization for Scientific Research (NWO). Their financial assistance is gratefully acknowledged.

Moreover, I would like to thank my family, especially my mum and dad, as well as my dearest friends who supported me during the last years and always. Their constant support and encouragement was invaluable throughout the progress of this thesis.

Despite the help of many people, I alone am responsible for the content of this doctoral thesis.

Contents

List of Tables

List of Figures

Symbols and Abbreviations

ABCC	Index that yields an estimate of the share of individuals who report a non-rounded age
AD	Anno Domini
GDP	Gross Domestic Product
IPUMS	Integrated Public Use Microdata Series
IV	Instrumental Variable
OECD	Organization for Economic Co-Operation and Development
OLS	Ordinary Least Squares Estimation
Wh	Whipple Index that reports age heaping
WLS	Weighted Least Squares Estimation

Country abbreviations according to DIN ISO 3166

1. Introduction

"The most valuable of all capital is that invested in human beings"

Alfred Marshall (1890)

1.1 The Role of Human Capital in Economics

Human capital can be defined as the stock of knowledge and skills that increase the productivity of the workforce. The term human *capital* derives from the fact that financial costs arise for the necessary training to acquire those skills and these investments in people generate future profitability for individuals and society. The highlight of the role of human capital in microeconomic as well as macroeconomic phenomena has led some important economists to define the 20th century and the beginning of the 21st century as the "era of human capital" (Becker 2002). However, the conceptualization was developed centuries ago (Kiker 1966). The most prominent economists who addressed issues related to human capital in their works were Adam Smith, John Stuart Mill and Alfred Marshall (Sweetland 1996). The microeconomic foundations of human capital theory were mainly laid by Theodore W. Schultz and Gary S. Becker in the sixties of the past century. They provided a theoretic framework to individual decisions on human capital investments (Schultz 1961), as well as to parental decisions on investing into the education of their children (Becker et al. 1960). In the last three decades, a vast empirical literature has been dedicated to the quantification of human capital and to the assessment of its effects on economic growth and wellbeing.

In which way does human capital contribute to economic growth? Since the emergence of *endogenous growth theory*, scholars generally agree on the stylized fact that human capital is a decisive factor explaining why some countries remain poor and others become rich (Cinnirella and Streb 2013). In the 1980s, a number of economists plead for human capital entering the aggregate production function in the neoclassical growth model of Solow (e.g. Romer 1986, Lucas 1988 and Rebelo 1991). These models aimed to *endogenize* technology, which had been treated as exogenously given in neoclassical economic theory. Mankiw, Romer and Weil (1992) tested this theory empirically by including human capital in their cross-country growth regressions. Human capital, calculated as the fraction of the population aged 15 to 19 enrolled in secondary school, was shown to increase the significance of their model considerably. The result was that similar countries in terms of savings, population growth and other characteristics might not converge in terms of economic growth due to differences in education.

Around the millennium, economists tried to explain the very long run development and included explicitly a historical dimension in their growth models. *Unified Growth Theory* argued that Europe and some of the former colonies were freed from the Malthusian trap[1] due to an increase in the technologically-driven demand for human capital and the simultaneous fertility decline (Galor 2005, 2010).The growth model proposed by Galor and Weil (2000), with human capital at its core, builds largely on the microeconomic foundations of Theodore W. Schultz and Gary S. Becker, the pioneers of “human capital theory”. According to this model, human capital is a complement to technology, since it is a necessary tool for the capability to innovate and, especially for emerging countries, to adopt foreign technology.

The roots of global disparities in human capital go well back into historical times. Joel Mokyr (2009) believes that “the Industrial Revolution was the outgrowth of the social and intellectual foundations laid by the Enlightenment and the Scientific Revolution”. Thus, he explains the growth explosion in the modern west, more specifically in England,

[1] The Malthusian trap is named after the economist Thomas Robert Malthus. He suggested that for most of history, income was largely stagnant because technological advances only resulted in an increase of the population, not in improvements in living standards. This changed with the Industrial Revolution when per capita income increased dramatically.

as a byproduct of the improved access to knowledge, as made possible, for example, through universities. Baten and van Zanden (2008) show that human capital, measured by per capita book production as a proxy for advanced literacy skills, can explain differences in economic growth long before industrialization. However, different opinions exist on whether human capital was already a driving factor of economic growth before the Industrial Revolution. Allen (2003) and Mitch (1993), for instance, argued that during the first phase of the Industrial Revolution in England, the basic school system was not yet implemented and literacy levels even stagnated. Thus, in their view, human capital did not contribute to the onset of industrialization. It was rather during the further advancement of the economy that education became increasingly important because of the demand for basic skills in new occupations. This need of a skilled workforce made basic education more widely available since the mid-19th century, and the state took over the responsibility for it from the church.

Apart from the direct effect on economic growth, human capital has been claimed to improve the quality of institutions (e.g. by Lipset 1960). First of all, human capital is needed in order for courts to operate and secure people's rights. Moreover, educated people are more prone to solve their problems by discussion and voting than through violence (Swanson and King 1991). Education makes individuals also better aware of the malfeasance of governments and more likely to try to change the situation. Empirically, this link has been tested by Glaeser et al. (2004), Alvarez et al. (2000) and Barro (1999). That "good institutions" are a very important driving factor for economic growth is the opinion of a number of important scholars (e.g. Baten and van Zanden 2008; Acemoglu et al. 2005; Easterly and Levine 2003).

Moreover, human capital differences across regions, genders, ethnic or religious groups reflect the inequality of opportunities. Income inequality is often a direct consequence of the inequality of opportunities which arise from an unequal access to education. If one part of the population is deprived of education this means that resources are probably dislocated, meaning that not the most talented receive the best opportunities and positions, which is in turn detrimental for economic growth. Another link of human capital to wellbeing derives from gender equality in education. Depriving women from education can have negative consequences for wellbeing since better educated women tend to improve nutrition levels and prevent illnesses in their families, thereby reducing child mortality (see Klasen 2002; Manzel and Baten 2009).

1.2 Methodological Aspects: Measurement of Human Capital

Although human capital investments include health and nutrition (Schultz 1981), generally education is considered the prime human capital investment in empirical studies. The main reason for using education to approximate human capital is that it improves health and nutrition and that it is better quantifiable (Mincer 1974). Apart from formal education, also informal education at home and at work (Schultz 1981) and on the job training and apprenticeships (Mincer 1974) play a role. However, the most common measures of human capital for today are based mainly on formal education. They include the average years of schooling (Barro and Lee 1996), enrollment rates (Barro 1991; Mankiw, Romer, and Weil 1992), literacy (Romer 1989), and educational attainment (Barro and Lee 1993). In this thesis I also use a measure of cognitive skills for the late 20th century that was constructed by Hanushek and Woessmann (2012) out of the mean results of PISA tests in mathematics and sciences in the period of 1960 to 2000. Nevertheless, most of these measures cannot be used in a historical perspective since data are not available. Therefore, the most important historical measures of human capital are literacy, sometimes approximated by signature ability, and more recently, numeracy. Other measures that have been used by economic historians include book production (Baten and van Zanden 2008), number of patents

(Labuske and Baten 2006, Baten et al. 2007), school enrolment ratios (Nehru et al. 1995) or student-teacher ratios.

1.2.1 Literacy

Literacy is the most popular human capital indicator for historical times. The knowledge of writing has a very long tradition in the western world, but for most of history it was only accessible to the elites of society. Still in 1750, more than 90% of the worldwide population was not able to write (Cipolla 1969). Thus, the majority of the population was excluded from literacy. It was with the introduction of universal formal schooling that literacy spread in Europe. Some censuses include the self-reported information on whether an individual can read or write. However, such sources are only available from the mid-19th century for part of the world. Lacking this sort of data, historians have often recurred to the signature ability as a proxy for literacy. People were sometimes required to sign their name on legal documents such as marriage records or wills. Those who could not write and therefore could not sign their names would make a cross instead. Jaime Reis (2005) presented a comparative survey of the signature ability for late 18th century Europe. The author himself acknowledges the disadvantages of this measure. For example, it is not always possible to discern if the person herself signed the document. Also, some people might have learned to sign their name but could not write. Literacy mostly does not allow any gradations. Individuals could have been able to read but not to write in many cases, since for a long time students learned one skill after the other (Cohen 1982). Moreover, as Woessmann (2003) puts it, using literacy as a proxy for human capital neglects educational investments made on top or instead of basic literacy, such as the acquisition of numeracy, technological knowledge or logical and analytical reasoning.

1.2.2 Numeracy

Numeracy can be expressed as “the ability to count, keep records of these counts, and make rational calculations” (Emigh 2002). A steep rise in numeracy took place in the early 19th century western world. Market capitalism furthered numeracy, since being able to calculate became an increasingly important skill for trade and finance. Numerical skills can be influenced by additional factors such as state bureaucracy. Population statistics in form of censuses or tax assessments carried out by the government may oblige individuals to keep records. However, the most important driving factor for numeracy has been claimed to be formal education (A’Hearn et al. 2009). Numeracy is much more understudied than literacy, also because it was long thought to be unmeasurable (Cohen 1982).

However, recently an increasing literature has used the age heaping technique to measure numerical skills (A’Hearn et al. 2009). This measure of human capital is the most frequently used in this thesis. The age-heaping phenomenon applies to historical populations as well as to people in the poorest countries today, in which a substantial share of the people was unable to state their exact age and hence reported a rounded age. The preferred numbers of “heapers” are multiples of five, such as “30” or “45”. Age heaping is calculated as the ratio between the preferred ages and others. It can be measured by several indices, the most common of which is the Whipple index. This index reports the proportion of people who state an age ending in five or zero, assuming that each terminal digit should appear with the same frequency in the “true” age distribution. Usually only individuals in the age range of 23 to 62 are included because for the younger ones it is more probable that another household member reported the age and the elder persons tend to exaggerate it. A Whipple index value of 500 means an age distribution with ages ending only on multiples of five, whereas the value of 100 indicates no heaping patterns on multiples of five; that is exactly 20 percent of the population reported an age ending in a multiple of five. The formula to calculate the Whipple index is the following:

$$(1)\ Wh = \left(\frac{(Age25 + Age30 + Age35 + ... + Age60)}{1/5 * (Age23 + Age24 + Age25 + ... + Age62)} \right) \times 100$$

For an easier interpretation, A'Hearn, Baten, and Crayen (2009) introduced the so-called ABCC index. The name resulted from the initials of the authors' last names plus Gregory Clark's, who suggested this transformation of the Whipple index in a comment on their paper. The ABCC index yields an estimate of the share of individuals who correctly report their age:

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

The share of persons able to report an exact age has been shown to be highly correlated with other measures of human capital, such as literacy and schooling (Mokyr 1983; A'Hearn et al. 2009; Crayen and Baten 2010). A'Hearn, Baten, and Crayen (2009) found a close correlation in less developed countries between illiteracy and age heaping. The correlation coefficient with illiteracy was as high as 0.7. The correlation with the PISA results for mathematical abilities was even higher. They conclude that the Whipple index is more closely correlated with mathematical abilities than with literacy. These authors also tested whether the extent of bureaucracy and government statistics are predictors of individuals' awareness of their own exact age. It turned out that except for countries with a very long tradition of census taking the number of previous censuses had no significant impact, meaning that the independent effect of bureaucracy is rather weak.

The age heaping technique has the advantage of being an outcome measure of education and therefore does not suffer from the shortcomings of input measures such as, for example, enrollment ratios or years of schooling. Those input measures do not account for the fact that the quality of education varies substantially between countries or regions. A further critique point of enrollment ratios is that an enrolled child may never attend school, but this escapes the statistics as well. Scholars who use years of schooling as a measure of education mistakenly assume that every year of schooling increases the human capital stock by an equal amount (Woessmann 2003). Another advantage of the age heaping methodology is that age statements are more widely available than other human capital proxies, such as signature ability or school attendance. The most common source used to capture age heaping are census lists, but also marriage registers and mortality records have recently been taken into account for this purpose. In this thesis age statements made at Inquisition courts are also used to calculate numeracy for an early period before the spread of government statistics.

The age heaping method is a measure of very basic skills (Reis 2008). For most of the western world, it cannot provide information on education for after the late 19th century, but it is a useful measure for Latin America in the period under study in some chapters of this thesis.

1.3 Aim of this Thesis

The ambitious aim of this thesis is to contribute to the assessment of the genesis of wellbeing and help understand the divergent development paths of different countries and world regions in the last five hundred years. Its focus lies on the analysis of human capital formation, an important component of wellbeing, in the Iberian world. Although a lot has been done in the last decade regarding the quantification of human capital in historical times, there are still regions and time spans to be covered.

An important contribution of this doctorate is the construction of new micro data bases with information on human capital and other social and geographic variables for the Iberian world (and beyond) that will allow for future analyses and testing of economic models. The construction of new databases is still an important task for economic historians. It is tedious work and often requires long hours transcribing information in archives and typing data, but the availability of good quality data is an important prerequisite for performing empirical analyses that encompass a longer time frame than the last half century.

Apart from the quantification of human capital in new world regions, this work contributes to the analysis of educational differences between ethnic and religious groups as well as to the assessment of the determinants of long-run human capital formation. It addresses important topics related to human capital that can be relevant also from today's perspective. For example, how persistent is human capital and what influences human capital formation in the long run? Does the culture and religion with which individuals are raised affect their educational investments? Since Jews have always had a significant educational advantage over other religious groups in the Diaspora, what can the expulsion of this religious minority imply? Did European immigration contribute to raise human capital levels in Latin America?

1.4 Outline of the Thesis

This thesis comprises six chapters of which four (chapters two, three, five and six) are intended for publication. I therefore refer to the respective chapter as paper. At the time this thesis is completed, two of the papers are published and one of them is submitted for publication.

Chapter two assesses a potential "pre-colonial legacy" hypothesis for the case of the Andean region by studying the basic numeracy skills of Inca Indios. This hypothesis holds that exploitative institutions that were already present in colonies before the European conquest and persisted well into colonial times, such as the "mita"-system, could be blamed for the slow growth path of some regions. The age heaping technique is applied to population lists carried out by the Spanish administration in Peru soon after the Conquest. It allows to trace back human capital levels at the time before the Spanish invasion of the Inca Empire. Numeracy values for Peruvian Inca Indios of the cohort born before the

Conquest are close to zero and evidence on inequality in pre-Columbian times is found. The low educational level and the high inequality reigning before the Spanish Conquest in the Andean region could be a hint of the pre-colonial legacies as the genesis of the long term path of only modest economic growth in Latin America.

Human capital can be influenced by exogenous variables (not economic incentives) such as religious rules or values (Botticini and Eckstein 2007, Becker and Woessmann 2009).[2] The Jewish is one religious group that is well-known for their immoderate investments into education. Chapter three assesses Jews' educational advantage over the Catholic majority in the Iberian Peninsula and Latin America, and addresses potential (historical) reasons for their educational superiority. An innovative source, Inquisition trial documents, is used to assess the human capital of various religious groups with the age-heaping method from the 1450s until the 1850s. Chapter number four is a direct continuation of the precedent study that assesses the educational levels of Jews and other religious groups in Latin America. Other more advanced measures of human capital such as years of schooling and attainment rates are used in this chapter, since this chapter encompasses the early 20th century.

Selective immigration is one potential determinant of long-run human capital formation in a country that is addressed in chapter five. This chapter aims to shed more light on the skill selectivity of European migrants to Cuba, Uruguay and Argentina during the age of mass migration as well as on their comparative human capital advantage relative to natives in those destination countries. Using new micro data compiled from passenger lists and census data this study analyzes the numeracy of immigrants with the use of the age heaping technique. The results indicate that the numeracy level of immigrants relative

[2] Religious rules that promote literacy, for example, are considered of exogenous character because they were originally intended for the devotees to be able to read religious texts and were not an educational investment decision for profit.

to natives in the destination countries varies across destinations. Compared to the population of origin, the analysis shows that the selectivity with respect to the sending region was negative for emigrants from the Canary Islands and emigrants from the Spanish mainland, whereas no particular selection applied for Spanish emigrants to Cuba.

Chapter six is not a regional study of the Iberian world, but an international cross-country analysis. Whether a government wants to invest into education or not, can be determined by the influence and the preferences of powerful elites. This chapter addresses the question of whether inequality hampers human capital accumulation in the long-run. At times in which land was the most important asset, the landed elite had the power to influence public decisions on educational investments. Large land owners were probably not interested in the introduction of public universal schooling since education does not bring high returns in the agricultural sector. Thus, a landed elite avoided paying taxes for public schools. Moreover, they could risk that their laborers left their estates to seek work in other industries where education is better paid off. The cross-country analysis shows that inequality in the 19th century, especially land concentration, had a substantial negative impact on present educational outcomes, even after controlling for other potential determinants of long-run human capital formation. Chapter seven concludes this thesis.

1.5 References

Acemoglu, D., Johnson, S., and Robinson, J. (2005). The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. *The American Economic Review*, 95(3): 546-579.

A'Hearn, B., Baten, J., and Crayen, D. (2009). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. *The Journal of Economic History*, 69(3): 783-808.

Allen, R.C. (2003). Progress and Poverty in Early Modern Europe. *Economic History Review*, 56(3): 403-443.

Alvarez, M., Cheibub, J. A., Limongi, F., and Przeworski, A. (2000). Democracy and Development: Political Institutions and Material Well-Being in the World, 1950-1990. Cambridge: Cambridge University Press.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, 106(2): 407-443.

Barro, Robert J. (1999). Determinants of Democracy. *Journal of Political Economy*, 107(6-2): 158-183.

Barro, R. J., and Lee, J. W. (1993). International Comparisons of Educational Attainment. *Journal of Monetary Economics*, 32(3), 363-394.

Barro, R. J., and Lee, J. W. (1996). International Measures of Schooling Years and Schooling Quality. *The American Economic Review*, 86(2): 218-223.

Baten, J., Spadavecchia, A., Streb, J., and Yin, S. (2007). What Made Southwest German Firms Innovative Around 1900? Assessing the Importance of Intra-and Inter-industry Externalities. Oxford Economic Papers, 59(suppl 1), i105-i126.

Baten, J.,and Van Zanden, J. L. (2008). Book Production and the Onset of Modern Economic Growth. *Journal of Economic Growth*, 13(3): 217-235.

Becker, G. S., Duesenberry, J. S., and Okun, B. (1960). An Economic Analysis of Fertility. In: *Demographic and Economic Change in Developed Countries*. Columbia University Press, 209-240.

Becker, G. S. (2002). The Age of Human Capital. In: *Education in the Twenty-First Century*, edited by E. P. Lazear. Palo Alto: Hoover Institution Press, 3-8.

Benhabib, J., and Spiegel, M. M. (1994). The Role of Human Capital in Economic Development - Evidence from Aggregate Cross-Country Data. *Journal of Monetary Economics*, 34(2): 143-173.

Botticini, M., and Eckstein, Z. (2007). From Farmers to Merchants, Conversions and Diaspora: Human capital and Jewish History. *Journal of the European Economic Association*, 5(5): 885-926.

Cinnirella, F., and Streb, J. (2013). The Role of Human Capital and Innovation in Prussian Economic Development. CESifo Working Paper (No. 4391).

Cipolla, C. M. (1969). *Literacy and Development in the West*. Baltimore: Penguin Books.

Cohen, P. C. (1982). *A Calculating People: The Spread of Numeracy in Early America*. Chicago: University of Chicago Press.

Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its implications for long-term growth. *Explorations in Economic History*, 47(1): 82-99.

Easterly, W., and Levine, R. (2003). Tropics, Germs, and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics*, 50(1): 3-39.

Emigh, R. J. (2002). Numeracy or Enumeration? The Uses of Numbers by States and Societies. *Social Science History*, 26(4): 653-698.

Galor, O. (2005). From Stagnation to Growth: Unified Growth Theory In: Handbook of Economic Growth, Vol. 1A, edited by P. Aghion. and S. N. Durlauf. Amsterdam: North Holland, 171-293.

Galor, O. (2010). The 2008 Lawrence R. Klein Lecture – Comparative Economics Development: Insights from Unified Growth Theory. *International Economic Review*, 52:1-44.

Galor, O., and Weil, D. N. (2000). Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review*, 90(4): 806-828.

Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004). Do Institutions Cause Growth?. *Journal of Economic Growth*, 9(3): 271-303.

Hanushek, E. A., and Woessmann, L. (2012). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *Journal of Economic Growth*, 17(4): 267-321.

Kiker, B.F. (1966). The Historical Roots of the Concept of Human Capital. *Journal of Political Economy*, 74(5): 481-499.

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): 345-37.

Labuske, K., and Baten, J. (2006). Patenting Abroad and Human Capital Formation. Under review at Cliometrica.

Lipset, S. M. (1960). Political Man: The Social Basis of Modern Politics. New York: Doubleday.

Lucas, R. E. Jr. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(2):3-42.

Mankiw, N.G., Romer, D., and Weil, D.N. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107 (2): 19-36.

Mincer, J. (1974). *Schooling, Experience, and Earnings.* New York: Columbia University Press.

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*, 27(1): 37-74.

Marshall, A. (1890). *Principles of Economics*. London: Macmillan and Company.

Mitch, D. (1993). The Role of Human Capital in the First Industrial Revolution. In: *The British Industrial Revolution: An Economic Perspective*, edited by J. Mokyr. Boulder: Westview Press, 267-307.

Mokyr, J. (1983). *Why Ireland Starved.* London: George Allen and Unwin, Chap. 8.

Mokyr, J. (2009). *The Enlightened Economy*. New York/London: Yale University Press.

Nehru, V., Swanson, E. and Dubey, A. (1995). A New Data Base on Human Capital Stock: Sources, Methodology, and Results. *Journal of Development Economics*, 46(2): 379-401.

Rebelo, S. (1991) Long Run Policy Analysis and Long Run Growth. *Journal of Political Economy*, 99(3): 500-521.

Reis, J. (2005). Economic Growth, Human Capital Formation and Consumption in Western Europe before 1800. In: *Living Standards in the Past*, edited by R. C. Allen, T. Bengtsson and M. Dribe. Oxford: Oxford University Press, 195-225.

Reis, J. (2008). How Helpful is Age Heaping? An Exploration Using Portuguese Data. Univ. of Lisbon working paper.

Romer, P. M. (1986). Increasing Returns and Long Run Growth. *Journal of Political Economy*, 94(5):1002-1037.

Romer, P. M. (1989). Human Capital and Growth: Theory and Evidence (No. w3173). National Bureau of Economic Research.

Schultz, T. W. (1961). Investment in Human Capital. *The American Economic Review*, 51(1): 1-17.

Schultz, T. W. (1981). *Investing in People: The Economics of Population Quality*. Los Angeles: University of California Press.

Swanson, A. D., and King, R. A. (1991). *School Finance: Its Economics and Politics*. New York: Longman.

Sweetland, S. R. (1996). Human Capital Theory: Foundations of a Field of Inquiry. *Review of Educational Research*, 66(3): 341-359.

Woessmann, L. (2003). Specifying Human Capital. *Journal of Economic Surveys*, 17(3): 239-270

2. On the Human Capital of Inca Indios before and after the Spanish Conquest. Was there a "Pre-Colonial Legacy"?

Abstract:

Not only the colonial period, but also the pre-colonial times might have influenced later development patterns. In this study we assess a potential "pre-colonial legacy" hypothesis for the case of the Andean region. In order to analyze the hypothesis, we study the human capital of Inca Indios, using age-heaping-based techniques to estimate basic numeracy skills. We find that Peruvian Inca Indios had only around half the numeracy level of the Spanish invaders. The hypothesis holds even after adjusting for a number of potential biases. In addition, we find evidence on inequality in pre-Columbian times. Given the low educational level and the high inequality reigning before the Spanish Conquest in the Andean region, we argue that more attention should be paid to the pre-colonial legacies when assessing the genesis of the long-term path of only modest economic growth in the countries of Latin America.

This chapter is based on a paper published in *Explorations in Economic History*, co-authored with Jörg Baten (University of Tübingen). The concept of the paper was developed jointly, analyses and writing were equally shared.

2.1 Introduction

One commonly accepted view on the genesis of Latin America's economic backwardness blames mainly the legacy of colonial institutions for today's underdevelopment (Engerman and Sokoloff 2000, Engerman et al. 1999, Acemoglu et al. 2001, 2002, Coatsworth 1998, 2008, Bulmer-Thomas et al. 2006). Acemoglu et al. claim that bad institutions originated in areas in which small European elites exploited large populations of native or African descent. These so-called "extractive institutions" – favoring the concentration of power by a small fraction of the population, hampering property and human rights, and restricting public investment in schools and other growth-inducing infrastructure – have tended to persist until today, hindering GDP growth. Engerman and Sokoloff argued that geography and initial factor endowments at the time of the conquest – both natural resources and labor supply (the latter determined by population density) – defined the later inequality of the colonies. Where climate and soils were suitable for cultivating highly valued commodities and indigenous population was dense, they argue, extractive institutions were more likely created in order to facilitate exploitation in form of large plantations or cattle and grain haciendas. This was the case in the Andes and Mexico, for example.

While we agree with the role of colonial development obstacles, we argue in this study that there might also have been an additional legacy of pre-colonial societies for economic success. Especially low human capital investment in the Inca Empire might have initiated a path-dependent process of agriculture which was not human-capital intensive in the following centuries. Recently, Comin, Easterly and Gong (2010) constructed a measure of technological adoption at a regional level reaching back to 1500 AD. Applied to Latin America and specifically to Peru, they find that the Inca Empire not only lagged far behind most of Eurasia and Northern Africa, but was also outperformed by the Aztecs in terms of the utilization of a set of basic technologies (communications, agriculture, military, industry and transportation). Using this measure they report a positive correlation between pre-colonial technology history and per capita income today. Thinking of human capital as a core determinant of economic development, in this paper we contribute to the literature on the genesis of underdevelopment in the Andean region by constructing a data set that permits us to measure pre-colonial educational levels and compare them with European and Asian values. We also argue that educational inequality was substantial in Peru before the Spanish conquest.

How can human capital of the "Inca" Indios be measured?[1] Obviously, no school enrollment or literacy rates are available for this early period of human history. However, a considerable number of studies have recently used an innovative measure of basic numerical skills, which takes the share of people who report an exact age as an indicator. In historical populations (as well as in the poorest countries today), a substantial share of the people was not able to state their exact age and hence reported a rounded age, such as "I am 40", when they were in fact e.g. 39 or 41. In hundreds of samples and dozens of studies, it has been proven that there is a significant correlation between the share of exact reported ages and other human capital indicators (see section 2 and Appendix A for longer methodological treatment). For example, one population register on a remote Andean region (Huanuco) contains both age statements of Indios born before the Spanish conquest and thereafter. Another register of Indios in the capital of Peru reflects the birth cohorts of the later 16th century and is therefore an interesting source for the history of numeracy under Spanish colonial rule. We also study a census of the population of Lima born during the 17th century, which allows us to perform a comparison between urban Peruvians of indigenous, European, black and mixed ancestry.

[1] "Inca" Indios are used here in citation marks, because the Inca Empire consisted of nearly 200 tribes, and only the family of the ruler was called Inca. But as the Habsburg Empire was named after the ruling family, it seems reasonable to name the whole group "Inca Indios". Hence, we will drop the quotation marks in the following.

Of course, our study faces a number of challenges. Cultural counting differences, selective mortality, and other human capital components apart from numeracy are all potential lacunae of our research design and hence will be discussed in detail below. But before doing so, we will give a brief history and chronology in section 1, discuss our sources and the methods of basic numeracy measurement in section 2, and present the main results in section 3. After the "Potential Objections section" we will dig slightly deeper and assess social and regional differences in section 5, before ending with a conclusion.

2.2 History and Chronology

Sources used by historians writing about pre-Columbian Peru comprise the legacy from officers of the Crown – like the *visitas* from which we derive information for our dataset, often conducted on the basis of detailed questionnaires and including remarks of the visitor –, Spanish travelogues and native chronicles.[2] Most important travelogues of conquistadors told from a Spanish point of view include Pedro Pizarro's (brother of Francisco Pizarro). Other authors like Pedro de Cieza de León were more interested in the Indian world and cross-referenced their evidence with information provided by natives (Esteban 1997; p.112). Works from mestizo and Peruvian writers such as Garcilaso de la Vega and Felipe Guaman Poma de Ayala, both descendants of the Inca nobility, brought up in close contact to Quechua culture and language, provide a valuable insight into the Inca history and culture, though probably also positively exaggerating their advance (Esteban 1997; p.118).

The date of creation of the Inca Empire in the city of Cuzco is assigned to the early thirteenth century, about three hundred years before the Spanish arrived in Peru (Bakewell

[2] As the Incas knew no writing, their oral tradition was written down in the early colonial time often in a mixture of Quechua and Spanish language (Wachtel 1977, p. 5). Members of the intellectual elite soon adopted cultural opportunities as the written language, offered by the interaction with the Spanish.

2004; p. 25).[3] At the time of its greatest expansion, the beginning of the 16th century, the Inca Empire included today's Ecuador, Peru, Bolivia and a large part of Chile, as well as smaller territories in Argentina and Colombia.[4] The Spanish conquistador Francisco Pizarro reached Inca territory arriving from Panama in 1526. Three years later he obtained the permission from the Spanish crown to conquer the region and become governor of New Castile –as the administrative unit reaching from Ecuador in the north to Cuzco in the south would be called until 1542. At that time, the smallpox epidemic that had already devastated the population in the Caribbean and Mexico, had also reached Inca territory, causing a population catastrophe in this Empire as well, as reported by a number of studies (Lockhart 1968, Denevan 1976, Cook 1981, Cook and Lovell 1991).[5] McCaa et al. (2004) argue, however, that smallpox epidemics and disease factors in general were probably less important, whereas the effects of civil war and of exploitation were rather more decisive for the demographic catastrophe of the Andean region. When the Spanish troops arrived in Ecuador with the intention of conquering the Inca Empire, Atahualpa had just defeated his half-brother Huascar and his supporters in a bloody assault of Cuzco. In 1532 the Spanish soldiers led by Pizarro captured Atahualpa at the Battle of Cajamarca. This was the first step of a long fight to subdue the Inca Empire. The Inca ruler was held as a prisoner for eight months. During this time Pizarro received a ransom of "enough gold to fill a room 22 feet long by 17 feet wide to a height of over 8 feet" (Diamond 1997, p. 68) in exchange for the promise to free Atahualpa (which was never fulfilled). The Inca was sentenced to execution in a mock trial because the Spanish suspected he was plotting his rescue by a

[3] Their own creation myths link Inca people's point of origin to the Lake Titicaca, 300 kilometers south-east of Cuzco.

[4] The territory of the Inca Empire was then called *Tahuantinsuyu* and included *Chinchay Suyo* (the north), *Anti Suyo* (the Amazon jungle in the east), *Colla Suyo* (the south) and *Conti Suyo* (the west). The word *Tahuantinsuyu* derives from the Quechua *tawa* (meaning "four"), to which the suffix *–ntin* ("together" or "united") is added, followed by *suyu* ("region" or "province"), which roughly renders as "The four regions together".

[5] Following Crosby (1976, p. 207), the smallpox epidemic reached Inca domain in 1525 or 1526.

large troop under the Inca general Rumiñahui. The charges against Atahualpa were polygamy, incestuous marriage, and idolatry – all common rites in the Inca culture – and having killed his brother Huascar. Following Atahualpa's death, further Spanish troops arrived in their mission to conquer Ecuador and Peru. Benalcázar's soldiers defeated the great Inca warrior Rumiñahui in Ecuador and occupied the city of Quito with the help of Cañari tribesmen who served as allies against the Incas (Prescott 1847). The Indio resistance continued for the next forty years with frequent attacks against Pizarro, who was based in Cuzco, until the last Inca Tupac Amaru was murdered in Vilcabamba in 1572.

One particularly important aspect of pre-contact history is the inequality of education. The Inca culture had a comparatively advanced social system for the 16th century. The Inca Indios were great architects and developed an admirable sophistication in advanced agriculture (Klein 2011). But while some advanced cultures in the world were characterized by a relatively broad participation of middle and sometimes even lower classes, other advanced cultures were mainly based on a thin upper class and a large quantity of uneducated people. We will argue in the following that the Inca culture was more of the latter type, with an extremely high educational inequality. The 16th century chronicler of Inca origin, Garcilaso de la Vega, described the elitist attitudes of the 15th century Inca ruler Tupac Inca Yupanqi in the following way: *"Science was not intended for the people; but for those of generous blood. Persons of low degree are only puffed up by it, and rendered vain and arrogant. Neither should such meddle with the affairs of government; for this would bring high offices into disrepute, and cause detriment to the state"*.[6] This view reflects the ideology of the Inca monarchy, which would take care of their people, but not educate them or let them participate in any decision-taking process.

[6] Note: translated by Prescott (1847, p. 39) from Garcilaso de la Vega, Inca (1609, part 1, book 8, chapter 8).

The majority of the population received a very modest education under the Inca Empire, while the small ruling Inca cast and those who dominated the military and the religious sectors would receive training by the *amautas* (wise people) in order to be prepared for their future positions in the administration of the government or as priests. They were trained in different disciplines according to their future profession, which was most often hereditary, like military competence, religious rites or basic mathematics (Julien 1998). Though it was the common people who were in charge of moving earth and stones in order to build the irrigation systems, the massive stone buildings, fortress temples and the rest of the impressive architectural and engineering marvels for which the Inca Empire is known, the state would not provide education to them.[7]

Concerning the level of economic development reached in the Andes before the arrival of the Spanish, Comin et al. (2010) have noted that a number of economically useful and important inventions were missing in the Inca Empire, making it technologically inferior to Eurasia and even to the Aztec Empire. They construct a dataset on technological adoption – determined by the utilization of a set of basic technologies (communications, agriculture, military, industry and transportation) –, which allows to show persistence of technological differences between the predecessors to today's nation states over long periods. In their study on the persistence of technological differences between nations over long periods, they find that their pre-colonial technology measure of 1500 AD is a statistically significant predictor of per capita income and technology adoption at the present.

Not only education, but also income was distributed very unequally during the Inca Empire. Engerman and Sokoloff (2000) argue that in pre-conquest Peru (and similarly in

[7] People in charge of building palaces and other buildings in Cuzco were part of the coerced rotating labor system called *mita* (Julien 1998, p. 85).

Mexico) extreme inequality was already latent before the Spanish conquerors arrived. These regions were geographically blessed with a dense population and abundant mineral resources. The strictly hierarchical organization of the native American societies allowed the powerful elite to take profit of these factors. Europeans to some degree adopted the social organization in which small Indio elites extracted tribute from the general population, contributing to the path-dependency of inequality. Also the institution of coerced labor called the *mita*[8], on which mining relied during colonial times, already existed in a similar version for the native population ruled by the Incas. It was reintroduced in the 1570s by the viceroy Francisco de Toledo among other territorial and tax reforms in order to allow a more profitable exploitation of the Potosí silver mines. It required Indians to work by force in return for wages well below the existing market rate (Bakewell 1984). About 16 percent of the male native population from communities lying between Cuzco and Potosí were sent to mine and refine silver (or mercury in Huancavelica) on a rotating basis during one year (Bakewell 2004; p. 202). Despite Toledo's efforts to protect *mitayos* from the abuses of overwork and severe punishments, these evils combined with harsh natural conditions are supposed to have taken a heavy toll of the native conscripts. However, it would be difficult to know whether the *mita* reintroduced by the Spanish colonial regime was more or less extractive than the previously existing. The colonial version distinguished itself from the pre-Hispanic one in some key factors. The Spanish only employed this institution for mining and not for other sectors of the economy, as it was done during the Inca Empire (Julien 1998). Further, the colonial *mita* was a source of coerced, albeit salaried, labor and it coexisted with a free, comparatively expensive workforce. In the late 1500s the demand for labor rose because of the decimation of the

[8] *Mita* is derived from the Quechua word *mit'a*, meaning "turn" or "period of service" (Cole 1985, Dell 2010).

indigenous population due to disease, war, civil war, exploitation and the Spanish newcomers needing native workers for their *encomiendas* (Bakewell 2004, Tandeter 2002).[9] Thus, many *mitayos* preferred to remain in Potosí selling their skilled mining labor at the end of their stints, rather than turning back to the more deprived community life in the rural sector (Bakewell 1984). By the early seventeenth century, as suggested by Bakewell, more than half of the Potosí labor force of some 10,000 consisted of free workers. A very interesting fact is that, for all its imperfections, a market for labor appeared *ex novo* after the Spanish conquest and increased in importance over time in the Andes during the colonial period. This has implications for numeracy, because the return to this basic human capital investment increased after a market for labor was created.[10]

Does this imply that the well-being of the native population improved with the change from the Inca to the Spanish colonial regime? The arrival of the Spanish clearly increased the speed of technological progress (see, for example, Comin et al. 2010) and also offered new cultural opportunities, which were soon adopted by members of the Inca elite, such as written language. On the other hand, Klein (2011) suggests that the Inca Empire "functioned as a major distributer of goods and services in a nonmarket manner and probably created a well-being and wealth among all the population unmatched from those times to the present".[11] This very positive view of the late old American Indio empires does not harmonize with anthropometric and other health evidence presented by Steckel and Rose (2002). They found that South Americans' health deteriorated since the switch from hunter gathering to the agricultural system and the creation of the Inca Empire (and thereafter). For the post-conquest period, Dobado and García (2010) recently argued that Andean and Mesoamerican people were "not so short" during the colonial period,

[9] But see McCaa et al. (2004) on the limited role of disease until the 1560s.
[10] At this point we thank an anonymous referee for his comments.
[11] Klein (2011), p. 19.

although the evidence of well-being directly before and after Spanish conquest is not based on a large number of observations.

2.3 Sources and Methods of Basic Numeracy Measurement

The Incas of Peru made regular counts of the population in order to raise tribute, and to know the number of potential warriors and defenders of the community (Cieza de León 1609). They recorded the numbers on knotted-string mnemonic devices called *quipus*. Within a short time after the conquest, Spanish administrators also conducted a number of surveys of the native population in order to learn about the number of tribute-paying subjects being granted to colonists. The Spanish administrators of Peru generated a number of population lists, of which some have survived and contain valuable information for our study. The population list of Huanuco 1562 includes only indigenous people of this province, which was situated some 400 km northeast of Lima in a valley of the Andean mountains (Table 2.1). The population centers of Huanuco lie between 250 and 6,631 meters over the sea level. This information is particularly valuable for our study, because the administrators documented this district already in 1562, i.e., only three decades after the Spanish conquest of Peru. Hence, many surveyed persons were born when the Inca Empire was still independent. They also confirmed directly if somebody did not know his or her age. As most basic numeracy is acquired during the first decade of life, their age heaping behavior should reflect the educational investment during the Inca Empire. We also mention in Table 2.1 whether we would expect a bias relative to the overall population of the country. In the case of Huanuco, there might be a negative bias due to its remote and rural geography.

Our second source, the census of Indios in Lima 1613, was taken under the viceroyalty of Juan de Mendoza y Luna. The viceroy sent a copy of this Indio census to Spain, and this copy is the only one surviving.[12] Lima was newly founded by the Spanish in 1535. Hence, the indigenous population structure in 1613 is characterized by immigration of Indios from the countryside, who worked as craftsmen and servants in the capital. The male population was about twice as large as the female population. Indios represented only eight percent of the total population of Lima at the time. The age group of 23-32 was by far the largest age group.[13] The officials asked every person for his or her name, age, occupation, and birthplace. Most people came from today's Peru and Ecuador, and some were immigrants from other Latin American regions. In total this sample consists of 749 individuals aged 23 to 72.

The Inquisition of Lima provides another small source for people born in Peru and Spain (N=64). A part of it was published by Ricardo Palma in 1897, containing the documents of the Sacred Court in Lima from 1625 to 1761.[14] 42 percent of the Inquisition victims were female, who were mostly judged for heresy. Men were mostly accused of bigamy, trigamy and similar reasons.[15]

A fourth source that provides a list of all male household heads in Lima from 1700, also recorded a (small) number of Indios. The source contains 2,946 household heads in total and lists 76 persons of (at least partly) Indio origin, as well as a small number of mulatto and black persons. Hence, this source actually informs us better about the

[12] Unfortunately, the population records of whites, blacks, mulattos and mestizos have been lost.
[13] This structure could also be observed at the time in mining towns as Potosí or Huancavelica. Furthermore, the great majority of the households have two children or less. More than fifty percent had no children. The age groups running from end Figure 3 to 2 are standard in the age heaping literature.
[14] Other original documents are not available anymore, because they were burnt during an occupation of Lima by Chilean soldiers in 1881. The first edition of Palma's "Anales" was published before the fire when he could still have access to those documents.
[15] The Tribunal of the Santo Oficio was not occupied with the cases of Indios in Peru. There were during the colonial times other institutions that accounted for this purpose (Moreno de los Arcos 1990).

numeracy of the white household heads of Lima. Unfortunately, there is no information given about the place of birth of the inhabitants in this census. We assume that the majority was either born in, or close to the capital, or moved quite early to the capital. Thus, we assigned in the following regression the value 1 to the control variable 'large capital'.

These four sources can inform us about the situation in the Andean region, and particularly about the Indio numeracy. For better comparison, we obtain four sources on the European side, as we want to place our argument into a comparative perspective. As for Peru, we work with population censuses and also with Inquisition sources, although the latter might be upwardly biased, as we will discuss below. For Portugal and Spain, we have evidence on the four Inquisition places in Évora (southern Portugal), Logroño, Cuenca and Llerena (north, east central and western Spain, respectively).[16] The Llerena data set (N=259) consists of a large proportion of Portuguese immigrants, whereas some Spanish immigrants can be located in the Évora data set (N=2,368). We have one source stemming from other sources than Inquisition, namely a tax list of 1717 from the city of Badalona in Catalonia in north-eastern Spain (N=359). It is one of the early population registers that survived until today. For the 16th and 17th century in contrast, the Inquisition files are relatively abundant, but it is clear that we will need to control for possible biases below.

In Table 2.2, we provide some comparative characteristics of the data sets. The share of Jewish accusations is important for the Inquisition-based data sets, and is as high as 88% for the Portuguese data set from Évora. For the census-based data sets, we assume that the Jewish share was close to zero, because the Jewish were expulsed and persecuted

---

[16] The Cuenca data set (N=475) was kindly provided to us by Sara Nalle, for which we cannot thank enough, and for the Portuguese data entry we thank Rosemarie Triebe. We have no information on place of birth and accusation; hence we will insert dummy variables in the regressions below which will control for "partly Jewish", 'partly migrant' and "partly large city".

in the Spanish and Portuguese Empires under study here.[17] The female share was particularly high in the Logroño data set, because this Inquisition case was mainly directed against witches. Some of the sources contained only males (Badalona 1717, Lima 1700), whereas in most data sets the share was between one fifth and one half. The Huanuco data set of 1562 had also a quite high female share, which might be caused by the fact that males had died in larger numbers from the infectious diseases or suffered from the violence during conquest.

We also define the variables 'large city' and 'migrant' which assign these characteristics to the sample for which the information was given and relevant. Sources from Lima were naturally quite urban, whereas migrants were numerous in the Llerena and Lima Inquisition files. Finally, the "age 23-32" variable controls for the age structure in our samples. This is particularly important because younger individuals in their twenties round sometimes on multiples of two rather than five, and as the heaping index employed here does not capture this rounding, we need to control for this as well. The share of young adults was normally in the range of 25-35 percent, but in Lima 1604 it was substantially higher.

How is the heaping index calculated? The ratio between the preferred ages and the others can be measured by several indices, one of them being the Whipple index.[18] To calculate the Whipple index of age heaping, the number of persons reporting a rounded age ending with 0 or 5 is divided by the total number of people, and this is subsequently multiplied by 500. Thus, the index measures the proportion of people who state an age

[17] Following the Alhambra Edict, an Edict of Expulsion against the Jews issued by Ferdinand and Isabella in 1492, about half of the Jewish population living in Spain emigrated and the other half stayed and converted to Christianity. Thus, there were practically no more 'official' Jews in Spain and, since 1497, neither in Portugal.

[18] A'Hearn, Baten and Crayen (2009) found that this index is the only one that fulfils the desired properties of scale independence (a linear response to the degree of heaping), and that it ranks samples with different degrees of heaping reliably.

ending in a five or zero, assuming that each terminal digit should appear with the same frequency in the 'true' age distribution.[19]

$$(1)\ Wh = \left(\frac{(Age25 + Age30 + Age35 + ... + Age60)}{1/5 * (Age23 + Age24 + Age25 + ... + Age62)} \right) \times 100$$

For an easier interpretation, A'Hearn, Baten, and Crayen (2009) suggested another index, which we call the ABCC index.[20] It is a simple linear transformation of the Whipple index and yields an estimate of the share of individuals who correctly report their age:

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

The share of persons able to report an exact age turns out to be highly correlated with other measures of human capital, like literacy and schooling, both across countries, individuals, and over time (Mokyr 1983, A'Hearn et al. 2009, Crayen and Baten 2010).

2.4 Regression Estimates of Numeracy Trends

The aim of the following figures and regression tables is to estimate the numeracy trends for the Indio and white population of the Andean region under study, as well as those of Spain and Portugal during the 15th to early 18th century. However, we identified above a number of potential biases that could create positive or negative selectivity and hence need to be controlled for. The variables which should be taken into account in order to avoid biases are the following:

1. *Inquisition sources*. One could imagine that victims of the Inquisition were not a random sample of the underlying population. We assume that individuals included in

[19] A value of 500 means an age distribution with ages ending only on multiples of five, whereas 100 indicates no heaping patterns on multiples of five, that is exactly 20 percent of the population reported an age ending in a multiple of five.

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

Inquisition samples are positively selected, especially because the accused were to a great part "major heretics", including Judaists, Protestants and other spiritual devotees. Arguments for our assumption of Jewish and converts' superiority in terms of education are given in the next paragraph. Concerning Protestants, many studies have shown a better performance in education, especially in literacy, compared to Catholics (Becker and Woessmann 2010). Illuminists, "Erasmians", Lutherans and other Protestants, were usually individuals who had thought critically about theological (and political) issues and adopted innovative views on spiritual and intellectual life. They were often familiar with devotional literature censored by the Inquisition (Rawlings 2006, pp. 90-113). Moreover, it has been argued that the Inquisition profited from persecuting heretics by confiscating the goods of the accused (Rawlings 2006; p. 42). In fact, expropriation was a common conviction besides wearing the *sanbenito* (penitential cloth)[21], scourge, exile, jail and death. For this reason, one could imagine that victims of the Inquisition were relatively wealthy – and probably on average more educated –, so that the Sacred Tribunal could make larger profits from condemnations. Even if we control for the effects of birth in large cities and being of Jewish religion separately (see below), the remaining victims of the Inquisition might have been more educated, hence we include a dummy variable for age statements from this source, as opposed to census data and other population lists.

2. *Jewish religion.* Being accused of practicing Jewish religious activities could generate a bias for various reasons. Firstly, the Jewish community in the Middle Ages – and the converts' afterwards[22] – had acquired important positions in the Spanish society thanks to their excellent educational preparation. They kept posts in medicine, universities,

[21] The *sanbenito* or "sacred cloth" was a penitential tunic worn by the condemned during the public penance and sometimes for a certain time afterwards in public (Rawlings 2006; p. 159).

[22] Since after the edict of expulsion was issued in 1492, about half of the Jewish population converted to Christianity, the Inquisition sued against "New Christians" accused of practicing Jewish rites.

in the church and in public administration. Furthermore, they were active in the world of finance and the economy, working as bankers, merchants and tax collectors (Prado Moura 2003; p. 18). Secondly, accusations of practicing Jewish religious activities could focus on the relatively rich and educated, even if their connections to Jewish belief were remote or non-existent. Therefore, we expect that this group has a higher numeracy.[23]

3. *Female*. During most of history, female children received less education than males. Given that our samples feature different gender shares, we control for this and adjust the estimates below to represent the male numeracy level, because in some sources exclusively males were included.

4. *Large cities*. In large cities, it is typically less costly to organize schools than in the countryside or in small towns, because the pupils have to walk shorter distances (Bouccekine et al. 2007). Moreover, the decision makers of the central administration were living mostly in the capitals and large cities, and they preferred to finance schools which their own children could attend. In contrast, very rural regions like Huanuco in the 16th century were certainly educationally disadvantaged. Hence, we controlled for a possible urban or rural bias by including an indicator variable.[24]

5. *Migrants*. Individuals migrate to other countries if they expect a higher income and a higher standard of living in the new location. However, migration requires an investment, both in terms of money and psychological cost. In the early modern period, travel costs were very high, thus relatively rich and educated individuals might have been overrepresented.

[23] As the share of Jewish was quite low, our estimates below refer to the non-Jewish population (and they refer to non-inquisition sources).

[24] This variable is one for the following large cities which were included in our sample: Berlin, Bordeaux, Cadiz, Cartagena, Coimbra, Cuzco, Cordoba, Geneva, Granada, La Paz, Lima, Lisbon, Madrid, Malaga, Manila, Naples, Porto, Quito, Santiago de Chile, Seville and Toledo.

6. *Age 23-32*. We organized our age-heaping-based numeracy estimates by age group, because those in their twenties normally displayed a different heaping pattern: Similar to the older persons, they reported ages rounded on multiples of five, but some of their rounding was also on multiples of two, such as 24, 26, 28 etc.[25]

The results are reported in Table 2.3. Column 1 displays the marginal effects of a logistic regression in which the dependent variable is 1, if the individual reported an age that was not a multiple of five, and zero otherwise. We run a logistic regression because our dependent variable is a binary. The marginal effects were multiplied by 125, so that they can be interpreted as percentage changes of numeracy, taking into account that 20% of ages would correctly end on 0 or 5.[26] For example, age statements that come from the five Inquisition sources are characterized by an additional numeracy of 13.4 points, even when controlling for urban bias and religion separately.

We further find that those who were persecuted for practicing Jewish religious elements were 7.7 percent more numerate.[27] The occupational constraint of the Jewish to some money-related activities and the religious demand to acquire reading abilities might have caused this advantage. In contrast, females were less educated than males during this early period, by as much as 16.9 percent, and inhabitants of large cities displayed higher education (by 12.8 percent). All these coefficients are statistically significant at the 5 percent level, and we will see below that they are also economically significant, i.e., that their size is relevant relative to overall differences of numeracy between countries. The coefficient of migrants was positive as expected, but not statistically significant. The

[25] This part of the age heaping is not reflected in the ABCC index, which focuses on heaping on multiples of five. Hence, we need to control for these young age groups and expect to see a seemingly positive bias of this age group, relative to the age groups 33-72.

[26] We included a note in Appendix C explaining the derivation of the multiplication with 125 for estimating numeracy.

[27] This effect is even higher if we do not control for Inquisition sources, as in Table 3 Model 3.

adjustment of those aged 23-32 in contrast, is highly significant and should be definitely taken into account.

We also estimated two different specifications in order to assess the robustness of our results. In column 2, we removed the Inquisition of Cuenca, for which we have no information about the accusation[28] and place of birth. Most of the coefficients are relatively stable, except for Spain in the late 15th century, which had rested mostly on this data set (the remaining few observations from other sources are too few for a reasonable estimate). As a second robustness test, we omitted the source-specification variable "Inquisition", which again resulted in fairly robust coefficients.

In order to estimate numeracy trends, we created a large set of coefficients that refer to each country, ethnic group and birth period. We have chosen half centuries as birth periods, in order to estimate the long-run trends and assess the difference between the Inca Indios before and after the Spanish conquest, as well as to compare their human capital with that of the European samples.

In order to calculate adjusted ABCC levels for all these groups, we run a regression of "not reporting a multiple of five" as the dependent variable. The independent variables are those country-ethnicity-period indicator variables and the full set of control variables (Table 2.3, column 4). We use a linear probability model due to the fact that this tool provides a constant, which we need for the calculation of ABCCs.[29]

As a result, we obtain here the first human capital estimates for Spain and Portugal for the late 15th century and thereafter. In Figures 2.1 and 2.2, we can compare the raw values and the adjusted estimates, which were calculated with the coefficients of Table 2.3,

---

[28] The explanatory variable "Partly Jew" is only relevant for the Cuenca Inquisition source because of this. Thus this variable is not included in Model 2.

[29] The correlation between coefficients of the logistic and the linear probability model is 0.96 and the slope in a scattergram comparing the two is equal to one (Figure available from the authors), hence there is almost no difference.

column 4. The adjusted Spanish and Portuguese numeracy values were between 20 and 30 percent in the late 15th century, and started a slow, but steady increase from this level until the early 17th century, when they reached between 50 and 70 percent. The 17th century was characterized by a stagnation or slight decline in Spain, and a temporary increase between the early and late 17th century in Portugal, but this was lost again in the early 18th century.

We estimate numeracy for Peruvian Inca Indios for the early 16th century, i.e., mainly before the Spanish conquerors arrived. The Peruvian Inca Indios had only about half the numeracy level of the Spanish invaders. This holds true even after adjusting for the downward bias, which might have been typical for the remote and isolated area of Huanuco, on which our early Peruvian data is based. The values during the late 16th century seem to have converged strongly towards Spanish and Portuguese levels. In fact, the numeracy estimate of Indios from Ecuador was even higher than Portuguese levels in the late 16th century. The high numeracy during this period might be slightly biased by the fact that we can only observe Inca Indios who had moved to the newly founded city of Lima and who might have been a quite particular selection of Indios. This view is supported by the fact that the numeracy of 17th century Indios fell back to substantially lower values. It could also be that Indios living in the capital exploited the cultural opportunities offered by the colonial society in form of numerical skills, the way members of the new intellectual elite – Garcilaso de la Vega and Guaman Poma de Ayala – did with the written language.

Nevertheless, it is interesting that the Ecuador Indio values were the highest, because the Indio collaboration with the Spanish conquerors was concentrated in the region that later became Ecuador.[30] This was the area of the Cañari Indios, some of the most

[30] The specification "Ecuador Indios 1550" refers to migrants coming from the area of today's Ecuador shown up in the Indio population list of Lima in 1613.

decided collaborators of the Spanish. The high levels of numeracy of the Cañari could have been a positive consequence of the Spanish rule for them, which might have implications for the colonial legacy. However, the levels of the late 16th century might be overestimated due to the immigration selectivity to Lima.

In the Lima census of 1613, those were explicitly identified as Cañaris, because belonging to this tribe was associated with a privileged position. In fact, it seems that this group of Indios was exempt from tribute and the *mita* as a reward for their help during the Indio's siege of Cuzco in 1536 (Cook 1981, p. 83; Livi Bacci 2008, p. 162). It is not astonishing that they could provide somewhat better education to their children than the Peruvian Indios, or the Southern Indios, which we indicated with the label "Chile/Bolivia/Argentina" (most of them were born in what is today Chile). [31]

The numeracy of white Peruvian household heads that we obtained from the 1700 census of Lima was situated in the middle between Peruvian Indios and Iberians in the 17th century. In contrast, the relatively few mulattos and blacks in Lima during this period reported only rounded ages. As a caveat, we should note that it might have been possible that census enumerators did not even bother asking them, as they did not expect a true age statement. The temporary increase of Indio numeracy during the late 16th century, even if it might have been partly caused by selectivity that we could not capture in our adjustment regressions above, also proves that Indios were not generally unable to reach numeracy levels which were comparable to Spanish or Portuguese levels during the 16th century. There was no cultural or perhaps even genetic hurdle which would have kept them from developing substantial age numeracy, if the educational level and other context parameters would have been sufficient.

[31] Individuals included in the variable "Chile/Bolivia/Argentina 1550" are also Indian migrants to the city of Lima captured in the Lima Indio count of 1613.

How high was Spanish and Inca Indio numeracy in comparison with other populations in Europe and Asia (Figure 2.3)? In the late 15th century, the Dutch, Italians and Germans had reached a higher numeracy than Iberians, and the gap between Northwest and Southwest Europe might have remained roughly constant thereafter, although numeracy grew in both regions. For the late 17th century, we have the first evidence for China, which had a very high numeracy. During the early 19th century crisis of China, its numeracy fell back below the European Northwest (see also Gupta and Ma 2010, pp. 274-275).[32] Summing up, we found that our sources based on Inquisition records were upward-biased, not only because the Inquisition victims were often of Jewish origin (or were so numerate that the accusers thought they might be Jewish) or urban, but also because Inquisition victims per se were more educated than the average population. Moreover, and crucial for the main question of this study, we find that among the adjusted figures, Spanish and Portuguese had a moderate numeracy during the late 15th and early 16th century, but it was at least twice as high as the numeracy of the Inca Indios. Hence, the low human capital level of Inca Indios indicates that a pre-colonial legacy burden is quite likely.

2.5 Potential Objections

We are aware that this study might be confronted with a number of possible objections.

Objection 1: Is the age heaping technique informative in the case of the Inca Indios? Or did they have culturally determined number preferences which were different from European ones, perhaps partly based on their knot-based counting method? Are the sources used reliable and valid for estimating numeracy of the aboriginal population in the late pre-Hispanic period? Our response to this objection is: It seems that all human beings

[32] Gupta and Ma (2010) report that literacy in the early 19th century was quite low.

start to count with the fingers of their hands, hence there is a strong preference to express numbers as multiples of five and ten. This is also observable when the Inca Indios were asked for their age, i.e., the data do not point to a preference for another digit. Some additional insights about number preferences can be gained from considering the Inca number system, which relied on the science of the quipus. Quipus, which translated means "knots", were fabricated by civil servants who were in charge of compiling and reporting information on the population structure or the agricultural inventories and the provisions delivered to the royal storehouse. They registered births, deaths, marriages and the number of tributaries as well as the number of people who could be recruited for war.[33] The distance between the knots represented the numbers and the colours of the cord were meant to identify the different counted units (Livi Bacci 2008; p. 161). Quipus were the only system of written numerals and, though little is known today on this system, it was used for a decimal order, with knots that represented singles, tens, hundreds, and higher numbers (Menninger 1934, p. 196; Espinoza Soriano 1997, p. 422). Hence, we would argue that it is consistent to assume that the Inca Indios would, as other cultures do, have a preference for multiples of five, and not for any other digit.

Another answer to this question is the fact that within the responses of the Indios, we see a clear stratification between richer and poorer – and presumably more and less educated – Indios, as indicated by their taxation records. This provides credibility to the indicator (see section 5 for a more detailed treatment). Finally, we observed special groups of Indios in the previous section – the Cañari Indios of Ecuador – who were as numerate as the Europeans.

[33] After the new census was planned, the official counters were sent by the Inca to the places in question. Quipu records of the previous count were brought and after noting deaths and births, new divisions of the population into age categories were recorded.

Thus we argue that the sources and the technique used here are valid for estimating numeracy of the aboriginal population in the late pre-Hispanic period and early colonial time and that the results can be compared directly with other samples, i.e. those of Indios born after the conquest and of Europeans.

Objection 2: What about selective mortality? Who survived the infectious diseases and hence could be asked for the age later-on? Our answer to this question is: we do not know exactly. The nature of most epidemics is that mortality is not very selective, rich and poor both die from it. It is slightly different in the case of mortality after bad harvests, when it is more selective. But even if the native mortality from the imported epidemic diseases were selective, our argument would be strengthened: as better educated Indios would have survived in this scenario, our numeracy estimate for the pre-Spanish period would even be over-estimated. However, the findings of McCaa et al. (2004) suggest that this might not be a major issue.

Objection 3: But is the Inca culture not famous because they had achieved such a high cultural level? Is this not a reflection of a high human capital? Yes, indeed, the Inca culture had many fascinating elements, including their architectural achievements – such as impressive stone buildings –, a broad infrastructure of paths connecting the whole Empire and a highly developed agricultural system. However, as discussed in section 1, the overall impression is that educational inequality was very high in the Inca Empire. A small elite group received training and education (as described in section 2), but the majority of the population did not.

Moreover, the overall scientific and technological development in the Andes was not very impressive, as suggested by Diamond (1997) and recently supported by Comin et al. (2010). Diamond stresses the fact that the Spanish conquerors – even though poorly educated themselves[34] – were coming from a culture with a certain share of literacy, whereas the Inca Indios had not developed a comparable literacy culture.[35] Measured in terms of technological adoption, Comin et al. (2010) find out that Incas not only lagged far behind most of Eurasia and Northern Africa, but were also outperformed by the Aztecs in terms of the utilization of a set of basic technologies (communications, agriculture, military, industry, and transportation).

Objection 4: Are the samples used here comparable? Is the validity of the sources given? The Huanuco sample of Indians born before Spanish conquest might cover the whole surviving population of this region. But Huanuco had no large urban settlement, which would have been characterized by typically better educational standards, probably also in the Inca culture. Hence we have worked with multiple regression models which allow to control for a large urban share in order to make the Huanuco sample comparable with the other samples. This is the reason why our numeracy estimate for Inca Indios born in the early 16th century (before the conquest) increased from 2.4 points to 12.5 points.[36] We applied the same procedure to the "Inquisition bias", because otherwise the samples affected by it would have been unrepresentatively high.

We have to admit that for the Inca economy before the conquest, we need to work like archeologists who find artifacts and interpret them based on what they know about the context based on other evidence. But the validity and comparability of the samples is given, as argued above.

---

[34] The educational level of the conquerors was mixed, but Pizarro himself could not read and write (Bakewell 2004; p. 100).

[35] In this context Diamond points out that Peruvians had never heard (or read) of the existence of the Spanish until Pizarro's first landing in Ecuador. They also remained ignorant about Spain's conquest of Central America. Diamond (1997, p. 80) emphasizes that "literacy made the Spaniards heirs to a huge body of knowledge about human behavior and history".

[36] Further, a gender bias characterizes some of the sources, such as the Lima Indio count. This is due to the fact that the Indio population of the new Peruvian capital comprised a large share of male working migrants (see Escobar Gamboa 1968; p. VII).

What would make a source unsuitable for the age heaping methodology?[37] Not asking individuals directly for their age, as well as counterchecking, would be an insurmountable obstacle. We have taken care that the persons listed were indeed asked for their age and that it was not guessed by the counting officials. It is also important that the age statements were not counterchecked with birth registers by the officials. In the introduction of the Lima and Huanuco sources, the explanations given to the counting officials are explicitly to ask them personally for their age (Escobar Gamboa 1968, Murra 1967, Cook 1985). Sources which were counterchecked, typically do not display age heaping.

2.6 Social and Regional Differences of numeracy

We will observe in the following whether numeracy differences existed among the social strata of Inca Indios. Our aim is to assess whether the age heaping technique can be applied as a proxy for human capital of the Indio population within their specific culture. If different Indio status groups show different age heaping levels, then this would be a hint that the proxy is informative. In the following we assess the differences in education by occupational groups or, in the case of Huanuco, approximated by three social groups and by the tribute paid to the Spanish *encomendero*.

The Huanuco region lies in the Andean highlands, about 400 kilometers away from Lima. It is thus a very remote region and we assume that the population had had nearly no contact with Spaniards at the time the population count was made. Our Huanuco sample includes three social groups called *chupacos*, *mitimaes* and *yanaconas*. What defines these groups? The first two groups belong to the common rural community, whereas *yanaconas*

[37] For caveats discussed in the literature, see also Manzel et al. (2010) and Stolz et al. (2011). In our source, we find direct remarks indicating that even the poorest women were individually asked for their age.

were a caste of slaves in service of the privileged classes (Baudin 2003, p. 390). The rural community lived in small clan groups. A certain amount of land was assigned to each of these clan groups depending on the number of its members. They mainly worked for their subsistence and not for the market, but they gave a part of their agricultural output as a tribute to the royal storehouses.[38] A special group within the rural population was that of the *mitimaes*, who were forced to move from the imperial centre to other valleys in the Andes for political reasons, especially in order to populate newly conquered regions (Bakewell 2004, p. 15).

Yanaconas were born into slavery as descendants of an ethnic group that had raised great revolts against the colonizing Inca tribe centuries before the Spanish conquest (Espinoza Soriano 1997, p. 287). We could expect this group to be disadvantaged compared to the other social groups.

In order to assess whether we could find differences in numeracy between social groups in Huanuco, we divide our data set into the three categories, *chupacos*, *mitimaes* and *yanaconas* and run a regression with *chupacos* as reference category. Although the slaves had the expected negative sign, our results (Model 1 in Table 2.4) show no significant effect of the two independent variables representing social groups. This might be explained by the fact that these three groups could have suffered from the same low quality of education.[39] As described above (section 1), the main divide was between the small elite and the rest of the Inca Indio population.

---

[38] A second sort of tribute was paid in form of labour, within the structure of the so-called *mita*. The *mitayos* generated income for the state by the cultivation of its land, keeping the cattle, exploiting its mines or fabricating arms and other artisanal objects and buildings or infrastructure. It consisted of a labour service by millions of men between 18 and 50 years old working exclusively for the construction of the state or in service of the elite.

[39] *Yanaconas* lived between the common rural population and could even aspire to social mobility, to a greater extent than the rest (Espinoza Soriano 1997, p. 287). In our sample we can find a few *principales*, and

At this enumeration, the visitors took into account the amount and form of the tribute the Indio population paid, which is useful for our purposes. Of the 747 people counted between 23 and 62 years of age, 638 paid some type of tribute to their *encomendero.*[40] The five different agricultural products in which the tribute was paid were wheat, maize, cotton balls, cotton pieces and poultry. It was not possible to quantify the value of the tributes, but we can assume that the wealthier Indios paid in more diverse items. In order to assess whether the relative wealth could have been a correlate of numeracy, we construct five different tribute categories depending on the amount of items delivered as a tribute, and run a logistic regression, observing marginal fixed effects. In column 2 of Table 2.4 we build dummy variables for the number of tribute categories: five categories, four categories, and so forth ending with 'no tribute'. Moreover, we control for other variables discussed in section three, such as gender and age structure. Our regression outputs show a positive significant correlation of the four and five tribute category with the numeracy level. There is also a positive significant correlation with paying any various tributes (Column 3).

In a second step we assess whether numeracy differences existed between Indios belonging to distinct social classes. Therefore we divide our complete data set into occupational groups according to the scheme Armstrong suggests for classifying enumerated populations (Armstrong 1987, p. 214).[41] We first run a logistic regression for the full data set of both Indios and Europeans taking the lowest skill group 0 as a reference

they belong to the three of the population groups. Unfortunately, the number of cases of *principales* in our sample is too small to be included in the analysis of differences in numeracy.

[40] After the Spanish conquest, the people continued paying tributes, but instead of to the Inca state, they paid it to their Spanish landlord, or *encomendero*. *Principales* (Indio village chiefs) were still in charge of collecting it.

[41] Farmers form an extra professional group, because they can belong to a wide range of groups along the social hierarchy depending on if they own the land or not.

category, which refers to the ones with no profession (some of them disabled). Our results in Model 1 of Table 2.5 show significant positive effects on numeracy for all higher skilled groups, relative to the constant, which refers to having no occupation. In our second model we consider only the Indios included in our data set. Here the reference category is the occupation group of "semi-professionals" and "professionals" together. We opted for pooling both groups, because only four Indios in our data set belong to the professional group, which includes doctors, interpreters, village chiefs, etc. We obtain in Model 2 significant negative effects of the variables "semi-skilled" and "skilled" relative to the most educated group. The "unskilled" also have a negative coefficient. However, it is not significant. This confirms that Indios who are assumed to have benefited from a better education would tend to round their age on multiples of 5 less often than poorly educated Indios, also those born before contact with the Europeans. We therefore conclude that the age heaping methodology is a good proxy for the numeracy of Peruvian Indios that also allows us to establish comparisons with European counterparts.

2.7 Conclusion

One of the most crucial moments in history was the meeting of Indios of the ancient American cultures and the European invaders of the 15th and 16th centuries. We launched the hypothesis that low educational investments in the pre-colonial Andean region could have influenced later long-term development patterns. We based our argument on the literature about the "colonial legacy" (Acemoglu et al., Engerman and Sokoloff, Coatsworth, Bulmer-Thomas) and argued that also pre-colonial conditions hampered economic growth. Especially, low human capital investment in the Inca Empire might have initiated a path-dependent process of agriculture which was not human-capital intensive in the following centuries.

The human capital of the native Inca society was measured with age heaping methods. The sources employed allow us to gain insight into pre-colonial numeracy levels for the first time and to compare them with European values. We discussed adjustments that were necessary for some types of sources, and some characteristics of the sample, performing a joint logistic regression with these adjustment variables, and a large set of country-half century dummy variables. The results of the adjustment variables were interesting by themselves. For example, age statements which come from the five Inquisition sources were characterized by an additional numeracy of 13 percent, even when controlling for urban bias and religion separately. We estimated numeracy for Peruvian Inca Indios for the early 16th century, i.e. mainly before the Spanish conquerors arrived. Crucial for this study is the fact that the Peruvian Inca Indios had only about half the numeracy level of the Spanish invaders. This held true even after adjusting for the downward bias which might have been typical for the remote and isolated area of Huanuco on which our early Peruvian data is based. As for the evolution of human capital in the colonial period, our adjusted figures show that Spanish and Portuguese numeracy remained twice as high as in the Latin American regions of the Inca Empire, though Northwest-Europe had even higher numeracy. Finally, our numeracy estimates for the indigenous population remain nearly ten points lower than the White Peruvians'.

Our sources and methodology further allowed us to find evidence on inequality in pre-Columbian times. Given the low educational level and the high inequality reigning before the Spanish conquest in the Andean region, we argue that more attention should be paid to the pre-colonial legacies when assessing the genesis of the long-term path of only modest economic growth in the countries of Latin America.

2.8 References

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

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics*, 117 (4): 1231–1294.

A'Hearn, B., Baten, J., and Crayen, D. (2009). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. *Journal of Economic History*, 68(3): 783–808.

Armstrong, W. A. (1972). The Use of Information about Occupations: Part 2. An Industrial Classification 1841-1891. In: *Nineteenth-Century Society,* edited by E. A. Wrigley. Cambridge University Press, 226-310.

Bakewell, P. (1984). *Miners of the Red Mountain: Indian Labour in Potosi*. Albuquerque: The University of New Mexico Press.

Bakewell, P. (2004). *A History of Latin America: c.1450 to the Present*. Malden: Blackwell Publishers.

Baten, J., and Mumme, C. (2010). Globalization and Educational Inequality During the 18th to 20th Centuries: Latin America in Global Comparison. *Revista de Historia Económica - Journal of Iberian and Latin American Economic History*, 28(2): 279–305.

Baudin, L. (2003). *Daily Life of the Incas*. Mineola, New York: Dover Publications.

Becker, S. O., Cinnirella, F., and Woessmann, L. (2010). The Trade-Off Between Fertility and Education: Evidence from Before the Demographic Transition. *Journal of Economic Growth*, 15(3): 177-204.

Bulmer-Thomas, V. (2006). *The Cambridge Economic History of Latin America*, 2 Vols. Cambridge: Cambridge University Press.

Boucekkine, R., de la Croix, D., and Peeters, D. (2007). Early Literacy Achievements, Population Density, and the Transition to Modern Growth. *Journal of the European Economic Association,* 5(1): 183–226.

Cieza de León, P. (1609). *The Discovery and Conquest of Peru: Chronicles of the New World encounter*. Edited and translated by Alexandra Parma Cook and Noble David Cook. Durham, N.C: Duke University Press., 1998.

Coatsworth, J. H. (2008). Inequality, Institutions and Economic Growth in Latin America. *Journal of Latin American Studies*, 40(3): 545–569.

Coatsworth, J. H. and Taylor, A. M. (1998). *Latin America and the World Economy Since 1800*. Cambridge: Harvard University Press.

Cole, A. (1985). *The Potosí Mita, 1573–1700: Compulsory Indian Labor in the Andes*. California: Stanford University Press.

Comin, D., Easterly, W., and Gong, E. (2010). Was Wealth of Nations Determined in 1000 BC?. *American Economic Journal: Macroeconomics*, 2(3): 65–97.

Cook, N. D. (1981). *Demographic Collapse, Indian Peru, 1520–1620*. New York: Cambridge University Press.

Cook, N. D. (introduction and edition) (1985). *Numeración General de todas las Personas de Ambos Sexos, Edades y Calidades que se ha hecho en esta Ciudad de Lima Año de 1700*. Lima: Ediciones COFIDE.

Cook, N. D., and Lovell, W.G. (1991). *Secret Judgment of God: Old World Disease in colonial Spanish America*. Norman, London: University of Oklahoma Press.

Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History*, 47(1): 82–99.

Crosby, A. W. (1972). *The Columbian Exchange: Biological and Cultural Consequences of 1492*. Westport, CT: Greenwood publishing.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, 78(6): 1863–1903.

Denevan, W. M. (1976). *The Native Population of the Americas in 1492.* Madison: The University of Wisconsin Press.

Diamond, J. (1997). *Guns, Germs and Steel: The Fates of Human Societies.* New York: W.W. Norton & Co.

Dobado González, R., and Garcia Montero, H. (2010). Colonial Origins of Inequality in Hispanic America? Some Reflections based on New Empirical Evidence. *Revista de Historia Económica - Journal of Iberian and Latin American Economic History*, 28(2): 253–277.

Easterly, W., and Levine, R. (2003). Tropics, Germs, and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics*, 50(1): 3–39.

Engerman, S. L., Mariscal, E.V., and Sokoloff, K.L. (1999). The Persistence of Inequality in the Americas: Schooling and Suffrage, 1800–1945. Working Paper, University of California, Los Angeles.

Engerman, S. L., and Sokoloff, K. L. (2000). History Lessons: Institutions, Factor Endowments, and Paths of Development in the NewWorld. *The Journal of Economic Perspectives*, 14(3): 217-232.

Escobar Gamboa, M. (1968*). Padrón de los Indios de Lima en 1613. Seminario de Historia Rural Andina*. Lima: Universidad Nacional Mayor de San Marcos.

Espinoza Soriano, W. (1997). *Los Incas, Economía, Sociedad y Estado en la Era del Tahuantinsuyo.* Lima: Amaru Editores.

Esteban, A. (1997). *Literatura Hispanoamericana: Introducción y Antología de textos*. Granada: Comares.

Garcilaso de la Vega, Inca (1609). *Comentarios Reales de los Incas*. Edited by Aurelio Miró Quesada. Caracas: Biblioteca Ayacucho, 2 Vols., 1976.

Gupta, B., and Ma, D. (2010). *Europe in an Asian Mirror: the Great Divergence*. In: *The Cambridge Economic History of Modern Europe*, Vol. 1, edited by S. Broadberry,. and K. H. O'Rourke. Cambridge, UK: Cambridge University Press, 264-285.

Henningsen, G. (1980). *The Witches' Advocate: Basque Witchcraft and the Spanish Inquisition (1609–1614)*. Reno, Nevada: University of Nevada Press.

Julien, C. (1998). *Die Inka: Geschichte, Kultur und Religion.* München: Verlag C. H. Beck.

Klein, H. S. (2011). *A Concise History of Bolivia*. Cambridge: Cambridge University Press.

Lea, H. C. (1998). The Inquisition in 17th –Century Peru: Cases of Portuguese Judaizers. [http://www.fordham.edu/halsall/mod/17c-lea-limainquis.html]

Lisón Tolosana, C. (1979). *Brujería, Estructura social y Simbolismo en Galicia*. Madrid: Akal Editor.

Livi Bacci, M. (2008). *Conquest: The Destruction of the American Indios*. Cambridge, US: Polity Press.

Lockhart, J. (1968). *Spanish Peru: 1532–1569: A Colonial Society*. Madison, Milwaukee, London: The University of Wisconsin Press.

Manzel, K., Baten, J., and Stolz, Y. (2012). Convergence and Divergence of Numeracy: the Development of Age Heaping in Latin America from the Seventeenth to the Twentieth Century. *The Economic History Review*, 65(3): 932-960.

Menninger, K. (1934). *Zahlwort und Ziffer: Aus der Kulturgeschichte unserer Zahlsprache, unserer Zahlschrift und des Rechenbretts*. Breslau: Ferdinand Hirt Verlag.

McCaa. R., Nimlos, A., and Hampe-Martínez, T. (2004). Why Blame Smallpox? The Death of the Inca Huayna Capac and the Demographic Destruction of

Tawantinsuyu (Ancient Peru). Working Paper presented at the American Historical Association Annual Meeting, Washington, D.C.

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

Moreno de los Arcos, R. (1990). *La Inquisición para Indios en la Nueva España (Siglos XVI a XIX)*. Servicio de Publicaciones de la Universidad de Navarra, 2(1): 1471–1484.

Murra, J. V. (ed.) (1967). *Visita de la Provincia de Leon de Huánuco en 1562*. Tomo I. Iñigo Ortiz de Zúñiga. Huanuco: Universidad Nacional Hermilio Valdizan.

Nalle, S. T. (1989). Literacy and Culture in Early Modern Castile. *Past and Present*, 125(1): 65–96.

Palma, R. (1897). *Anales de la Inquisición de Lima*. Madrid: Ricardo Fé.

Perez Villanueva, J., and Escandell Bonet, B. (2000). *Historia de la Inquisición en España y América*, Vol. 3. Madrid: Biblioteca de autores cristianos, centro de estudios inquisitoriales.

Poma de Ayala, G. F. (1615). *Nueva Corónica y Buen Gobierno*. Edition and preface by Franklin Pease. Caracas: Fundación Biblioteca Ayacucho, 1980.

Porras Berrenchea, R. (1986). *Los Cronistas del Perú (1528–1650) y otros ensayos.* Lima: Ediciones del Centenario, Banco de Crédito del Perú.

Prado Moura, A. (2003). *El Tribunal de la Inquisición en España (1478–1834).* Editorial ACTAS, Madrid.

Prados de la Escosura, L. (2000). International Comparisons of Real Product, 1820–1990: An Alternative Data Set. *Explorations in Economic History,* 37(1)**:** 1–41.

Prescott, W. H. (1847). *History of the Conquest of Peru*. London: R. Bentley.

Rawlings, H. (2006). *The Spanish Inquisition*. Oxford: Blackwell Publishing.

Shea, D. E. (1976). A Defense of Small Population Estimates for the Central Andes in 1520. In: The Native Population of the Americas in 1492, edited by W. M. Denevan. Madison: The University of Wisconsin Press.

Steckel R. H., and Rose J. C. (2002). *The Backbone of History*. Cambridge: Cambridge University Press.

Stolz, Y., Baten, J. and Botelho, T. (2011). Growth Effects of 19th Century Mass Migrations: "Fome Zero" for Brazil. University of Tübingen Working Papers in Economics and Finance.

Tandeter, E. (2002). *Coacción y Mercado: la Minería de la Plata en el Potosí Colonial, 1692-1826*. Madrid: Siglo XXI de España Editores.

Wachtel, N. (1977). *The Vision of the Vanquished: The Spanish Conquest of Peru through Indian Eyes, 1530–1570*. Sussex: The Harvester Press.

Wedin, A. (1965). *El Sistema Decimal en el Imperio Incaico*. Madrid: Insula.

Williamson, J. (2009). Five Centuries of Latin American Inequality. NBER Working Paper No.w15305.

2.9 Tables

Table 2.1: Sources, ABCC Indices and Possible Biases by Sample

Label	Source	ABCC	Selectivity hypothesis?
Huanuco 1562	Murra (1967)	2	Negative (remote region, no contact to Spanish)
Lima 1700	Cook (1985)	45	Positive (large city)
Lima Indios 1613	Escobar Gamboa (1968)	36	Positive (large city, migrants)
Lima Inquisition (1625 – 1761)	Palma (1897)	44	Positive (Inquisition, large city)
Badalona 1717	National Archive Catalonia	66	Positive (large city)
Cuenca Inquisition (1521 – 1723)	National Historical Archive, Madrid (Nalle, 1989)	60	Positive (Inquisition)
Évora Inquisition (1487 – 1894)	National Archive Torre do Tombo, Lisbon	53	Positive (Inquisition)
Llerena Inquisition (1695 – 1730)	Pérez Villanueva and Escandell Bonet (2000)	81	Positive (Inquisition)
Logrono Inquisition (1609 – 1633)	Henningsen (1980)	49	Positive (Inquisition)

Table 2.2: Characteristics by Sample

Source	Cases	% Jew. accus.	Female	Large city	Migrant	Age 23-32	Source Inquis.
Huanuco	746	0.00	0.63	0.00	0.00	0.46	0
Lima 1700	2946	0.00	0.00	1.00	0.00	0.38	0
Lima Indios 1613	749	0.00	0.33	0.11	0.10	0.64	0
Lima Inquisition	64	0.05	0.42	0.41	0.23	0.38	1
Badalona	359	0.00	0.00	0.00	0.00	0.30	0
Cuenca Inquisition	475	0.00	0.19	0.00	0.00	0.32	1
Évora Inquisition	2368	0.88	0.52	0.03	0.02	0.36	1
Llerena Inquisition	259	0.48	0.44	0.16	0.33	0.29	1
Logroño Inquisition	157	0.00	0.80	0.00	0.00	0.22	1

Table 2.3: Determinants of Numeracy (as percentages)

	(1)	(2)	(3)	(4)
Dependent variable	Numerate	Numerate	Numerate	Numerate
Estimation technique	Logit mfx	Logit mfx	Logit mfx	Linear prob. Model
Data included	All	Except Cuenca	All	All
Source: Inquisition	13.38**	13.88**		11.78**
	(0.02)	(0.02)		(0.03)
Accusation: Judaism	7.71**	8.04**	10.05***	7.04**
	(0.03)	(0.03)	(0.00)	(0.03)
Female	-16.88***	-16.88***	-16.13***	-15.13***
	(0.00)	(0.00)	(0.00)	(0.00)
Large city	12.75***	12.46***	14.25***	12.04***
	(0.01)	(0.01)	(0.00)	(0.01)
Migrant	7.36	9.10	9.24	7.33
	(0.25)	(0.17)	(0.15)	(0.20)
Age 23-32	18.63***	18.25***	18.75***	17.13***
	(0.01)	(0.01)	(0.01)	(0.01)
Mulatto and Black 1650	-19.75	-20.13	-9.49	-18.25
	(0.28)	(0.26)	(0.66)	(0.24)
Chile, Bolivia and Argentina 1550	39.00***	37.50***	37.13***	19.75*
	(0.00)	(0.00)	(0.00)	(0.09)
Ecuador Indios 1550	47.63***	46.50***	45.88***	31.00***
	(0.00)	(0.00)	(0.00)	(0.01)
Spain 1450	35.63***	12.39	44.38***	16.00*
	(0.00)	(0.65)	(0.00)	(0.06)
Spain 1500	40.50***	37.13***	49.13***	21.63***
	(0.00)	(0.00)	(0.00)	(0.00)
Spain 1550	51.00***	51.00***	58.75***	34.25***
	(0.00)	(0.00)	(0.00)	(0.00)
Spain 1600	64.75***	64.38***	67.88***	58.75***
	(0.00)	(0.00)	(0.00)	(0.00)
Spain 1650	62.00***	62.25***	64.50***	50.13***
	(0.00)	(0.00)	(0.00)	(0.00)
Peru 1600	36.50***	35.75***	47.63***	17.63
	(0.00)	(0.01)	(0.00)	(0.13)
Peru 1650	44.50***	43.63***	55.75***	26.50**
	(0.00)	(0.00)	(0.00)	(0.02)
Peru 1700	41.00***	40.75***	51.13***	24.00
	(0.00)	(0.01)	(0.00)	(0.13)
Peru Indios 1550	42.88***	42.88***	42.88***	24.13***
	(0.00)	(0.00)	(0.00)	(0.00)
Peru Indios 1600	15.38	14.50	28.50	2.46
	(0.48)	(0.51)	(0.15)	(0.88)
Peru Indios 1650	30.00*	29.25*	41.63***	11.64
	(0.06)	(0.07)	(0.00)	(0.42)
Portugal 1450	29.00**	28.00**	40.00***	11.28
	(0.02)	(0.03)	(0.00)	(0.30)
Portugal 1500	31.38***	30.38***	42.13***	13.25**
	(0.00)	(0.00)	(0.00)	(0.04)
Portugal 1550	42.63***	41.75***	52.13***	24.50***
	(0.00)	(0.00)	(0.00)	(0.00)

Portugal 1600	49.00***	48.25***	57.63***	32.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Portugal 1650	58.50***	58.00***	64.63***	46.38***
	(0.00)	(0.00)	(0.00)	(0.00)
Portugal 1700	54.50***	54.25***	61.25***	41.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Constant				12.75***
				(0.02)
Observations	8124	7649	8124	8124
Pseudo R-Squared	0.0827	0.0821	0.0822	
Adjusted R-Squared				0.0970

Notes: Robust P-Values in parenthesis
*** p<0.01, ** p<0.05, * p<0.1
Constant refers to Peruvian Indios born between 1500 and 1549, age 33 to 72, non-migrant, rural, male, other sources, not accused of being Jewish. The year numbers refer to the beginning of a half century of birth (for example "Mulatto and Black 1650" to those born 1650-99). We scaled the coefficients of all independent variables up by 125, for a more convenient interpretation of changes in numeracy (see Appendix C for the details). We also included dummy variables for "Jew partly" and "Large city unknown", which refer to the Cuenca sample only, "migrant unknown" to this one and the Lima 1700 sample. All year statements refer to the beginning of a half century. "Chile, Bolivia and Argentina 1550" refers to migrants to Lima born in those countries. "Ecuador Indios 1550" refers to migrants coming from the area of today's Ecuador, who mainly identified themselves as "Cañari".

Table 2.4: Logit Regression – Social Inequality in Huanuco

	(1)	(2)	(3)
Dependent variable	Numerate	Numerate	Numerate
Female	-1.25	-1.40	-1.06
	(0.32)	(0.23)	(0.39)
Five tributes		13.38*	
		(0.09)	
Four tributes		6.01*	
		(0.10)	
Three tributes		4.04	
		(0.19)	
Two tributes		-0.69	
		(0.69)	
One tribute		6.36	
		(0.33)	
Mitimaes (forced migrants)	4.11		
	(0.41)		
Yanaconas (slaves)	-1.04		
	(0.59)		
Any tribute			2.65**
			(0.01)
Observations	746	746	746
Pseudo R2	0.37	0.42	0.38

Notes: Robust P-Values in parentheses
*** p<0.01, ** p<0.05, * p<0.1
In Model (1) constant refers to "chupacos" (common population).
In Models (2) and (3) constant refers to men who do not pay any tribute.
Age group dummies for the youngest group included, which had to cover only age 23 – 27 in this case, because the number of unrounded ages was extremely small.

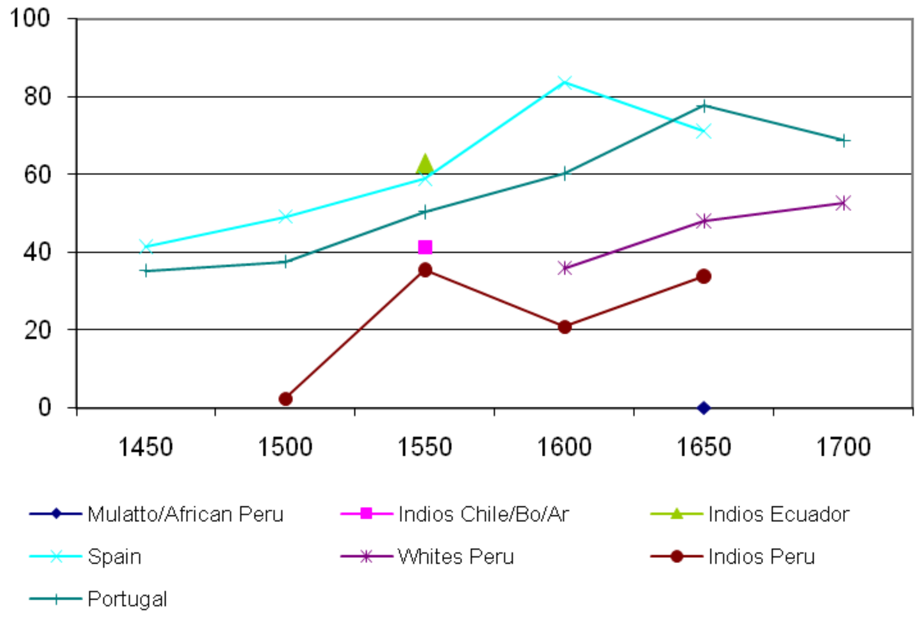
Table 2.5: Logit Regression – Inequality among Social Skill Groups: Full Data Set and Only Indios

	(1)	(2)
Dependent variable	Numerate	Numerate
Population included	All	Only Indios
Reference category	No profession	Semi professionals and professionals
Source: Inquisition	14.63**	
	(0.04)	
Jew	7.26**	
	(0.03)	
Female	-17.13**	-4.05**
	(0.02)	(0.03)
Large city	12.05**	8.08*
	(0.04)	(0.05)
Migrant	7.00*	-42.25**
	(0.05)	(0.02)
Unskilled	15.25**	-7.70*
	(0.05)	(0.05)
Semi-skilled and skilled	17.75**	-14.13**
	(0.05)	(0.05)
Semi-professional	18.38**	
	(0.05)	
Professional	26.25**	
	(0.05)	
Farmer	15.75*	-8.39*
	(0.06)	(0.06)
Skill unknown	17.13**	-13.16*
	(0.04)	(0.07)
Observations	8124	1558
Pseudo R2	0.08	0.15

Notes: Robust P-Values in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Time and country fixed effects are included in both regressions.
In Model (1) reference category is skill group 0 (no profession).
In Model (2) reference category is skill group 4 and 5 together (semiprofessionals and professionals).
Age group dummies for the youngest group included.

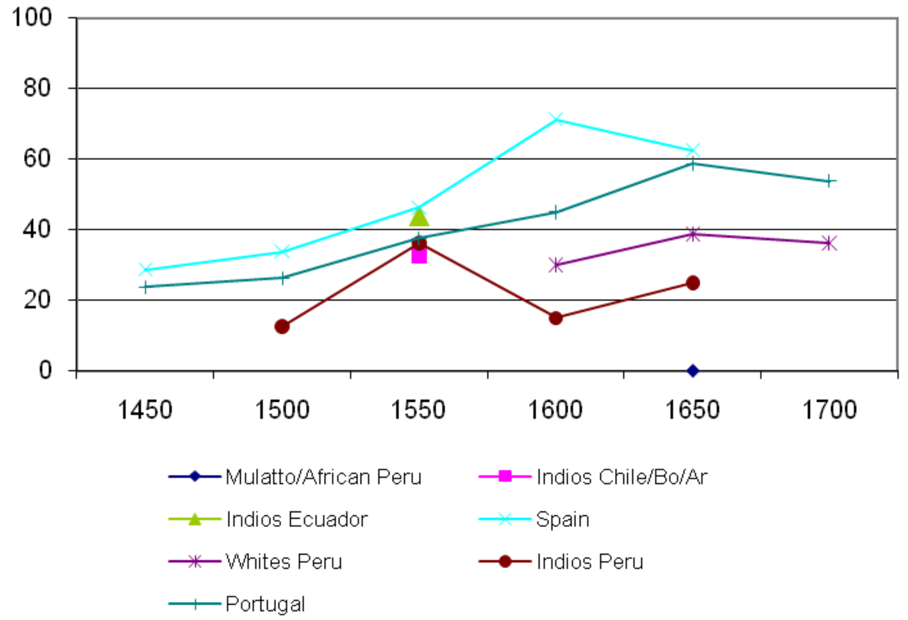
2.10 Figures

Figure 2.1: Raw Averages of ABCC by Half Century, Country and Ethnic Group



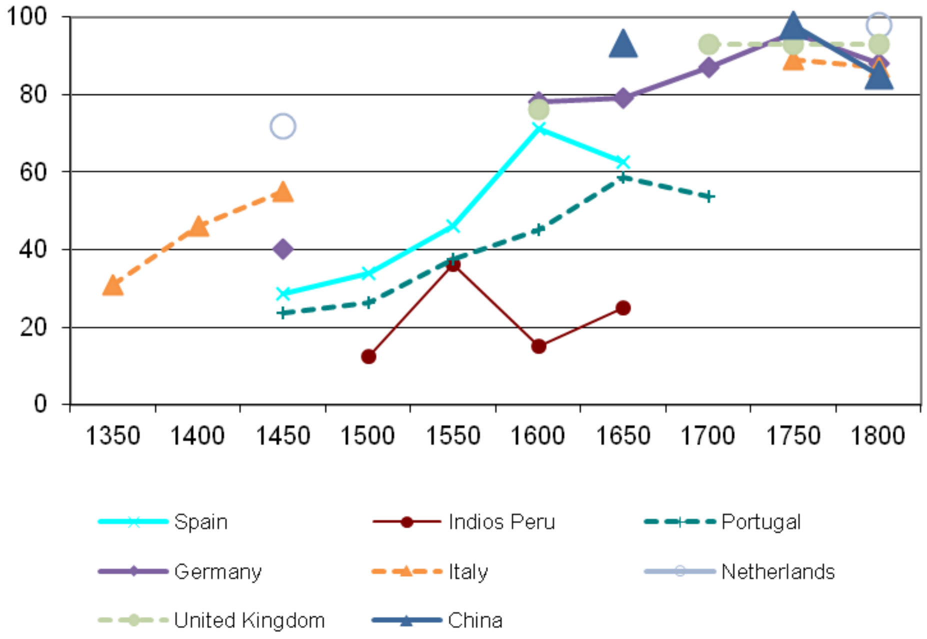
Note: The year on the horizontal axis represents the first year of a half century of birth.

Figure 2.2: Adjusted Estimates of ABCC by Half Century, Country and Ethnic Group



Note: The year on the horizontal axis represents the first year of a half century of birth.

Figure 2.3: Peruvian ABCC in Comparison with Other European Countries and China

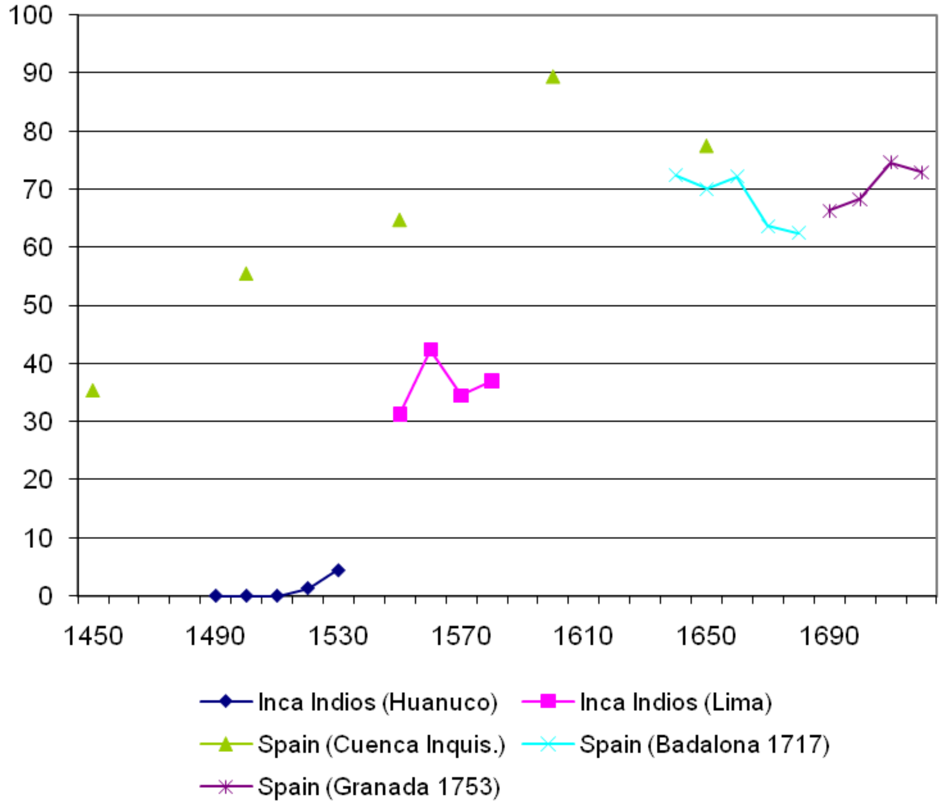


Notes: In general, the years on the axis represent the first year of a half century of birth. China 1650 refers to 1660-1700, China 1750 refers to 150-1759 and China 1800 refers to 1820. Italy refers to Northern Italy before 1750, Germany to Swabia in 1450. The large unfilled circles represent the Netherlands.

Source: A'Hearn et al. (2009)

2.11 Appendix

Appendix A



Numeracy Trends among "Inca" Indios and various Spanish samples. For each birth decade we obtain plausible ABCC values.

Appendix B: Notes on numeracy estimations

Assume that $\frac{1}{m}$ of the population are numerate and that age is uniformly distributed.

$1 - \frac{1}{m} = \frac{m-1}{m}$ is not numerate and will state a multiple of five as their age anyway.

$\frac{1}{5} \cdot \frac{1}{m}$ of the population will correctly and non-accidently report a multiple of five.

In total, $\frac{1}{5m} + \frac{m-1}{m} = \frac{1-5m-5}{5m} = \frac{5m-4}{5m} = 1 - \frac{4}{5m}$ will claim to be a multiple of five years old.

Conversely, $1 - \left(1 - \frac{4}{5m}\right) = \frac{4}{5m}$ will answer with an age that is not a multiple of five.

The fraction of the population assumed to be numerate is recovered by multiplication with $\frac{5}{4}$, since $\frac{4}{5m} \cdot \frac{5}{4} = \frac{1}{m}$.

3. Dangerous Education? The Human Capital of Iberian and Latin American Jews and Other Minorities during the Inquisition

Abstract:

We study the impact of exogenous religious motivations on the generation of human capital. For this purpose, we consider 13,000 victims of the Spanish Inquisition and 47,000 "control group" observations. The age-heaping technique to measure the numeracy component of human capital makes it possible to compare, for the first time, both the Jewish minority and the non-Jewish majority from approximately 1450 to the 19th century. We confirm the hypothesis that the Jewish population had a substantial advantage in numeracy. However, other elite groups of society had a similar level of numeracy (and became targets of the Inquisition). With this study, we contribute to a growing body of literature examining the effect of religion on human capital.

This chapter is based on a paper submitted to the *European Review of Economic History*, co-authored with Jörg Baten (University of Tübingen). The concept of the paper was developed jointly, analyses and writing were equally shared.

3.1 Introduction

The question of whether specific values of religious groups have an exogenous impact on human capital formation is a core topic for growth economics. Max Weber's famous argument that Protestant Christians (particularly Calvinists) are prone to save and to work long hours has been a standard of the literature. Recently, Becker and Woessmann (2009) proposed a modified theory that emphasizes the role of human capital given the Lutheran rule that Protestant children should become able to read the Bible. A number of religious and ethnic minorities, such as the Jains and Parsi in India, the Christians in Syria, and the Chinese in Southeast Asia, have a similar reputation for exceptional educational investment and subsequent economic success. Another remarkable minority group is the Jewish. Botticini and Eckstein (2007) offered a human capital interpretation of Jewish economic history in which they ascribed Jewish occupational selection to a decisive moment in the first century CE. Although the Jewish religion previously had no specific rules about education, in this century, the religious group of the Pharisees gained dominance over the competing Sadducees, and the main religious rule became one that required male individuals to teach their male children to read the Torah. Botticini and Eckstein argued that this was a relatively exogenous religious decision that was not determined by demand because the Jewish remained mostly farmers for another seven centuries. Only in the eighth and ninth centuries, when parts of the Middle East urbanized, could the Jewish finally make use of their exogenously created human capital, and the Jewish elite began to specialize in trade, banking, medicine, and similar professions. Later in the Middle Ages, Christian rulers invited Jewish groups to Northwestern Europe. They promised privileges to these well-educated experts; the legal constraints came later (for a different view, see Roth 1938).

In this study, we aim to measure the human capital differences of religious groups, especially Jews, living in Iberia and Latin America at the time of the Inquisition and until the twentieth century. A data set composed of defendants for various crimes at the courts of the Spanish Inquisition, which was also active in the Spanish colonies, provides insights into these educational differences. The focus on the Spanish-speaking world regions has the advantage that the majority of the population was culturally relatively homogenous. Moreover, both the Jewish minority and the non-Jewish majority can be documented quite well. Lastly, a well-documented census-based comparison sample exists, which has been carefully assessed for its representativeness and the absence of sample selectivity bias (SSB) (Manzel et al. 2012 confirmed the representativeness). In contrast, if we would consider the question whether the Jewish accused by the Inquisition are representative for the Jewish population of Spain during the 15th century, the answer might be indeterminate: we could well imagine that those who left for Amsterdam, for example, were the wealthier and more educated Jewish persons, although empirical evidence is not available.[1] Our study focuses on the majority of previously Jewish who remained in Iberia or the Spanish and Portuguese New World.

The age-heaping technique designed to measure the numeracy component of human capital makes it possible, for the first time, to compare both population groups back to the fifteenth century.[2] We confirm the hypothesis that the Jewish had a substantial advantage in numeracy skills. However, other elite groups of society had a similar level of numeracy (and became targets of the Inquisition). With this study, we contribute to a growing body of literature that has examined the effect of religion on human capital (Chiswick 1988, 1991; Lehrer 1999; Becker and Woessman 2009; Botticini and Eckstein 2007) and other economic outcomes (Barro and McCleary 2006; Guiso, Sapienza, and Zingales 2003). We

[1] The majority of Spanish Jews left for North Africa, and again, no empirical data are available. One could imagine that the selectivity might be less likely to have been positive for the latter migrant group.

[2] The age-heaping strategy considers the share of people being able to report an exact age rather than a rounded age. For the details, see below.

also discuss various hypotheses about Jews' (cultural) motivations for investing immoderately in education.

3.2 Religion and Economics

Max Weber's famous article "Protestant Ethic and the Spirit of Capitalism" (1905) strongly influenced the subsequent literature on religion and economics. His argument that Protestant Christians (particularly Calvinists) are prone to save and to work long hours, thereby contributing to the rise of capitalism, inspired economic theories on why adherents of one particular religious belief are more successful than others (see also Iannaccone (1998) for an extensive review of the literature on this subject). In answer to Weber's seminal work, the economist Werner Sombart (1911) argued that the characteristics Weber ascribes to Protestants apply to a greater degree to Jews. In "The Jews and Modern Capitalism", Sombart hypothesizes that there may be a connection between the shifting of the economic centers from Southern to Northern Europe and the movement of the Jews, beginning with the Jewish populations who were expelled from Spain and Portugal. Although these populations were small in quantity, one can imagine learning effects on the non-Jewish population, knowledge transfer, and crucial entrepreneurial input.[3] For example, at the end of the sixteenth century, Holland enjoyed sudden economic development that coincided with the establishment of a prosperous Jewish community in Amsterdam formed by refugees from Portugal.

More recently, Becker and Woessman (2009) offered an alternative explanation to Weber's theory. Analyzing differences in literacy between nineteenth-century Prussian Protestant and Catholic counties and using instrumental variable methods, they concluded

[3] In a similar vein, Ashraf and Galor (2011) argued that the Jewish and other minorities played a beneficial role for economic development.

that the human capital that was exogenously gained by Lutherans (due to the need to read the Bible) allowed Protestant regions to prosper economically and caused Catholic regions to fall behind.

A number of empirical studies have focused on the relationship between Jewish affiliation and economic outcomes, such as earnings, female employment, or education, during the twentieth century. Chiswick (1983, 1988) and Lehrer (1999) found a higher level of earnings and education as well as higher returns from human capital investments for American Jews relative to non-Jews,[4] and Tomes' (1983) study showed a similar pattern for Canadian Jews. These authors attributed the higher demand of schooling by Jewish Americans to the "diaspora" hypothesis, which holds that, due to their history, Jews favored investments in human capital rather than physical capital because the former is portable and more easily transferable. Moreover, they suggested that Jews have a higher preference for the quality rather than quantity of children; fertility and the rate of young mothers' employment among Jews is relatively low.

In historical studies, not all Jewish population groups have higher literacy or numeracy levels, compared to other religious groups. For example, O'Grada (2006) studied Jewish immigrants to Dublin using the Irish census of 1911 (the birth decades of the adult population represent mostly the mid- to late 19th century). The Jewish had substantially lower values of both literacy and numeracy compared to the Catholic population of Dublin at the same time, not to speak of other religious groups such as protestants. The reason was that the Jews of Dublin were mostly born in the Russian Empire, especially in the Baltic provinces. Educational levels were substantially lower in the Russian Empire during this period, and this also influenced literacy and numeracy of the Jewish minority in the Czarist

---

[4] Jews being defined as second-generation Americans raised in a home in which Yiddish, Hebrew, or Ladino was spoken either in addition to or instead of English

Empire. Compared to Jewish literacy there, the Dublin immigrants were positively selected. This study of the 19th century birth cohorts suggests that the Jewish were not always positively selected, hence it is even more important to assess their human capital empirically.

At the macro-level, a broader series of cross-country empirical studies, including McCleary and Barro (2006), Glaeser and Sacerdote (2008), and Guiso, Sapienza, and Zingales (2003), have recently analyzed the association between religion in general and economic outcomes. Glaeser and Sacerdote (2008) found a positive relationship between education and the extent of religiosity (among Jews and others), which they attributed to the fact that more educated people have a greater incentive to participate in group activities, including church services. However, these authors hold that human capital levels are higher for religious doctrines with low attendance, such as the Jewish today. Guiso, Sapienza, and Zingales (2003), in contrast, argue that the direction of causality runs from religious denominations to educational choice and other attitudes conducive to growth. They conclude that the effect is positive; when comparing religions, Christian religions are, on average, more positively associated with these attitudes than other religions.

Jewish minorities in urban centers of the Muslim world and subsequently in Europe have historically been engaged in relatively high-skilled occupations, such as trade, medicine, governmental administration, and finance (Roth 1995, p. xi).[5] The Jewish community in the Iberian Peninsula existed in Roman antiquity but began to flourish under the Muslim reign of *Al-Andalus* (which lasted in some parts of Spain and Portugal from 711 until 1492). Many Jews from Europe and the remaining Muslim world were attracted by the Muslim Caliphate of Cordoba, where they enjoyed tolerance and economic

[5] We will here refer to Jews as those practicing the Jewish religion, either having being forced to convert yet or not.

opportunities. During the Middle Ages, some Spanish cities became centers of Jewish intellectual and cultural life. After the creation of the Christian kingdoms of Aragon and Castile in the eleventh century, Jews continued to occupy important positions, including the posts of royal counselors and administrators, chief treasurers, and heads of the king's chancellors. Other major occupations for Jews in the fifteenth century included traders, physicians, advocates, tailors, jewelers, and smiths (see Kamen 1965 and Roth 1995). From approximately 1100 on, money lending, a stigmatized activity in medieval Europe, began to play a key role in Jewish economic life (although the specialization of Jews in the financial sector was more pronounced in England, France, Germany, and northern Italy than in the Iberian Peninsula).[6] Farming, however, was not part of their typical professions. In late medieval Spain, Sephardim formed a small middle class living in urban centers that controlled commerce and capital. However, their situation deteriorated dramatically with their expulsion from Spain and Portugal (in 1492 and 1497, respectively) and declined even further when Philip II supported a terrible type of Inquisition (Sombart 1911).

A number of scholars have attempted to identify the reasons for the selection of Jews into urban high-skilled occupations. Some historians, such as Cecil Roth (1938) and Solomon Katz (1937), argue that *restrictions* imposed on Jewish minorities by local rulers restrained them from engaging in agricultural activities and encouraged them to specialize in trade and arts and, later, in money lending.[7] An alternative explanation was suggested by

[6] Making profits from money lending was condemned by the Catholic Church. Classical writers conceived profits from trade - and more so from lending money - as morally suspect ("sinful") for not being derived from physical labor. Thus, Christians were not allowed to engage in this sinful economic activity. In the late Middle Ages, commerce increased due to agricultural surpluses and urbanization developed. Under those circumstances, money-lending became more necessary for the expanding European economy. In order to resolve this dilemma, Jews were allowed in the twelfth century to engage in the forbidden activity of money-lending. Soon Jews became the indirect tax collectors for the royal treasuries, at times when the nobility and the clergy were exempt from royal taxation (Muller 2010).

[7] Nevertheless, this view was rejected by Botticini and Eckstein (2005, 2007), who argue that at the time and places the occupational transition occurred, no constraints prevented Jews from owning land (or engaging Christian laborers to work on their farms).

the economist Simon Kuznets (1960, 1972),[8] who attributes the engagement of Jews in non-farming occupations to a non-economic decision common to small *minorities*. To maintain their group identity and customs, the Jewish community (like other small minorities, such as the Parsi in India or the Huguenots in early modern Western Europe) preferred to be concentrated in selected industries, which happened to be urban occupations. Before Kuznets, Max Weber had hypothesized that Jews voluntarily segregated into certain occupations to correctly observe their strict religious rituals, which, in his view, was a trait common to all religious minorities (Botticini and Eckstein 2012). In a similar vein, Baron, Gross, and Kahan (1975) relate the fact that Jews in the diaspora turned to mercantile activities to a common sociological phenomenon among immigrants, familiar with more than one language and culture, which allows them to mediate between distant places. Another line of scholars, including Werner Sombart (1911), maintain that Jews and other persecuted religious or ethnic minorities preferred to invest in *human capital* rather than physical capital because it could less easily be expropriated. Therefore, Jews did not invest in land, and their educational advantage allowed them to engage in highly skilled urban occupations. Botticini and Eckstein (2007, 2012), in contrast, argue that the Jewish community enjoyed an *exogenously gained high level of human capital*, which provided them with a comparative advantage in entering professions that required a considerable level of education. These authors attribute the advantage in terms of education to a religious law issued in the first century CE by the religious group of the Pharisees that forced Jewish fathers to teach their male children to read Hebrew or to send

[8] Kuznets analyzes the occupational structure of the Jewish population in nineteenth century Eastern Europe and North America.

them to school to do so.[9] During this time, the Pharisees became the dominant group, and Judaism transformed from a religion of "sacrifices in temples" to a religion of the study of the Torah in the synagogue. Although not its goal, this religious emphasis on literacy was a precondition for the specialization of Jews in highly skilled occupations when urbanization progressed in the eighth century Muslim world. The ability to read and write constituted an advantage when engaging in occupations such as commerce, crafts, medicine, and finance. Because most Jewish farmers were literate whereas most of the non-Jewish population was not, the former could more easily enter high-skilled and well-paid professions. When they dispersed in Europe, Jews had already specialized into urban professions and continued to pursue these professions in the diaspora.

3.3 "Judaizers" and the Modern Inquisition[10]

The Inquisition was originally a tribunal introduced by the Roman Catholic Church that was meant to discover and abolish heresy to the unity of Christendom. It was founded at the synod of Toulouse in 1229 and first operated in Bohemia, France and Italy. The so-called Modern Spanish Inquisition was officially established by the Catholic Monarchs Isabella I of Castile and Ferdinand II of Aragon in 1478 after the unification of both kingdoms and was suppressed definitively in 1834 under Queen Isabella II of Spain.[11] Since the first tribunal, based in Seville, commenced its activity in 1480, more than

[9] As the result of the destruction of the Jewish commonwealth and the Second Temple in the year 70 CE, Rabbis were required to face a new reality in which the old system of oral scholarship could not be maintained. It is during this period that rabbinic discourse began to be recorded in writing.
[10] The term "Modern Inquisition" is used here to differentiate this institution from the apostolic Inquisition that existed in parts of the Iberian Peninsula before the ascendency of the Catholic monarchs. More about this distinction is explained in the text.
[11] It had been abolished for the first time during Napoleon's dominion with Joseph Bonaparte on the Spanish throne in 1808 and was re-established and abolished several times until its definitive suppression.

100,000 trials were conducted against so-called Judaizers,[12] converted Muslims (or *moriscos*),[13] Protestants, and other "heretics" (Vidal-Robert 2011, p. 3). This institution differed from the preceding one that had been established in the kingdom of Aragon in 1232 to address Catharism insofar as it was uniquely responsible to the Crown because the monarchs had been granted legitimacy over the Holy Court by Pope Gregory IX. Its role at that time was mainly to assure the orthodoxy of New Christians who converted from Judaism and, later, Islam, especially after mass conversions followed the royal decrees of expulsion for Jews and Muslims issued in 1492 and 1501, respectively. Although the court dealt almost only with converted Jews at its origin,[14] with the rise of Lutheranism in Germany at the beginning of the sixteenth century, it turned its focus to religious and intellectual reformers and so-called "minor heresies".[15] After the menace of Protestantism was mainly suppressed in Iberia, the prosecution of Jewish converts again became the main duty of the tribunal from the 1560s forward. In comparison to other Inquisitorial courts, for example, the tribunal in the Canary Islands tried a much larger share of Protestants – all of them foreigners[16] – given the active relations with Dutch and German traders and high immigration in the archipelago (Fajardo Spinola 2005, p. 113).

The Inquisition courts soon acquired a reputation for being a repressive instrument of racial and religious intolerance that regularly employed torture, restricting Spain's

---

[12] "Judaizer" is one of the terms used to refer to Jews who converted to Christendom and continued to practice their former religion. The Inquisition did not have any jurisdiction over actual Jews, as they were not members of the Catholic Church and therefore could not be accused of heresy. One exception was that of the Jews alleged of proselytism, which means those specifically accused of influencing Christians to convert (Roth 1995, p. 213).

[13] Muslims who converted to Christianity mostly following the "Reconquista" suspected of secretly practicing Islam were called *Moriscos* (meaning "Moorish").

[14] Also called *conversos.* This term is used to refer to Jews who had voluntarily converted to Christianity and were mostly regarded as secretly practicing their former religion. Because of this, also the term "crypto-Jews" (*gr.* secret Jews) has been ascribed to them.

[15] Rawlings (2006) uses the term "minor heresies" to refer to the behavioral attitudes that attempt against Christian rules, such as bigamy, superstition, etc.

[16] The processes against foreign Protestants, mainly Englishmen, contributed to the creation of the Black Legend myth in Europe.

intellectual development for centuries (Rawlings 2006, p. 1). Some scholars have maintained that in addition to creating an atmosphere of fear in Spain, the main objective of the court was to extract wealth for its financing (Llorente[17] 1822) or to serve as an instrument of repression to avoid internal revolts (Vidal-Robert 2011).

Until the end of the fifteenth century, the Spanish population was of mixed confessions (Muslim, Christian, and Jewish). Jews already played an important role in the economic and cultural life of Spain under the Visigoths; they had prospered socially during the Liberal Muslim Caliphate and occupied important positions in the public administration during the "Reconquista" in Christian Spain. However, anti-Semitic tensions had long been latent and worsened during the fourteenth century, leading to the massacres of 1391 and 1412. From then on, several pogroms brought about forced mass conversions and culminated in the issuance of a royal Edict of Expulsion of the Jews in 1492 (Roth 1995, Netanyahu 1995). Kamen (1965) attributes the growing anti-Semitic tendency to the hostility of the feudalistic nobility against a rising middle class that controlled commerce and capital and that mainly consisted of Sephardim. Furthermore, Jews' engagement in financial activities – which were condemned by the Christian church – and their supposed "unwillingness to take part in manual labor" (Kamen 1965) might have turned the hostility of the Christian population against them. However, the Catholic Monarchs' official reason for the expulsion was their desire to preserve Christian unity. As a result, of approximately 80,000 Sephardic Jews living in Spain in 1492, an estimated 40,000-50,000 left for exile; most of them likely returned a few years later and converted to Christianity (Rawlings 2006, p. 66).[18] Because the Jews who remained in Spain or

---

[17] Juan Antonio Llorente was the last secretary and the first historian of the Spanish Inquisition (Kamen 1965, p. 49).

[18] In their majority they emigrated to Portugal, Navarre, North Africa, Turkey, Italy and Western Europe (Rawlings 2006).

returned from exile had been forced to convert, they were suspected of secretly practicing Jewish rites. Thus, distrusted converts came to be considered crypto-Jews (from the Greek "kryptos", meaning "hidden"), or *marranos*.[19] The edict of expulsion resulted in more potential victims of the Holy Office, which pursued Jewish converts accused of being false Christians.

In Portugal, the persecution of crypto-Jews was initially much less severe than in Spain, although they had been forced to convert in 1497. In fact, many secret Jews who had fled from Spain in 1492 continued to practice their religion in the neighboring country with relative freedom until the Portuguese throne was taken over by the Spanish monarch between 1580 and 1640. In this period, the Holy Office courts of Coimbra, Lisbon, and Evora became more aggressive against "Judaizers" and other heretics. This situation forced many Jews to migrate back to Spain or to make their way to the New World through Brazil, where the immigration restrictions for Jews were more lenient than in the Spanish viceroyalties. A relatively large community of Jews deported by the Portuguese Inquisition had already settled there. Furthermore, a few relaxations of the laws barring the migration of New Christians to the New World – for example, for only four years in 1514 – allowed the establishment of a New Christian society on the American continent (Liebman 1974, p. 20).

The Spanish Inquisition had offices not only in the Iberian Peninsula and the Balearic and Canary Islands but also throughout the Spanish Empire. In America, the Holy Office was officially established in 1569 and was subordinate to the authority of the Spanish monarchs. During the first decades of its existence, the Inquisition in Latin America was particularly occupied with American Indians, who worshipped non-Christian

[19] The word "marrano" in Spanish has a negative connotation, as it is also used to refer to pigs. It is unclear where this word comes from. Probably it is derived from a Hebrew word for "the true faith" (Kamen 1965, p. 15).

idols, and with bigamy. Bigamy was a defined crime against Christian rules and hence a target of the Inquisition. The Spanish conquistadores and their followers during the late sixteenth century were often already married in Spain but tended to marry a second time in the New World. In particular, daughters of the ruling class of the Aztec Empire occasionally became second wives. At the same time, the Catholic Church was struggling to convince American Indians to maintain their religious rules. The religious bureaucracy had no sympathy for Spanish bigamists who presented "bad examples" by uniting with Aztec princesses.

After a number of complaints from bishops in Havana and Puerto Rico about heretics arriving in the New World (despite the existence of decrees since the end of the fifteenth century preventing the entry of New Christians), the first tribunals were established in Mexico City and Lima, each with jurisdiction over the complete extension of the respective viceroyalty, New Spain and Peru.[20] In 1610, a third tribunal was established in Cartagena de Indias. The presence of Jews and New Christians in the Spanish viceroyalties was illegal during the entire colonial period, and it was the task of the Inquisition to keep them out. Only descendants of "Old Christians", who were entitled to certificates of purity of blood, were allowed to migrate to the New World (Liebman 1974: 18). Despite these official rules, New Christians found their way to the colonies and became the principal duty of the American Inquisition tribunals (Greenleaf 1969, Wachtel 2007).

Inquisitorial sources are a useful instrument for the historiography of the social and cultural situation of Sephardim given that a large amount of information was collected.

[20] The viceroyalty of New Spain included all of modern Mexico and Central America, as well as the southwestern part of the present United States and the Philippines (Liebman 1970, p. 16). The jurisdiction of the Holy Office comprehended the same territory. The viceroyalty of Peru included modern-day Peru, Chile, Colombia, Panama, Ecuador, Bolivia, Paraguay, Uruguay and Argentina.

Defendants were subjected to rigorous interrogations. They had to provide information on their family background, place of birth, occupation, and (occasionally) literacy and, depending on the accusation, to reveal details of their religious beliefs and practices or moral behavior.

3.4 Method and Data

3.4.1 Age Heaping

The method used to analyze education levels is the "age-heaping" technique. The age-heaping phenomenon applies to historical populations (as well as people in the poorest countries today), in which a substantial share of the people were unable to state their exact age and hence reported a rounded age, such as "I am 30", when they were in fact 29 or 31, for example. How is age heaping calculated? The ratio between the preferred ages and others can be measured by several indices, the most common of which is the Whipple index.[21] This index measures the proportion of people who state an age ending in a five or zero, assuming that each terminal digit should appear with the same frequency in the "true" age distribution (or the degree to which the distribution of age statements approaches an equal distribution).[22]

$$(1)\ Wh = \left(\frac{(Age25 + Age30 + Age35 + ... + Age60)}{1/5 * (Age23 + Age24 + Age25 + ... + Age62)} \right) \times 100$$

[21] A'Hearn, Baten and Crayen (2009) found that this index is the only one that fulfills the desired properties of scale independence (a linear response to the degree of heaping), and that it ranks samples with different degrees of heaping reliably. It is also used by the UN Statistics Division.

[22] A value of 500 means an age distribution with ages ending only on multiples of five, whereas 100 indicates no heaping patterns on multiples of five, that is exactly 20 percent of the population reported an age ending in a multiple of five.

For an easier interpretation, A'Hearn, Baten, and Crayen (2009) suggested another index, which we call the ABCC index.[23] This is a simple linear transformation of the Whipple index and yields an estimate of the share of individuals who correctly report their age:

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

The share of persons able to report an exact age has been shown to be highly correlated with other measures of human capital, such as literacy and schooling, across countries and individuals and over time (Mokyr 1983; A'Hearn, Baten, and Crayen 2009; Crayen and Baten 2010).

How large was the educational gap between the average population in the Iberian Peninsula and the Sephardic Jewish people in the diaspora? How did this gap evolve over time? This is the first economic study to attempt to quantify the human capital levels of this religious minority and to compare them to the average gentile population of Spain, Portugal, and Latin American countries in a long-term period.

For this purpose, we constructed a large database that included 1) a sample based on Inquisition trials from Spain, Portugal, and Latin America and 2) a comparison sample based on non-Inquisition sources. The second sample consisted of Latin American census records compiled by Manzel, Baten, and Stolz (2012) and both census and death records from Spain and Portugal. In Table 3.2, we present the date and place of recording of the single sources as well as the number of observations contained in each source.

Sample selection bias in the Spanish and Portuguese "control group" samples based on census and mortality cannot be very substantial because these sources aimed to include the entire population at the year of the census taking or death (in Spain and Portugal,

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

almost everyone was Catholic and was entered in the death registers). The Inquisition sources clearly represent a socially selected group, which is our focus of analysis. There may be a regional selectivity issue. Given the regional differences in education, especially in Spain, we carefully verified that our data were representative for Spain and Portugal as a whole by comparing the regional distribution of observations in our dataset with the regional distribution of the actual population in censuses of the eighteenth and nineteenth century, respectively. In Table 3.2, we assess the regional distribution of individuals included in our dataset. We compare the samples with the earliest censuses of Spain (Floridablanca 1787) and Portugal (national census on 1864).[24] How representative is our dataset for the population of Portugal and Spain? In general, the Spanish source seems quite representative by region. The larger units of the country have similar shares in the census and in this sample. The Castilian South and Granada were overrepresented in Inquisition sources because the previous density of Jewish and Muslims was higher. For Portugal, in contrast, the central region of the country is somewhat overrepresented, whereas the islands and the North are slightly underrepresented relative to the census evidence. This issue raises the question of whether numeracy in the center differs substantially from the latter two regions. We calculated numeracy for the period of 1500 to 1699 and found similar ABCC levels for all regions. The center has a numeracy level of 62, the north is only four points less at 58, and the south has 54 ABCC points. The center includes the largest city of Lisbon, which explains the numeracy advantage. However, the islands display even higher numeracy, with an ABCC of 73. This finding is not surprising because the Acores and Madeira were wealthy sugarcane producers and enjoyed high

[24] For Latin America, comparable census evidence across all regions is not available, but Manzel et al. (2012) have assessed the representativeness of their sample and rejected the existence of sample selection bias (SSB).

cultural diversity in the fifteenth to seventeenth centuries because they attracted many European merchants.

Our Jewish sample is most likely representative of the total Jewish population (despite being drawn from Inquisitorial records) because all Jews and New Christians who remained in the Iberian peninsula after the edict of expulsion or fled to Spanish America were potential victims of the Inquisition court, independent of their educational level and a large share was actually accused.

3.4.2 Religion, Elites, and Human Capital

As noted above, Inquisitorial trial documents provide comprehensive personal information on the defendants, such as their age, birthplace, occupation, marital status, and religion. The countries covered by the latter sources are Spain, Portugal, Mexico, Peru, Argentina, and Chile. Of the 13,529 individuals accused at Inquisition trials whose ages we could obtain from archival sources, 12,065 cases include information about the alleged crime. We classify the Inquisitorial crimes into categories based on Contreras and Henningsen (1986) in Table 3.3. [25] Judaism, Mahometism, and Protestantism were the three "major heresies". The category of Protestantism included Calvinism, Lutheranism, Erasmism, Illuminism, and Freemasonry because these schools were perceived as similar to each other in educational orientation. A number of victims of the Inquisition were tried in the sixteenth century for various "crimes" at the same time. For example, Maria Cazalla was arrested and accused of Lutheranism, Illuminism, Molinism (a Catholic theological line named after the Jesuit theologian Luis de Molina that claimed that divine grace and human free will are not mutually exclusive), and Erasmism in 1532 (see Kamen 1965, p. 73). "Minor heresies" were polygamy, other crimes related to sexual promiscuity (such as

[25] See also Rawlings (2006, p. 13).

sodomy, cohabitation with a partner, or infidelity), blasphemy and propositions,[26] superstition and witchcraft, offenses against the Holy Office or impeding its correct functioning (including false testimonies at the court or not serving an imposed sentence), and being an accomplice to a crime. Contreras and Henningsen include the category of “solicitation” in the “minor heresies”. “Solicitation” occurred if a priest seduced a woman during confession. It was considered by contemporaries to be a particularly heinous crime because women had to confess their sins (including those related to sexuality) to the priest, who should be a person who could be trusted (and who should be celibate). If the priest abused this position to seduce a woman during confession, the Inquisition applied particular attention to this persecution. We have ascribed this type of crime to a category we call “clergy crimes”, which includes seducing priests and preachers pretending to be clergymen. This category is particularly interesting because it represents Catholics who belonged to the presumably best-educated (and substantial) social group of early modern society. Hence, we can compare the other religious group with this Catholic elite. Naturally, the relative frequency of the different crimes addressed by the Inquisition varied significantly between the Inquisitorial districts and across time. For example, whereas the court originally dealt primarily with converted Jews, with the rise of Lutheranism in Germany at the beginning of the sixteenth century, it turned its attention to religious and intellectual reformists. Minor heresies became more frequent in the Inquisition records from the second half of the sixteenth century (see Appendix A.3).

We hypothesize that people accused at the Inquisition were, on average, more numerate than the rest of the population. The accused were largely “major heretics”, including Judaists, Protestants, and other spiritual devotees. Moreover, Illuminists,

[26] From the 1560s on, the Holy Office began assuming the responsibility to intervene in the correction of unorthodox beliefs and behavior of Old Christians. The most common propositions would include denying the existence of God.

"Erasmians", Lutherans, and other Protestants were usually individuals who had thought critically about theological (and political) issues and adopted innovative views on spiritual and intellectual life. They were often familiar with devotional literature censored by the Inquisition or had even written it themselves (Rawlings 2006, pp. 90–113; Kamen 1965). It has also been argued that successful Jewish or Protestant merchants and other professionals were accused by their competitors, or even that the Inquisition profited from persecuting wealthy "heretics" by confiscating their goods (Fajardo Spinola 2005, p. 113, Rawlings 2006, p. 42). In fact, expropriation was a common punishment, in addition to wearing the *sanbenito* (penitential cloth), scourging, jail, and death. For these reasons, one could imagine that a large share of the victims of the Inquisition were relatively wealthy and educated. When examining the average ABCC index for every crime category and the ABCC index for the complete comparison sample (without controlling for time differences), we can observe consistently higher numeracy for all crime categories when compared to the non-Inquisition sample. In Table 3.4 (Model 5) and Figure 3.1, we calculate the deviation of ABCC levels from the comparison sample. The Catholic elite had the greatest numeracy, as observed from the high values for solicitation (and false clergy) and detractors of the Holy Office. Jews and Muslims in the Inquisition had a substantial advantage over the (Christian) control group. The smallest numeracy advantage was evident for (other) sexual crimes, polygamy, and witchcraft. However, this descriptive figure may be misleading because the Jewish "crimes" were overrepresented in the early period, whereas numeracy was lower in general compared to later periods. In the following regressions, controls for time fixed effects are necessary.

3.5 Logistic Regression Results

We tested these descriptive differences in logistic regressions, controlling for other potential determinants of numeracy. We used a binary dependent variable that takes the value of 1 if the age reported was not a multiple of five and 0 otherwise. The results in Table 3.4 (Model 5) show that, in fact, individuals included in most crime categories performed significantly better than the (average) comparison population in terms of numeracy.[27] Thus, the hypothesis above can be confirmed. The difference in numeracy between "Judaizers" and the comparison sample is ten percent if we do not control for the fact that the control group includes a substantial share of individuals drawn from Inquisitorial records (Table 3.4, Model 1).[28] In column (2), we compare those accused of Jewish rites with the (positively selected) rest of the Inquisition sample. Even compared to this positive selection, those accused of being "Jewish" had a five percent advantage (Model 2).

We assumed that the most educated individuals – excluding "Judaizers" – would be those included in the crime categories of "clergy crimes",[29] Protestantism, and "detractors of the Holy Office" (as shown in the results of Model 5) as well as document falsifiers and prohibited book keepers, who were often included in the categories of "accomplice" or "witchcraft and superstition". To segregate the effect of this intellectual elite, we included an independent dummy variable ("elite crimes") in the regression (Table 3.4, Model 2-4). When controlling for migrant status, the coefficient for "Judaizers" was 14 percentage points over the average population (Model 3). In general, the coefficients of "Judaizers" and elite crimes were large (relative to the control group of the "average" population).

---

[27] We multiplied the coefficient of the Logit regressions by 125 as to adjust it for the 20% of ages that should truly be multiples of five given a normal distribution (see Baten and Juif 2013 and Appendix B for details). Furthermore, in all regressions, except Model 2, we controlled for the sample source being death registers, which did not change our coefficient of Judaism.

[28] Although our sample of Jews is drawn only from Inquisitorial sources, we have good reasons to assume that their numeracy is representative for the Sephardic Jewish population in general, given that most *conversos* practiced Jewish rites secretly and were potential victims of the Inquisition. Moreover their occupational structure coincides with the one ascribed in the literature to the Jewish population in Spain and Portugal at the time.

[29] The idea behind it is that clergymen have obtained a higher education (and "false clergymen" have enough knowledge to effectively pretend being a priest). Of course, people accused of practicing a more reformist confession have at least thought critically about religion and are therefore expected to have higher than average cognitive abilities.

3.6 Cross-Check: Literacy of Inquisition Victims

To counter-check our results, we compare them with literacy evidence available for the Inquisition court of the Canary Islands. In the original records, some courts recorded literacy, but the editions we used did not report it.[30] Wachtel (2007, p. 23), for example, observes a significantly higher level of instruction of individuals accused of "Judaizing" in Mexican Inquisition trials in the seventeenth century compared to the other defendants at this court: all of the men and half of the women could read and write. Furthermore, more than 80 percent were engaged in commercial activities. Our analysis of the records of the Inquisition in the Canary Islands also points to an important advantage of "Judaizers" in literacy skills (see Table 3.5). Whereas 82 percent of those accused of "Judaizing" could read and write, only 42 percent of the remaining defendants were literate. Because literacy is an important component of human capital and because Jews engaged in professions that required a certain amount of numeracy, we expect the Jewish population in Iberia and Spanish America to show a higher than average educational level. The occupational structure of the "Judaizers" included in our data set, reported in Table A.2, fits the expected pattern.

3.7 Cross-Check: Were Ages Determined Differently in Inquisition Sources?

Were age questions asked differently during Inquisition trials compared to census data collection efforts? In principle, we cannot reject the possibility that ages were reported in a different way, although there is no evidence that this was the case. Asking for age was most likely the least exciting part of an Inquisition trial because ages did not play a role in

[30] For example, in the publication of the Logroño Inquisition trials (Henningsen 2004, pp. 397-437) no literacy information is included.

the decision; even very young or very old people were treated similarly, and if mercy toward younger victims was given by some individual Inquisition officials, the difference between, for example, an age of 30 or 31 did not play a role. Are there indirect ways to assess whether Inquisition courts asked age questions differently? One possibility would be to compare basic numeracy based on age statements of priests and the average using census records only. Even if the normal village priest might not be identical to priests accused by the Inquisition, a similar magnitude of the differential of numeracy might suggest that age questions were asked in a similar way by the Inquisition and that numeracy differences are plausible. We consider the Spanish Ensenada census (which reflects late 17th and early 18th century birth cohorts) a source for this and find that priests had a numeracy of 99, whereas the average population had a numeracy of 77. This finding indicates that the order of magnitude of the difference of those accused of elite crimes and the average population was not unrealistic (compare Table 3.4).

3.8 Conclusion

In this study, we contributed to a growing body of literature that has recently examined the effect of religion on human capital. We confirmed the hypothesis that those accused as "Judaizers" by the Inquisition enjoyed a substantial advantage in numeracy skills. We quantified the educational gap in terms of numerical and reading abilities between this group and the Iberian and Latin American population for a time span of nearly 400 years. Their accumulated human capital over centuries continued to distinguish them from their Catholic compatriots for as late as we can measure. What is the potential effect of this advantage to the average population? In this context, we cannot seriously speculate about the size of spill-over effects of Jewish human capital to the rest of the population. However, if Inquisition would not have been so dangerous for this highly educated group, economic growth would probably not have been hampered as much.

We also find that other elite groups of society who became targets of the Inquisition had a substantial educational advantage. This finding implies that the Inquisition – unintentionally – provided very strong incentives against human capital acquisition, both among Christian and previously Jewish population segments. A significant retardation of economic development should be the consequence, as has often been speculated. However, we provide quantitative evidence for this speculation for the first time.

3.9 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*, 69(3): 783-808.

Alberro, S. (1988). *Inquisition et société au Mexique, 1571-1700*. Mexico City: Centre d'études mexicaines et centraméricaines.

Alvarez Alonso, F. (1999). *La Inquisicion en Cartagena de Indias durante el siglo XVII*. Madrid: Fundacion Universitaria Espanola.

Aranda Mendiaz, M. (2000). *El Tribunal de la Inquisición de Canarias durante el Reinado de Carlos III*. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria.

Ashraf, Q., and Galor, O. (2011). Cultural Diversity, Geographical Isolation, and the Origin of the Wealth of Nations (No. w17640). National Bureau of Economic Research.

Baron, S.W., Kahan, A., and Gross, N. (1975). *Economic History of the Jews*. New York: Schocken Books.

Barro, R. J. and McCleary, R. M. (2006). Religion and Economy. *Journal of Economic Perspectives*, 20(2): 49-72.

Becker, S. O., and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, 124(2): 531-596.

Bennassar, B. (1979). *L'Inquisition Espagnole xv-xix Siècle*. Paris : Hachette

Botticini, M., and Eckstein, Z. (2005). Jewish Occupational Selection: Education, Restrictions, or Minorities?. *Journal of Economic History*, 65: 922-948.

Botticini, M., and Eckstein, Z. (2007). From Farmers to Merchants, Conversions and Diaspora: Human Capital and Jewish History. *Journal of the European Economic Association*, 5(5): 855-926.

Botticini, M., and Eckstein, Z. (2012). *The Chosen Few: How Education Shaped Jewish History, 70-1492*. Princeton University Press.

Chiswick, B. R. (1983). The Earnings and Human Capital of American Jews. *Journal of Human Resources*, 18(3): 313–336.

Chiswick, B. R. (1988). Differences in Education and Earnings Across Racial and Ethnic Groups: Tastes, Discrimination, and Investments in Child Quality. *The Quarterly Journal of Economics*, 103(3): 571-597.

Chiswick, B. R. (1991). Jewish Immigrant Skill and Occupational Attainment at the Turn of the Century. *Explorations in Economic History*, 28(1): 64–86.

Chiswick, C. U. (2006). An Economic Perspective on Religious Education: Complements and Substitutes in a Human Capital Portfolio. *Research in Labor Economics*, 24: 419-467.

Contreras, J. (1982). *El Santo Oficio de la Inquisicion de Galicia*. Madrid: Akal Universitaria.

Contreras, J., and Henningsen, D. (1986). *Forty-four thousand cases of the Spanish Inquisition (1540-1700): Analysis of a historical data bank*, in (eds.) Gustav Henningsen and John Tedeschi: "The Inquisition in Early Modern Europe. Studies in Sources and Methods". Dekalb, IL: Northern Illinois University Press, 100-129.

Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Term Growth. *Explorations in Economic History*, 47(1): 82-99.

Fajardo Spinola, F. (2005). *Las Víctimas de la Inquisición en las Islas Canarias.* La Laguna: Editorial Universidad de La Laguna.

Garcia Fuentes, J. M. (2006). *Visitas de la Inquisición al Reino de Granada*. Granada: Editorial Universidad De Granada

Glaeser, E. L., and Sacerdote, B.I. (2008). Education and Religion. *Journal of Human Capital*, 2(2): 188-215.

Greenleaf, R. E. (1969). *The Mexican Inquisition of the Sixteenth Century*. Albuquerque: University of New Mexico Press.

Guiso, L., Sapienza, P., and Zingales, L. (2003). People's Opium? Religion and Economic Attitudes. *Journal of Monetary Economics*, 50(1): 225–28.

Henningsen, G. (1980). *The Witches´Advocate – Basque witchcraft and the Spanish Inquisition (1609-1614)*. Reno: University of Nevada Press.

Henningsen, G. (2004). *The Salazar Documents, Inquisitor Alonso de Salazar Frías and Others on the Basque Witch Persecution*. Leiden, Boston: Brill Publishers.

Iannaccone, L. R. (1998). Introduction to the Economics of Religion. *Journal of Economic Literature*, 36(3):1465-1495.

Kamen, H. (1965). *The Spanish Inquisition*. London: Weidenfeld and Nicolson.

Katz, S. (1937). The Jews in the Visigothic and Frankish kingdoms of Spain and Gaul. The Mediaeval Academy of America.

Kuznets, S. (1960). "Economic Structure and Life of the Jews." In *The Jews: Their History, Culture, and Religion*, edited by Louis Finkelstein. Jewish Publication Society of America.

Kuznets, S. (1972). Economic Structure of U.S. Jewry: Recent Trends. Working paper, Institute of Contemporary Jewry, Hebrew University of Jerusalem.

Lea, H.C. (1907). *A History of the Inquisition of Spain*. New York: The MacMillan Company.

Lee, R. (2007). *The Columbia History of Jews and Judaism in America*. New York: Columbia University Press.

Liebman, S. B. (1970). *The Jews in New Spain: Faith, Flame, and the Inquisition*. Coral Gables: University of Miami Press.

Liebman, S. B. (1974). *The Inquisitors and the Jews in the New World. Summaries of Procesos 1500-1810 and Bibliographical Guide*. Coral Gables: University of Miami Press.

Lehrer, E. L. (1999). Religion as a Determinant of Educational Attainment: An Economic Perspective. *Social Science Research*, 28(4): 358-379.

Llorente, J. A. (1822). *Historia crítica de la Inquisición de España*, Vol. 3. Madrid: Imprenta del Censor.

Manzel, K., Baten, J., and Stolz, Y. (2012). Convergence and Divergence of Numeracy: the Development of Age Heaping in Latin America from the Seventeenth to the Twentieth century. *Economic History Review*, 65(3): 932-960.

Medina, J. T. (1952). *Historia del Tribunal del Santo Oficio de la Inquisicion en Mexico*. Santiago de Chile: Impr. Ercilla.

Mokyr, J. (1983). Why Ireland Starved: A Quantitative and Analytical History of the Irish Economy, 1800-1850. London and Boston: Allen and Unwin.

Muller, J. Z.(2010). *Capitalism and the Jews*. Princeton: Princeton University Press.

Netanyahu, B. (1995). *The Origins of the Inquisition in Fifteenth-Century Spain*. New York: Random House.

Ó Gráda, C. (2006). Dublin Jewish Demography a Century Ago. The Economic and Social Review, 37(2): 123-147.

Palma, R. (1863). *Anales de la Inquisición en Lima: Estudio Historico*. Lima: Ediciones del Congreso de la Rebública.

Prado Moura, A. (2003). *El Tribunal de la Inquisición en Espania (1478-1834)*. Madrid: Editorial ACTAS.

Rawlings, H. (2006). *The Spanish Inquisition*. Malden, USA: Blackwell Publishing.

Rico Medina, S. (1990). *Los Predicamientos de la Fe, La Inquisición en Tabasco, 1567 – 1811*. Gobierno del Estado de Tabasco.

Roth, C. (1938). *The Jewish Contribution to Civilization*. London: Macmillan & Co.

Roth, N. (1995). *Conversos Inquisition and the Expulsion of the Jews from Spain*. Madison, WI USA: The University of Wisconsin Press.

Sierra, J. (2006). *Procesos de la Inquisición de Toledo (1575-1610). Manuscrito de Halle*. Madrid: Ed. Trotta.

Sombart, W. [1911]. *The Jews and Modern Capitalism*. Kitchener: Batoche Books, 2001.

Tomes, N. (1983). Religion and the Rate of Return on Human Capital: Evidence from Canada. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 16(1): 122-138.

United Nations Statistic Division: http://unstats.un.org/unsd/demographic/products/dyb/DYBcensus/V1_Notes1c.pdf

Vidal-Robert, J. (2011). An Economic Analysis of Spanish Inquisition's Motivations and Consequences. Working Paper.

Villanueva, J. P., and Bonet, B. E. (1984). *Historia de la Inquisición en España y América*, tomo III. Madrid: Biblioteca de Autores Cristianos.

Wachtel, N. (2007). *La Fé del Recuerdo: Laberintos Marranos*. Buenos Aires: Fondo de Cultura Económica.

Weber, M. (1905). *The Protestant Ethic and the Spirit of Capitalism*. London: Unwin, 1956.

3.10 Tables

Table 3.1: Inquisition and Non-Inquisition Sources (Comparison Sample)

Source	Region	Birth year min.	Birth year max.	N (cases)
Census/Mort.	Latin America	1502	1803	24800
Inquisition	Latin America	1531	1778	284
Census/Mort.	Portugal	1549	1799	4227
Inquisition	Portugal	1469	1778	10315
Census/Mort.	Spain	1540	1799	18300
Inquisition	Spain	1454	1795	2811

Table 3.2: Data Representativeness of Regional Units in Spain and Portugal

SPAIN				
Regions[31]	census of Floridablanca 1787[32]		Inquisition sources and comparison sample (1540-1750)	
	Total population	Percentage	Observations	Percentage
Aragon and Navarre	2,723,708	26.10	2,353	14.98
Center-Castille	3,272,591	31.36	4,358	27.74
North-Castille	2,164,402	20.74	2,380	15.15
South-Castille and Granada	2,106,810	20.19	5,752	36.61
Canary Islands	167,224	1.60	869	5.53
Total Spain	10,434,735	100	15,712	100
PORTUGAL				
Regions[33]	census of 1864[34]		Inquisition and comparison sample (1540-1750)	
	Total population	Percentage	Observations	Percentage
North	1,795,222	42.88	3,480	30.00
Center	1,532,099	36.60	4,547	39.20
South	500,193	11.95	3,415	29.44
Islands	358,792	8.57	158	1.36
Total Portugal	4,186,306	100.00	28,927	100.00

[31] The regional categorization for Spain was made along with the historical distribution that included the kingdoms of Aragón, Navarre, Castille and Granada. The regional units are composed of the following provinces: Centre-Castille: Albacete, Ciudad Real, Toledo, Cuenca, Guadalajara, Madrid, Badajoz, Cáceres, Salamanca, Avila, Segovia, Soria, Burgos, Valladolid, Zamora, León, Palencia, la Rioja. North-Castille: La Coruna, Pontevedra, Ourense, Lugo, Asturias, Cantabria, Vizcaya, Guipúzcoa, Álava. South-Castille and Granada: Huelva, Sevilla, Cadiz, Córdoba, Jaén, Murcia, Granada, Almeria, Malaga. Aragon and Navarre: Alicante, Valencia, Castellon, Tarragona, Lleida, Barcelona, Girona, Teruel, Zaragoza, Huesca, Navarra.

[32] Source: Instituto Nacional de Estadísticas de Espana (http://www.ine.es).

[33] The regions of Portugal are composed of the following provinces: North includes Braga, Bragansa, Vila Real, Vianna do Castello, Porto, Aveiro, Coimbra; Center includes Viseu, Guarda, Castelo Branco, Leiria, Santarem, Lisboa; South includes Setubal, Portalegre, Beja, Evora, Faro and Islands includes the Acores and Madeira.

[34] Source: Instituto Nacional de Estadistica do Portugal (http://www.ine.pt).

Table 3.3: Classification of "Crimes" Following Contreras and Henningsen (1986)

Crime	Frequency	Percent
Judaism	7,869	58.12
Blasphemy, propositions	989	7.30
Witchcraft and superstition	805	5.95
Polygamy	636	4.70
Detractors of the Holy Office (making offenses against the court or impeding its right functioning)	522	3.86
Mahometism	489	3.61
Clergy crimes (solicitation at confession and false clergymen)	372	2.75
Sexual crimes (sodomy, immoral statements, infidelity, cohabitation)	292	2.16
Protestantism, Calvinism, Illuminism, Freemasonry	275	2.03
Miscellaneous	96	0.71
Accomplice	29	0.21
Crime unknown	1,165	8.60
Total	13,539	100

Table 3.4: Determinants of Numeracy in Logit Regressions

	(1)	(2)	(3)	(4)	(5)
Samples included	Inq. And comparison	Only Inquisition	Inq. and comparison	Inq. and comparison	Inq. and comparison
Female	-6.83***	-18.24***	-6.34***	-10.79***	-6.00***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age 23-32	12.16***	14.61***	12.23***	13.42***	12.13***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Migrant				6.66***	
				(0.000)	
Judaism	10.89***	4.68***	13.92***	14.91***	19.71***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Elite crimes		14.09***	25.84***	24.40***	
		(0.000)	(0.000)	(0.000)	
Crime unknown		1.13	11.98***	18.93***	18.00***
		(0.616)	(0.000)	(0.000)	(0.000)
Mahometism					16.54***
					(0.000)
Protestantism					28.62***
					(0.000)
Polygamy					8.28***
					(0.002)
Blasphemy, propositions					21.73***
					(0.000)
Solicitation and false priests					32.46***
					(0.000)
Accomplice					24.76**
					(0.020)
Miscellaneous					25.16***
					(0.000)
Against Holy Office					28.53***
					(0.000)
Sexual Crimes					15.24***
					(0.000)
Witchcraft, Superstition					11.39***
					(0.000)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	59,193	13,394	59,193	26,558	59,193
Pseudo R-sq.	0.0194	0.0520	0.0216	0.0365	0.0238

Notes: Dependent Variable: numeracy (the inverse of “age statement is a multiple of five”). We controlled for time effects including half centuries of birth. All coefficients are multiplied by 125 (see Appendix B). Reference group in Model 1 is a man aged 33 to 72, born 1750 to 1800, not accused of Judaism. In Model 2 the reference group is a male defendant of an Inquisition court, aged 33 to 72, born 1750 to 1800, accused of one of the following crimes: witchcraft or superstition, sexual crimes, accomplice of a crime, blasphemy or propositions, polygamy or miscellaneous. In Model 3 the reference group is man of age 33-72, born 1750 to 1800, either of the control group or accused at an inquisition court for one of the following crimes: witchcraft or superstition, sexual crimes, accomplice of a crime, blasphemy or propositions, polygamy or miscellaneous. In Model 4 the reference group is the same as in Model 3, but being a non-migrant individual. In Model 5, the reference group is a man aged 33-72, born 1750 to 1800 and not accused at an inquisition court.

We also controlled for the sample source being a death register, as their ages were usually reported by relatives (and were sometimes estimated by priests) which would imply a downward bias with respect to census samples. “Protestantism” includes here Lutheranism, Calvinism, Illuminism, Freemasonry, because the latter two are represented by very small sample size.

Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table 3.5: Determinants of Literacy in Logit Regressions: The Canary Islands Sample

	(1)	(2)	(3)
Female	-0.53***	-0.45***	-0.43***
	(0.000)	(0.000)	(0.000)
Black and Mulatto		-0.23***	-0.21***
		(0.001)	(0.009)
Migrant			0.02
			(0.753)
Judaism	0.52***	0.56***	0.56***
	(0.000)	(0.000)	(0.000)
Elite crimes		0.50***	0.53***
		(0.000)	(0.000)
Time fixed effects	Yes	Yes	Yes
Observations	718	718	668
Pseudo R2	0.237	0.362	0.368

Notes: Dependent Variable: Literacy. The sample includes only defendants of the Inquisition court in the Canary Islands. We also controlled for time effects including half centuries of birth.
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.10

3.11 Figures

Figure 3.1: Deviation in Numeracy from Comparison Sample: Coefficients of Logit Regressions in Table 3.5 (Model 5)

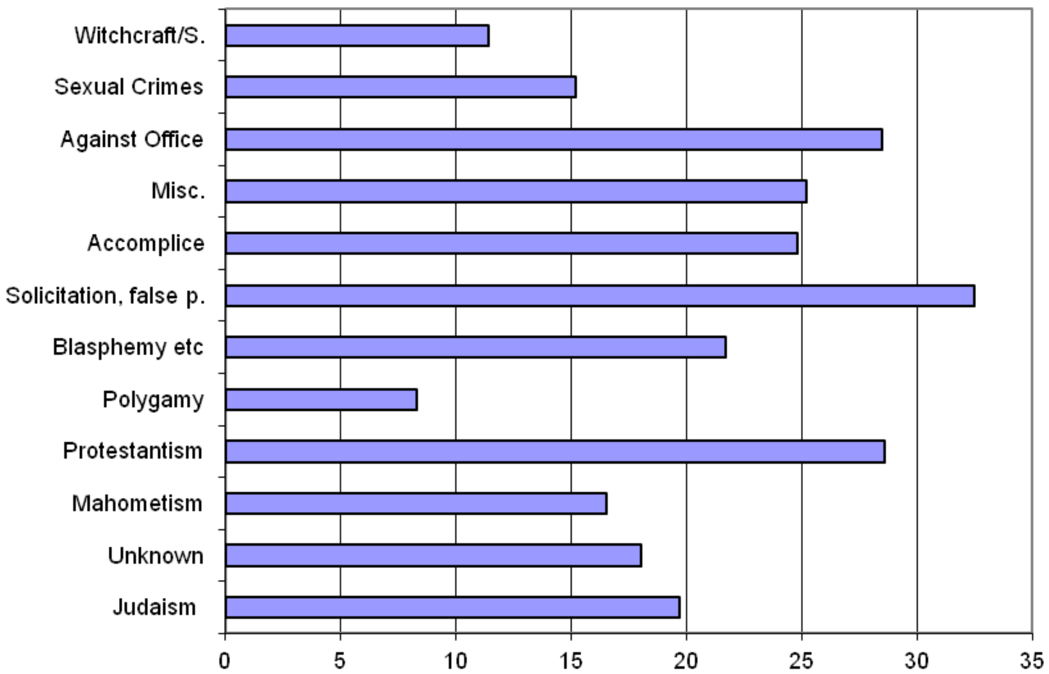
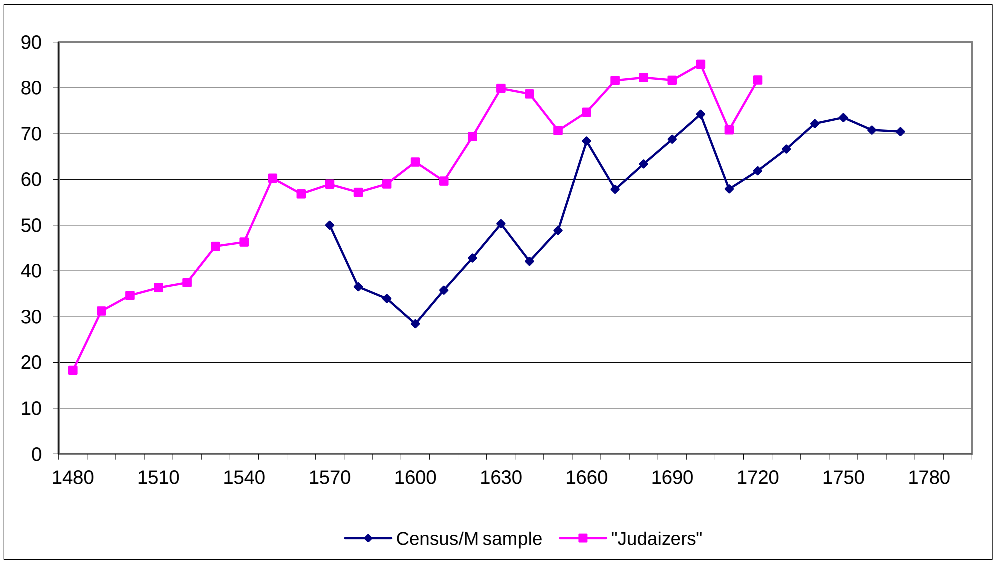


Figure 3.2: Numeracy of “Judaizers” and the Comparison Sample in the Iberian Peninsula



Note: Included are birth decades with 30 or more observations.

3.12 Appendix

Table A.1: Sources for Inquisition Data

Inquisition Tribunal	**Source**
Canary Islands (1521 – 1819)	Fajardo Spinola, F. (2005). Las Víctimas de la Inquisición en las Islas Canarias. La Laguna, Universidad de La Laguna
Cuenca (1521 – 1723)	National Historical Archive, Madrid (data provided by Sarah Nalle)
Galicia (1641)	Contreras, J. (1982). El Santo Oficio de la Inquisición de Galicia. Madrid: Akal Universitaria
Granada (1577–1614)	Garcia Fuentes, J. M. (2006). Visitas de la Inquisición al Reino de Granada. Granada: Editorial Universidad De Granada
Llerena (1695 – 1730)	Villanueva, J. P. and Bonet, B. E. (1984). Historia de la Inquisición en España y América, tomo III. Madrid: Biblioteca de Autores Cristianos
Logroño (1609 – 1633)	Henningsen, G. (2004): The Salazar Documents, Inquisitor Alonso de Salazar Frías and Others on the Basque Witch Persecution. Boston: Brill Leiden
Toledo (1577-1610)	Sierra, J. (2006). Procesos de la Inquisición de Toledo (1575-1610). Manuscrito de Halle. Madrid: Ed. Trotta.
Lima (1625 – 1761)	Palma, R. (1863). Anales de la Inquisición en Lima: Estudio Historico. Lima: Ediciones del Congreso de la Rebública
Chile, tribunal of Lima (1578-1785)	Toribio Medina, J. (1952). Historia del Tribunal del Santo Oficio de la Inquisición en Chile [1890]. Santiago de Chile: Impr. Ercilla
Mexico (1574-1804)	Medina, J. T. (1952). Historia del Tribunal del Santo Oficio de la Inquisición en México. Santiago de Chile: Impr. Ercilla
Tabasco, tribunal of Mexico (1567-1811)	Rico Medina, S. (1990). Los Predicamientos de la Fe, La Inquisición en Tabasco, 1567 – 1811. Gobierno del Estado de Tabasco
Évora, Coimbra, Lisboa (1487 – 1894)	National Archive Torre do Tombo, Lisbon [http://antt.dgarq.gov.pt/]

Table A.2: Occupations of Jews (23 to 62 years old)

Occupation	Freq.	Percentage
Merchant	898	29.24
Manufacture*	566	18.43
Tailor/ weaver	174	5.67
Farming/ fishery	137	4.46
Shopkeeper/ salesman	137	4.46
Lawyer/ judge/ investigator/ notary	127	4.13
Medicine/surgery/physician	114	3.61
Military	111	3.61
Domestic service	108	3.52
Lesser/tenant/owner	94	3.06
Subsistent farmer	79	2.57
No job	71	2.31
Clergy	52	1.69
Clerk	37	1.2
Donkey driver	37	1.2
Administrator	36	1.17
Pharmacy	36	1.17
University student	26	0.85
Barber	23	0.75
Nun	19	0.62
Arts/music	18	0.59
Seaman	18	0.59
University graduate	18	0.59
Miner	12	0.39
Teacher	12	0.39
Tax officer	8	0.26
Miller	6	0.2
Assistant	5	0.16
Butcher	5	0.16
Architect/ ship builder	4	0.14
Banker	4	0.13
Guard	4	0.13
Hunter	4	0.13
Politics	4	0.13
Scholar	4	0.13
Midwife	2	0.07
Beggar	1	0.03
Currency exchanger	1	0.03
Day laborer	1	0.03
Governor	1	0.03
Book printer	1	0.03
Insurance agent	1	0.03
Police	1	0.03
Rabbi	1	0.03
"Worker"	53	1.73
Total	3,071	100

*(blacksmith, carpenter, tanner, baker, etc.)

Table A.3: Number of Trials per "Crime" Category and Half Century of Birth

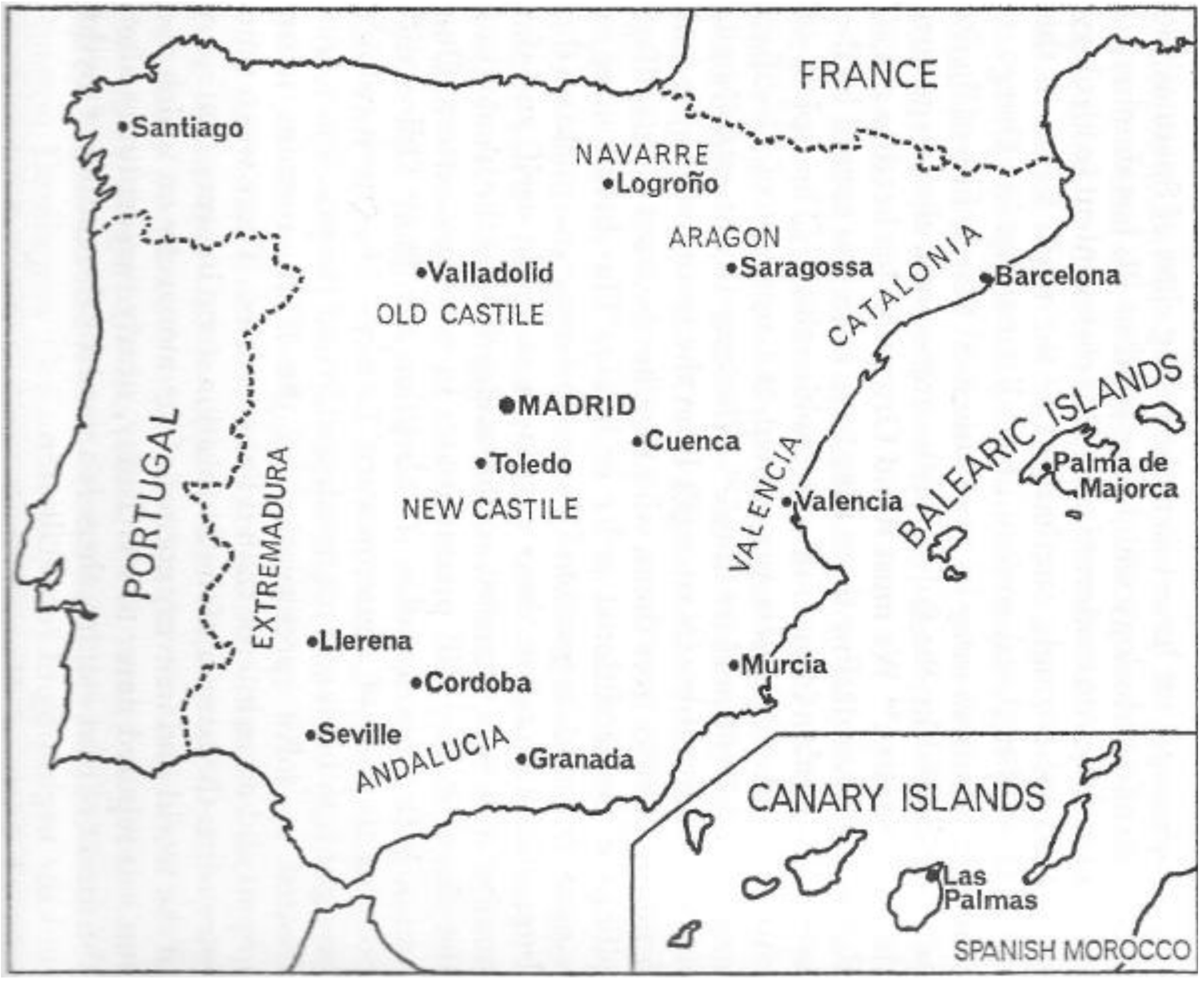
Half century of birth	1500	1550	1600	1650	1700	1750	Total
Polygamy	95	162	137	118	87	28	627
Blasphemy, Propositions	309	367	107	79	60	27	949
Clergy crimes	47	71	58	109	75	9	369
Accomplice	2	1	12	6	7	1	29
Unknown	167	479	174	209	27	4	1060
Mahometism	158	185	71	33	18		465
Judaism	1587	2349	2082	1263	322	1	7604
Misc.	30	29	11	9	11	5	95
Protestantism	90	97	15	27	24	20	273
Detractors of the Holy Office	95	202	95	56	43	9	500
Sexual crimes	62	153	54	14	6		289
Witchcraft and superstition	107	270	132	173	100	10	792
Total	2749	4365	2948	2096	780	114	13052

Table A.4: Numeracy in India and other World Regions 1450-1800

Regression Modell	(1) Logit	(2) Logit	(3) Linear prob. model
Female	-5.89***	-5.85***	-5.70***
	(0.00)	(0.00)	(0.00)
Age 23-33	13.13***	12.53***	12.27***
	(0.00)	(0.00)	(0.00)
Inquisition source	14.03***	11.51***	11.24***
	(0.00)	(0.00)	(0.00)
Death register		-5.57***	-5.26***
		(0.00)	(0.00)
India 1450-1749	-40.47***	-39.87***	-40.82***
	(0.00)	(0.00)	(0.00)
Northafrica and Middle East 1450-1539	-48.73***	-48.44***	-53.49***
	(0.00)	(0.00)	(0.00)
Northafrica and Middle East 1540-1750	-15.79*	-15.00*	-14.01*
	(0.02)	(0.07)	(0.09)
Subsaharan Africa 1450-1749	-45.79***	-45.83***	-50.26***
	(0.00)	(0.00)	(0.00)
Subsaharan Africa 1750-1799	-37.45*	-36.35*	-36.91*
	(0.07)	(0.08)	(0.09)
Ireland 1540-1709	44.02***	44.37***	33.32***
	(0.00)	(0.00)	(0.00)
Ireland 1710-1799			44.79***
			(0.00)
Brasil 1540-1599	-10.26	-9.19	-8.82
	(0.59)	(0.63)	(0.64)
Brasil 1600-1699	-7.25	-6.35	-6.01
	(0.16)	(0.22)	(0.24)
Brasil 1700-1799	3.40	4.29	3.91
	(0.64)	(0.55)	(0.55)
Spain 1450-1599	-19.39***	-18.54***	-18.45***
	(0.00)	(0.00)	(0.00)
Spain 1600-1699	Ref. Cat.	Ref. Cat.	Ref. Cat.
Spain 1700-1799	10.89***	12.61***	12.35***
	(0.00)	(0.00)	(0.00)
Portugal 1450-1599	-24.42***	-23.65***	-23.83***
	(0.00)	(0.00)	(0.00)
Portugal 1600-1699	-7.70***	-6.63***	-6.30***
	(0.00)	(0.00)	(0.00)
Portugal 1700-1799	0.34	3.48**	3.18*
	(0.84)	(0.04)	(0.06)
Southeast Europe 1510-1759	12.15	12.69	9.82
	(0.40)	(0.38)	(0.40)
Rest of the World 1450-1799	-6.06***	-7.65***	-7.44***
	(0.00)	(0.00)	(0.00)
Constant			63.27***
			(0.00)
Observations	59,186	59,186	59,193
Pseudo R2	0.0199	0.0203	0.03

Notes: Dependent Variable: numeracy (the inverse of “age statement is a multiple of five”). All coefficients are multiplied by 125. The reference category is a male individual, aged 33 to 72, born in Spain between 1600 and 1699 and not drawn from inquisitorial record. Rest of the World includes Latin American countries, Western Europe, Southeast Asia and East Asia.
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Figure A.1: Inquisition Courts in Spain



Source: Kamen (1965)

Appendix B. Notes on Numeracy Estimations

Assume that $\frac{1}{m}$ of the population are numerate and that age is uniformly distributed.

$1-\frac{1}{m}=\frac{m-1}{m}$ is not numerate and will state a multiple of five as their age anyway.

$\frac{1}{5}\cdot\frac{1}{m}$ of the population will correctly and non-accidently report a multiple of five.

In total, $\frac{1}{5m}+\frac{m-1}{m}=\frac{1-5m-5}{5m}=\frac{5m-4}{5m}=1-\frac{4}{5m}$ will claim to be a multiple of five years old.

Conversely, $1-\left(1-\frac{4}{5m}\right)=\frac{4}{5m}$ will answer with an age that is not a multiple of five.

The fraction of the population assumed to be numerate is recovered by multiplication with $\frac{5}{4}$, since $\frac{4}{5m}\cdot\frac{5}{4}=\frac{1}{m}$.

4. Jewish Educational Selection in 20th century Latin America

Abstract:

Chiswick and Lehrer among other economist have argued that the religion with which individuals are brought up has an influence on their educational achievements and their later earnings. The Jewish is one religious group that has been found to perform particularly well in terms of educational investments. An earlier study by Juif and Baten (2013) has confirmed this finding for the Iberian world in the era of the Spanish Inquisition. This study complements the previous one by analyzing the performance of Jews in early 20th century Latin America. After controlling for social, demographic and geographic variables, Jews perform significantly better in terms of human capital than Catholics and other religious minorities.

4.1 Introduction

A number of scholars have argued that the religion with which individuals are brought up has an influence on their educational attainment. The channels through which this occurs are various ranging from fertility choices and employment of mothers to religious laws and traditions that promote education. Judaism has been shown empirically to be one of the religious groups with the highest educational advantage in the United States and Canada. Chiswick (1983, 1988) and Lehrer (1999) find a higher level of human capital for U.S. Jews relative to non-Jews, and Tomes' (1983) study yields similar results for Canadian Jews. They suggest that Jews have a higher preference for the quality than the quantity of children, as fertility and the rate of young mothers' employment among Jews are relatively low in their samples. Moreover, these authors find that this religious group obtains higher returns on educational investments than others. Chiswick (1983, 1988) and Lehrer (1999) also bring the "diaspora" hypothesis into focus, which holds that, due to their history of prosecution, Jews favored investments into human capital over physical capital, because it is portable and more easily transferable. Those are some alleged reasons for Jews' relatively high investments into education.

Another explanation for the historically driven human capital advantage of Jews has been provided by Botticini and Eckstein (2007, 2012). They argue that their educational superiority derives from a religious law issued 100 B.C. which imposed fathers to teach their male children to read the bible. This literacy requirement which was originally not for professional use yielded Jews a comparative advantage to enter high skilled occupations later, when urbanization took place. The relatively good representation of Jews in high-skilled urban occupations persisted since the early Middle Ages.

Juif and Baten (2013) report a significant human capital advantage of Jews over the average population in the Iberian Peninsula and Latin America from the 1450s until the 19th century. In Spain, Sephardic Jews have a long tradition of holding important positions since the time of the Moorish Caliphate. A significant part of the Jews expelled in the 16th century secretly left for the Spanish colonies in America taking their human capital with them. This research note is an extension of Juif and Baten (2013) in which I analyze the human capital advantage of the Jews, many of Iberian origin, in late 19th and early 20th century Latin America. Since it focuses on a relatively recent period, I use measures of human capital that can capture more advanced skills than the age heaping technique. To my knowledge, no such study has been carried out for Latin America. I use individual census data from the Integrated Public Use Microdata Series (Ipums) that includes information on the religious affiliation for three Latin American countries. The human capital measures used are years of schooling, the highest school degree acquired and literacy. Holding social variables that are commonly argued to influence education constant, I find a significant human capital advantage of Jews over the average population.

This study contributes to a growing literature that has recently studied the effect of religion on human capital (Chiswick 1988, 1991; Lehrer 1999; Becker and Woessman 2009; Botticini and Eckstein 2007) and other economic outcomes (Barro and McCleary 2006; Guiso, Sapienza, and Zingales 2003, 2006). Section 2 summarizes this literature and discusses briefly the different hypotheses on Jews' (cultural) motivations for investing immoderately into education. Section 3 presents the data and method used in this research note, section 4 portrays the regression results and section 5 concludes.

4.2 Literature Review

4.2.1 Religion and Economy

At the macroeconomic level, a broader series of cross-country empirical studies including McCleary and Barro (2006), Glaeser and Sacerdote (2008) and Guiso, Sapienza and Zingales (2003) have recently analyzed the relation between religious affiliation and economic outcomes. McCleary and Barro (2006) report that strong religious believing, especially in afterlife, influences GDP growth positively. They explain this finding with Weber's hypothesis[1] that religion fosters personal traits such as work ethic, honesty (trust), thriftiness, charity, hospitality to strangers etc. and thereby affects economic outcomes. Jews perform slightly better than Christians in their cross-country analysis. Glaeser and Sacerdote (2008) find a positive relationship between education and the extent of religiosity (or active religious participation), which they attribute to the fact that more educated people have a greater incentive to participate in group activities, including church services. Thus, in their view, the direction of causality runs from educational level to religiosity. However, these authors find that human capital levels are higher for religious doctrines of low attendance, such as the Jewish. Guiso, Sapienza and Zingales (2003, 2006), in a similar vein as McCleary and Barro (2006), argue that religion as a cultural trait determines people's educational choice as well as values or preferences that in turn affect national economic outcomes. Individuals' attitudes towards trust, saving and certain political preferences, for example, are to a large extent a consequence of their religious upbringing and, even more so, of whether they still practice a religion. These authors report that the effect of religiosity on those values that foster economic growth is positive and, in a rank of religions, Christian religions are on average more positively associated with them than non-religiosity. However, in their study the effect of being raised as a Jew on growth enhancing values is not significant.

A number of empirical studies have focused on the relationship between Jewish affiliation and economic outcomes such as earnings, female employment or education during the twentieth century. Chiswick (1983, 1988) and Lehrer (1999) find a higher level of earnings and education as well as higher returns to human capital investments for

[1] Weber (1912) wrote in his work "The Protestant Spirit of Capitalism" that Protestant regions were more advanced economically due to the religious values characteristic for this confession.

American Jews relative to non-Jews[2], and Tomes' (1983) study shows a similar pattern for Canadian Jews. Chiswick (1983, 1988) and Lehrer (1999) attribute the higher demand of schooling by Jewish Americans to the "diaspora" hypothesis, which holds that, due to their history of frequent expels, Jews favored investments into human capital rather than physical capital, because it is portable and more easily transferable. Moreover, they suggest Jews have a higher preference for "quality" rather than quantity of children, as fertility and the rate of young mothers' employment among Jews are relatively low. Tomes' view is that Jews invest more into schooling due to the fact that for some reason they obtain higher returns to educational investments.

4.2.2 The History of Jews and Human Capital

Jews have historically been engaged in relatively high-skilled occupations such as trade, finance, medicine and governmental administration. Especially in the Iberian Peninsula in the Middle Ages, during the Muslim Caliphate, Jews enjoyed tolerance and good economic opportunities. When the kingdoms of Castile and Aragon were founded after the *Reconquista*, they continued to hold important posts, for example as royal counselors and administrators or royal physicians. However, after the implementation of the Inquisition court in Spain in the 15th century, their situation changed dramatically and Sephardic Jews who converted to Christianity after the edict of expulsion had to fear being accused of secretly practicing Jewish rites. This is why between the 15th and the 17th century large numbers of Jews and New Christians fleeing the harsh reprisals of the Inquisition took their way to North Africa, Northern and Eastern Europe and also to Latin America. One of the largest Jewish communities of South America was established in Brazil, especially

[2] Jews being defined as second-generation Americans raised in a home in which Yiddish, Hebrew, or Ladino was spoken either in addition to or instead of English

while it was a Dutch colony, and it consisted to a large extent of Sephardic Jews originally from Portugal and Spain, some of them having first found refugee in Amsterdam. In the British, Dutch and French colonies Jews could profess their religion openly, whereas in Spanish and Portuguese America they had to live as crypto-Jews. However, a great share opted for the latter regions, probably due to the cultural proximity with the Iberian Peninsula (Bernardini and Fiering 2001). In the last decade of the 19th century, Russian Jews fleeing the Tsarist prosecution landed in Latin America, many of them in Brazil. The highest numbers of Jews arrived in Brazil in the 1920s and 1930s, especially since Jewish rights in Eastern Europe were harshly restricted after the Polish annexation and other countries like the U.S., Argentina and Canada started restricting immigration. Moreover, the “holocaust” pushed a large group of Jews from Germany, Austria and other central European countries to leave for the Americas (Lesser 1995, Bernardini and Fiering 2001). Most of the Jews continued to live in cities and to be engaged in urban occupations after leaving Europe.

Apart from Kuznets’ “diaspora” hypothesis of Jews developing a preference to invest into education rather than into physical capital, different arguments have been raised on why Jews performed relatively well in terms of human capital and professional success. One view, introduced by Weber and also followed up by Kuznets, is that Diaspora Jews, such as other minorities, specialized in certain occupations in every city to maintain their cultural identity and follow their rites. Some historians, like Cecil Roth (1938), argued that restrictions on owning land imposed on Jews by the authorities made European Jews discard farming and specialize in urban occupations like crafts and trade. Yet Botticini and Eckstein (2007, 2012) reject the theory of restrictions and argue that at the time Jews entered high-skilled occupations in the Muslim Empire, during the eighth and ninth century, no such restrictions existed. These authors propose a human capital interpretation of the Jewish selection into certain high-skilled urban professions. In their view, the specialization of Jews in these occupations was exogenously driven by a religious law issued 100 B.C. that imposed fathers to teach their male children to read the Bible. This imposition of literacy provided Jews with an educational advantage to enter professions requiring high skills in the Middle Ages, when urbanization took place. Their argument is that their human capital investments were definitely an exogenous decision because Jews did not make any professional use of their literacy during more than seven centuries in which they were mainly engaged in farming (see also Juif and Baten (2013) on this passage).

4.3 Data Description and Method

I use census data from 1960 Brazil, 1970 Mexico and 1970 Chile from the Integrated Public Use Micro Sample International of the Minnesota Population Center (Ipums International) to assess the human capital advantage of Jews with respect to the whole population. All three datasets are representative samples of the national population. The Jewish share is less than one percent in all three censuses. Unsurprisingly, the Catholic share of the population represents between 80% in Chile and 97% in Mexico (Table 4.1). The summary statistics by country (Table 4.2 to Table 4.4) show that Jews perform better than any other religion in all educational variables. While in Brazil the average schooling is 2.1 years, Jews have on average attended school for 7.1 years. Literacy rates are 59% on average and 93% for the Jewish population. Moreover, while 22% of individuals included in the sample have obtained a secondary school degree, 53% of Jews have one (Table 4.2). Similar results were found for Chile and Mexico (Tables 4.3 and 4.4). In Chile, the mean years of schooling are 5.9 years for the whole population versus 10.6 years for Jews, whereas literacy rates are 88% and 98% respectively. While only 15% of individuals included in the Chilean sample have obtained a secondary school degree, 54% of Jews have (Table 4.3). In the Mexican sample, the average person has attended school for 3.3

years, whereas Jews went to school for a mean of 7.4 years. While the literacy rate is 73% in Mexico, 93% of Jews living in Mexico can read and write. The share of Jews having obtained a secondary school degree is 19% relative to a national average of 4% (Table 4.4). Moreover, the Jewish women have fewer children on average than Catholic women in all samples, suggesting that Jews preferred having less children and investing more into their education.

In order to test the effect of being raised with a particular religion on educational outcomes, I run regressions with different proxies of education as dependent variables: years of schooling, being literate and having obtained a secondary school degree. In the same vein as Glaeser and Sacerdote (2008) and Lehrer (1999), I perform ordinary least squares regressions (OLS) on religion when the dependent variable is "years of schooling" and logistic regressions when the dependent variable is a dummy variable that takes the value of 0 or 1, either for "literacy" or for "completed secondary track".[3] I control for other social and demographic factors that usually affect educational performance, such as gender, being married, years of schooling of the mother, number of children, persons living in the household, living in a rural region, being a migrant from a country with high or low educational levels and being indigenous. Married men have been observed to earn more than unmarried men, maybe because they have a higher ability that makes them successful also on the marriage market and because they show more professional commitment (Tomes 1983). Marriage could also have a negative influence on schooling for women, since they might quit studying after getting married. Being indigenous is a variable only available for Mexico and it is approximated here by the information that an individual speaks an indigenous language. The variable "yrschl mother" (for the mother's years of

[3] In order to test the influence of religion on attainment (ranging from 0 to 9) one could also use ordered logit regressions, which are multinomial methods for ordered variables without constant intervals between the attributes.

schooling) is not included in the regressions because it would cause many observations to drop, since this information is only available for the individuals living in the same household as their parents. However, the summary statistics show us that Jewish mothers are better educated than mothers of other religious affiliates (Tables 4.2 to 4.4). The Brazilian sample includes information on the ethnic background, which I use in the regressions to control for the effect of racial discrimination.

One should of course bear in mind that a substantial share of the Jews living in South America in the 1960s and 1970s were European immigrants. A large part of them had fled the Nazi regime before or during World War II. Those born in Northwestern and Central Europe in fact account for 21% of all the Jews in Chile, for 13% in Brazil and for 5% in Mexico. They could have a substantial educational advantage only because they were born and received an education in countries which have higher average educational levels. Another relatively large migrant group comes from Russia, since Jews were expelled at the beginning of the 20th century after the issuance of anti-Jewish pogroms and some of them ended in Latin America (Spitzer 2013). Those Jews, on the other hand, are on average less educated than non-migrant Jews. It is thus important to control for migrant status in the regressions when assessing the human capital advantage of Jews. I opted for including two migration variables, one for migrants from countries with high levels of education ("mig highed"), in which I include Germany, Czech Republic, Hungary, Denmark, Finland, Ireland, Norway, Sweden, United Kingdom, Austria, Belgium, France, the Netherlands and Switzerland, and another variable for migrants from the rest of countries ("mig lowed").

The age range included in the analyses is 26 to 90 years for those regressions in which "years of schooling" is the dependent variable, since before 26 the schooling track might not be completed yet. For those regressions with "literacy" and "completed secondary track" as dependent variables, 18 to 90 year-olds are included. In the case of Mexico, those born before the decade of the Revolution (1910) were excluded for the regressions, because educational conditions especially for the poorer section of the population took a dramatic turn afterwards. Regressions are performed first for males and females together and then separately by gender, because results might differ between the sexes.

4.4 Regression Results

The regression results in Tables 4.5 to 4.13 confirm that Jews have a substantial advantage in terms of education over the average population. For Brazil, Table 4.5 column 1 shows the results of the baseline model. Here, the Jewish advantage over Catholics in years of schooling is 5.2 in total. The gap is more pronounced for men than for women (5.5 versus 4.7 years) as shown in columns 2 and 5. However, when controlling for social and geographic variables, the difference in schooling between Jews and Catholics decreases significantly (to 3.9 years for men and 3.5 years for women). The most decisive factor in determining the amount of schooling besides religious upbringing seems to be whether the individual lives in an urban region. Controlling for all other possible determinants of schooling, living in an urban region increases the average years of education by 1.5 years for men and 2 years for women. Moreover, Jewish women are 38% more likely to be able to read and write and Jewish men are 31% more likely (Table 4.6, columns 2 and 4). Being Jewish increases the probability of obtaining a secondary school degree by 35% for men and 30% for women (Table 4.7, columns 2 and 4). However, the coefficient for secondary schooling decreases by over 20 points when controlling for other determinants of education (Table 4.7, columns 3 and 5).

Chilean Jews attended school for on average 5 years longer than Catholics (Table 4.8 column 1); men 5.3 years and women 4.6 years (columns 2 and 5). When including all social variables in the OLS regressions, the coefficients for Jewish religion sink to 2.6 for men and 2.4 for women, but they remain significant (columns 4 and 7). The most important social variables here are those related to being a migrant – both from countries with high educational levels and from the rest of countries –with positive and significant coefficients of between 2.45 and 6.26. Also living in an urban region increases years of schooling by more than three years. Other religions also perform better than Catholic Christians, but the coefficients are significantly smaller than for the Jewish. The second most educated group is the Orthodox Christian (however, the coefficient turns insignificant for women when including other control variables); “no religion” and “other religion” follow. When looking at literacy (Table 4.9), being Jewish increases the probability to be able to read and write by 10%; 9% for men and 11% for women. The coefficients decline to 4% and 7% respectively for males and females when controlling for social variables (Table 4.9, columns 3 and 5). Table 4.10 displays the probability of having obtained a secondary school degree, which is between 15% (for women controlling for other variables) and 44% (for men without including control variables) higher for Jews than for Catholics.

In 1970 Mexico, Jews have on average attended school 4.8 years longer than Catholics; similar results account for males and females (Table 4.11). When controlling for social, demographic and geographic variables, Jewish men have attended school 2.5 years longer than Catholic men and Jewish women have attended school for 3.0 years longer than Catholic women (Table 4.11, columns 4 and 7). Interestingly, “no religion” has a negative and significant coefficient in all specifications, contrary to the case of Chile. When looking at literacy, after controlling for other variables, female Jews were 16% more literate than female Catholics and male Jews were 11% more literate than their Catholic counterparts (Table 4.12, columns 3 and 5). Moreover, the probability of having obtained a secondary school degree was 2% higher for Jewish women and 6% higher for Jewish men (Table 4.13, columns 3 and 5). Also in Mexico, migrants not only from countries of high average schooling levels have a significant educational advantage.

4.5 Conclusions and Further Research

This research note showed that the educational advantage of Jews that has been found in previous studies for the United States and for Canada, as well as for Latin America in much earlier times (by Juif and Baten 2013), also accounts for 20th century Latin America. It aimed at disentangling the effect of being raised with a certain religion (and culture) on investments into education by controlling for social, demographic and regional aspects that affect schooling outcomes. Of course, I cannot rule out the possibility that in some cases the causality was reverse and well educated individuals chose to become Jews (as argued by Glaeser and Sacerdote). However, intergenerational continuity in religious affiliation has been shown to be the norm and thus the current confession is a good proxy of the religious upbringing (Tomes 1983). Jews were shown to have fewer children than other religious groups and to invest more into their education. A number of explanations for Jewish immoderate investments into schooling were presented in the paper. Some of them are based on historically fostered cultural preferences; an alternative one is that for some reason Jews obtain higher returns to investments in education than non-Jews, as claimed by Chiswick (1983, 1988) and Tomes (1983). Given the possibilities that these datasets offer, it would be interesting to analyze returns to schooling investments in terms of earnings for Jews and other religions. In a similar vein as Chiswick (1991), one could also use a proxy of “occupational prestige” to analyze Jews’ relative performance on the labor market in Latin America.

4.6 References

Barro, R.J. and McCleary, R.M. (2006). Religion and Economy. *Journal of Economic Perspectives*, 20(2): 49-72.

Becker, S.O., and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, 124: 531-596.

Bernardini, P. and Fiering, N. (2001). *The Jews and the Expansion of Europe to the West 1540-1800*. New York: Berghahn Book.

Botticini, M., and Eckstein, Z. (2007). From Farmers to Merchants, Conversions and Diaspora: Human Capital and Jewish History. *Journal of the European Economic Association*, 5(5): 855-926.

Botticini, M., and Eckstein, Z. (2012). *The Chosen Few: How Education Shaped Jewish History, 70-1492*. Princeton University Press.

Chiswick, B.R. (1983). The Earnings and Human Capital of American Jews. *Journal of Human Resources*, 18(3): 313–336.

Chiswick, B.R. (1988). Differences in Education and Earnings across Racial and Ethnic Groups: Tastes, Discrimination, and Investments in Child Quality. *The Quarterly Journal of Economics*, 103(3): 571-597.

Chiswick, B. R. (1991). Jewish Immigrant Skill and Occupational Attainment at the Turn of the Century. *Explorations in Economic History*, 28(1): 64–86.

Glaeser, E.L., and Sacerdote, B.I. (2008). Education and Religion. *Journal of Human Capital*, 2(2): 188-215.

Guiso, L., Sapienza, P., and Zingales, L. (2003). People's Opium? Religion and Economic Attitudes, *Journal of Monetary Economics*, 50(1): 225–28.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does Culture Affect Economic Outcomes?.*The Journal of Economic Perspectives*, *20*(2): 23-48.

Juif, D. and Baten, J. (2013). Dangerous Education? The Human Capital of Iberian and Latin American Jews and Other Minorities during the Inquisition. University of Tübingen Working Paper.

Kuznets, S. (1960). Economic Structure and Life of the Jews. In: *The Jews: Their History, Culture, and Religion*, edited by Louis Finkelstein. Jewish Publication Society of America.

Kuznets, S. (1972). Economic Structure of U.S. Jewry: Recent Trends. Working paper, Institute of Contemporary Jewry, Hebrew University of Jerusalem.

Lea, H.C. (1907). *A History of the Inquisition of Spain*. New York: The MacMillan Company.

Lehrer, E.L. (1999). Religion as a Determinant of Educational Attainment: An Economic Perspective. *Social Science Research*, 28(4): 358-379.

Lesser, J. (1995). *Welcoming the Undiesirables: Brazil and the Jewish Question*. Los Angeles: University of California Press.

Roth, C. (1938). *The Jewish Contribution to Civilization*. London: Macmillan & Co.

Spitzer, Y. (2013). Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914. Working paper presented at the Economic History Association Annual Meeting in Washington, 2013.

Tomes, N. (1983). Religion and the Rate of Return on Human Capital: Evidence from Canada. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 16(1): 122-138.

Weber, M. (1905). *The Protestant Ethic and the Spirit of Capitalism*. London: Unwin, 1956.

4.7 Tables

Table 4.1: Summary Statistics by Census (18 to 90 year old individuals)

	Mexico 1970		Brazil 1960		Chile 1970	
Religion	Total No.	Percentage of sample	Total No.	Percentage of sample	Total No.	Percentage of sample
Catholic	219,773	96.42	1,593,364	92.65	386,108	79.74
Protestant	4,737	2.08	68,373	3.98	28,465	5.88
No religion	2,282	1.00	10,730	0.62	11,292	2.33
Jewish	330	0.14	2,381	0.14	1,120	0.23
Other	797	0.35	9,054	0.53	1,018	0.21
Unknown			736	0.04	55,758	11.51
Buddhist			6,376	0.37	87	0.02
Muslim			214	0.01	111	0.02
Orthodox			1,719	0.1	274	0.06
Spiritist			26,850	1.56		
Total Obs. No.	227,927	100	1,719,797	100	484,233	100

Table 4.2: Means of Variables in Brazil 1960

Means of Variables	All	Catho-lics	Ortho-dox	Protes-tants	No reli-gion	Other	Jewish	Budd-hist	Mus-lim	men	women	rural	urban	Sao Paolo	Jews in Sao Paolo
yrschl	2.05	1.98	4.99	2.82	3.18	3.09	7.13	2.92	3.26	2.24	1.86	1.01	3.16	2.94	8.12
second	0.22	0.21	0.49	0.32	0.31	0.30	0.53	0.40	0.40	0.23	0.21	0.09	0.37	0.36	0.58
literacy	0.59	0.58	0.86	0.78	0.68	0.77	0.93	0.80	0.77	0.64	0.54	0.44	0.75	0.73	0.98
female	0.50	0.50	0.42	0.51	0.31	0.50	0.48	0.46	0.28			0.48	0.52	0.49	0.49
child. no.	3.79	3.80	2.70	3.83	2.95	3.48	2.16	4.35	3.64		3.79	4.34	3.25	3.27	1.78
persons in HH	5.82	5.84	4.72	5.65	4.97	5.54	4.59	7.46	5.14	5.81	5.83	6.14	5.47	5.31	4.10
urban	0.48	0.47	0.89	0.55	0.73	0.72	0.89	0.44	0.86	0.46	0.50			0.67	0.98
married	0.73	0.73	0.77	0.78	0.70	0.78	0.80	0.76	0.72	0.69	0.77	0.74	0.72	0.75	0.83
age	36.41	36.23	42.91	38.02	37.47	39.07	41.85	41.00	37.12	36.68	36.15	35.96	36.89	36.85	42.74
yrschl mother	1.11	1.07	1.87	1.72	1.52	1.49	3.80	1.22	1.65	0.99	1.29	0.54	1.77	1.39	4.41
migrant	0.03	0.03	0.76	0.05	0.08	0.04	0.57	0.70	0.69	0.04	0.03	0.01	0.06	0.10	0.72

Table 4.3: Means of Variables for Chile 1970

Means of Variables	All	Catho-lics	Ortho-dox	Protes-tants	No reli-gion	Other	Jewish	Budd-hist	Mus-lim	men	women	rural	urban	Metro-politan	Jews in Metro-politan
yrschl	5.91	5.87	8.13	5.20	7.59	7.82	10.66	6.06	6.46	3.31	6.68	6.09	5.74	6.94	11.57
second	0.15	0.14	0.32	0.10	0.31	0.31	0.54	0.13	0.21	0.03	0.18	0.16	0.13	0.20	0.62
literacy	0.88	0.88	0.89	0.86	0.90	0.90	0.98	0.93	0.85	0.73	0.93	0.89	0.87	0.94	0.99
female	0.52	0.52	0.47	0.55	0.35	0.46	0.50	0.48	0.52	0.46	0.54			0.55	0.51
child. no.	3.68	3.71	3.98	4.48	3.26	3.32	2.77	3.40	3.60	5.00	3.35	5.28	3.68	3.14	2.33
persons in HH	5.85	5.88	5.81	5.98	5.52	5.47	4.77	5.33	5.14	6.39	5.69	5.91	5.79	5.73	4.52
urban	0.77	0.76	0.92	0.80	0.87	0.83	0.93	0.78	0.89			0.74	0.80	0.94	0.99
married	0.70	0.71	0.81	0.77	0.71	0.74	0.84	0.74	0.79	0.68	0.70	0.68	0.72	0.71	0.83
age	39.14	39.15	47.30	40.18	37.96	40.14	47.24	42.97	41.79	40.09	38.86	38.89	39.37	38.66	47.24
yrschl mother	3.18	3.18	5.89	2.81	3.98	4.43	7.68	3.33	1.44	1.64	3.72	3.10	3.28	4.09	9.57
migrant	0.02	0.01	0.42	0.03	0.02	0.07	0.45	0.06	0.19	0.01	0.02	0.02	0.01	0.02	0.52

Table 4.4: Means of Variables for Mexico 1970

Means of Variables	All	Catho-lics	Protes-tants	No religion	Un-known	Jewish	men	women	rural	urban	D.F.	Jews in D.F.
yrschl	3.30	3.27	3.74	4.24	4.63	7.38	3.63	2.98	1.76	4.33	5.79	9.23
second	0.04	0.03	0.05	0.12	0.10	0.19	0.05	0.02	0.01	0.05	0.10	0.27
literacy	0.73	0.72	0.78	0.68	0.81	0.93	0.77	0.68	0.59	0.82	0.89	0.98
female	0.51	0.51	0.52	0.38	0.48	0.48			0.49	0.52	0.53	0.49
child. no.	4.05	4.05	4.20	3.85	3.93	3.70		4.05	4.55	3.74	3.27	3.19
persons in HH	7.38	7.39	7.16	6.99	6.33	6.38	7.93	6.85	6.77	7.79	6.81	5.93
urban	0.60	0.60	0.57	0.49	0.72	0.82	0.58	0.61			0.97	1.00
married	0.76	0.76	0.78	0.74	0.78	0.82	0.73	0.80	0.80	0.74	0.71	0.81
age	37.12	37.07	38.33	37.56	40.11	40.88	37.24	37.00	37.37	36.95	36.39	42.16
yrschl mother	2.24	2.22	2.46	2.41	2.64	5.50	2.03	2.54	1.10	2.90	3.96	7.79
indigenous	0.09	0.09	0.14	0.16	0.08	0.03	0.09	0.09	0.17	0.04	0.02	0.01
migrant	0.01	0.00	0.03	0.02	0.04	0.26	0.01	0.01	0.00	0.01	0.02	0.39

Table 4.5: Years of Schooling in Brazil 1960

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Included	all	males	males	males	females	females	females
Regr. Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS
persons in HH			-0.01***	-0.00*		-0.00*	0.05***
			(0.00)	(0.08)		(0.06)	(0.00)
urban			2.09***	2.01***		1.62***	1.50***
			(0.00)	(0.00)		(0.00)	(0.00)
mig highed			4.13***	4.04***		2.88***	2.69***
			(0.00)	(0.00)		(0.00)	(0.00)
mig lowed			0.29***	0.15***		-0.06***	-0.17***
			(0.00)	(0.00)		(0.00)	(0.00)
black			-1.49***	-1.39***		-1.43***	-1.35***
			(0.00)	(0.00)		(0.00)	(0.00)
brown			-1.24***	-0.98***		-1.13***	-0.91***
			(0.00)	(0.00)		(0.00)	(0.00)
married			-0.17***	-0.17***		-0.51***	-0.13***
			(0.00)	(0.00)		(0.00)	(0.00)
child. no.							-0.09***
							(0.00)
Religion:							
no religion	1.37***	1.59***	0.84***	0.78***	0.52***	0.11*	0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.06)	(0.75)
Buddhist	0.79***	0.91***	0.35***	0.32***	0.61***	0.48***	0.39***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Jewish	5.15***	5.53***	3.61***	3.54***	4.73***	3.43***	3.28***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Muslim	1.21***	1.42***	-0.03	-0.10	0.15	-0.76*	-0.76*
	(0.00)	(0.00)	(0.93)	(0.78)	(0.73)	(0.08)	(0.07)
Orthodox	3.05***	3.45***	1.87***	1.85***	2.43***	1.50***	1.39***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Protestant	0.84***	0.96***	0.50***	0.37***	0.73***	0.40***	0.28***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Spiritist	1.30***	1.54***	0.83***	0.69***	1.06***	0.51***	0.35***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Other	0.70***	0.90***	0.34***	0.20***	0.46***	0.11***	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.74)
Unknown	1.81***	2.05***	0.83*	0.68	1.52***	0.46	0.45
	(0.00)	(0.00)	(0.07)	(0.14)	(0.00)	(0.18)	(0.21)
Birth decades incl	no	no	yes	yes	no	yes	yes
Residence state incl.	no	no	no	yes	no	no	yes
Constant	1.83***	2.06***	1.89***	1.88***	1.60***	1.86***	1.68***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	1,073,853	542,103	541,566	541,566	531,750	531,282	519,360
R-squared	0.02	0.02	0.24	0.25	0.01	0.24	0.26

Notes: Dependent Variable: years of schooling. The age range included is 26 to 90 years. The reference category in all specifications is a Catholic individual. P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table 4.6: Literacy in Brazil 1960

	(1)	(2)	(3)	(4)	(5)
Included	all	males	males	females	females
Regr. Model	logit mfx	logit mfx	logit mfx	logit mfx	logit mfx
Jewish	0.35***	0.31***	0.22***	0.38***	0.29***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	1,527,092	762,604	761,904	764,488	750,852
Pseudo R2	0.00918	0.0116	0.157	0.00762	0.182

Notes: Dependent Variable: Literacy. The included age range is 18 to 90 years. The reference religion-category is Catholic. Birth decades are included in all models and other social variables included in Models 3 and 5 are married, urban, migrant, black, brown, persons living in the household number of children per woman (in Model 5) and state of residence in Brazil.
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table 4.7: Secondary Schooling in Brazil 1960

	(1)	(2)	(4)	(5)	(7)
Included	all	males	males	females	females
Regr. Model	logit mfx	logit mfx	logit mfx	logit mfx	logit mfx
Jewish	0.33***	0.35***	0.06***	0.30***	0.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	1,528,424	763,338	762,463	765,086	751,192
R-squared	0.00698	0.0087	0.189	0.0054	0.193

Notes: Dependent Variable: Secondary school track completed. The included age range is 18 to 90 years old. The reference religion-category is Catholic. Birth decades are included in all models and other social variables included in Models 3 and 5 are married, urban, migrant, black, brown, persons living in the household number of children per woman (in Model 5) and state of residence in Brazil.
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table 4.8: Years of Schooling in Chile 1970

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Included	all	males	males	males	females	females	females
Regr. Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS
persons in HH			-0.06***	-0.06***		-0.10***	-0.01***
			(0.00)	(0.00)		(0.00)	(0.00)
urban			3.37***	2.85***		3.01***	2.20***
			(0.00)	(0.00)		(0.00)	(0.00)
mig highed			6.26***	6.05***		5.30***	4.73***
			(0.00)	(0.00)		(0.00)	(0.00)
mig lowed			3.81***	3.59***		2.86***	2.45***
			(0.00)	(0.00)		(0.00)	(0.00)
married			0.51***	0.47***		-0.22***	0.36***
			(0.00)	(0.00)		(0.00)	(0.00)
child. no.							-0.23***
							(0.00)
Religion:							
no religion	1.89***	2.04***	1.47***	1.51***	1.38***	0.91***	0.94***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Buddhist	0.17	0.76	0.43	0.28	-0.39	-0.68	-0.74
	(0.74)	(0.37)	(0.57)	(0.70)	(0.52)	(0.23)	(0.18)
Jewish	4.97***	5.28***	2.40***	2.23***	4.64***	2.41***	2.12***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Muslim	0.47	0.95	-0.23	-0.17	-0.07	-1.16*	-1.16*
	(0.31)	(0.16)	(0.73)	(0.79)	(0.91)	(0.08)	(0.06)
Orthodox	2.79***	3.41***	1.21***	1.15***	2.03***	0.64	0.66
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.16)	(0.12)
Protestant	-0.69***	-0.55***	-0.75***	-0.60***	-0.80***	-0.91***	-0.62***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Other	2.37***	3.00***	2.35***	2.37***	1.53***	1.05***	1.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Unknown	0.15***	0.27***	0.27***	0.26***	0.04	-0.02	-0.03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.13)	(0.57)	(0.25)
birth decades incl.	no	no	yes	yes	no	yes	yes
residence region incl.	no	no	no	yes	no	no	yes
Constant	5.53***	5.71***	4.07***	5.04***	5.36***	4.71***	5.47***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	358666	171242	171156	171156	187424	187333	187171
R-squared	0.01	0.01	0.21	0.22	0.01	0.17	0.22

Notes: Dependent Variable: years of schooling. The age range included is 26 to 90 years. The reference category in all specifications is a Catholic individual.
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table 4.9: Literacy in Chile 1970

	(1)	(2)	(3)	(4)	(5)
Included	all	males	males	females	females
Regr. Model	Logit mfx	Logit mfx	Logit mfx	Logit mfx	Logit mfx
Jewish	0.10***	0.09***	0.05***	0.11***	0.07***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	484,233	231,424	231,295	252,809	252,497
Pseudo R2	0.000978	0.000958	0.146	0.00135	0.151

Notes: Dependent variable: Literacy. The age range included is 18 to 90 years. Reference religion-category is Catholic. Birth decades are included in all models and other social variables included in Models 3 and 5 are married, urban, migrant, persons living in the household, residence region and number of children per woman (only Model 5).
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.1

Table 4.10: Secondary Schooling in Chile 1970

	(1)	(2)	(3)	(4)	(5)
Included	all	males	males	females	females
Regr. Model	Logit mfx	Logit mfx	Logit mfx	Logit mfx	Logit mfx
Jewish	0.41***	0.44***	0.21***	0.37***	0.15***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	484,233	231,424	231,295	252,809	252,497
Pseudo R2	0.00890	0.0108	0.0949	0.00644	0.104

Notes: Dependent variable: Secondary school track completed. The age range included is 18 to 90 years. Reference religion-category is Catholic. Birth decades are included in all models and other social variables included in Models 3 and 5 are married, urban, migrant, persons living in the household, residence region and number of children per woman (only Model 5).
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.1

Table 4.11: Years of Schooling in Mexico 1970

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Included	all	males	males	males	females	females	females
Regr. Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS
persons in HH			-0.01***	-0.01***		-0.05***	-0.02***
			(0.00)	(0.00)		(0.00)	(0.00)
urban			2.68***	1.95***		2.00***	1.40***
			(0.00)	(0.00)		(0.00)	(0.00)
mig highed			7.49***	6.62***		7.17***	6.42***
			(0.00)	(0.00)		(0.00)	(0.00)
mig lowed			3.95***	3.47***		3.74***	3.28***
			(0.00)	(0.00)		(0.00)	(0.00)
indigenous			-1.10***	-0.93***		-1.48***	-1.28***
			(0.00)	(0.00)		(0.00)	(0.00)
married			-0.44***	-0.35***		-1.03***	-0.65***
			(0.00)	(0.00)		(0.00)	(0.00)
child. no.							-0.08***
							(0.00)
Religion:							
No religion	1.18***	1.54***	1.70***	1.49***	0.35	0.65***	0.53***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.11)	(0.00)	(0.00)
Jewish	4.80***	4.83***	3.39***	2.47***	4.73***	3.59***	3.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Protestant	0.53***	0.51***	0.62***	0.57***	0.57***	0.64***	0.62***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Other	1.39***	1.94***	1.48***	1.11***	0.78***	0.45*	0.31
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.08)	(0.21)
No religion	-3.12***	-3.46***	-2.86***	-2.56***	-2.79***	-3.62***	-3.44***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
birth decades incl.	no	no	yes	yes	no	yes	yes
residence state incl.	no	no	no	yes	no	no	yes
Constant	3.12***	3.46***	2.45***	2.29***	2.79***	2.92***	2.59***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	116,419	57,379	57,377	57,377	59,040	59,040	59,040
R-squared	0.00	0.01	0.15	0.21	0.00	0.16	0.22

Notes: Dependent Variable: years of schooling. The age range included is 26 to 50 years. The reference category in all specifications is a Catholic individual.
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.1

Table 4.12: Literacy in Mexico 1970

	(1)	(2)	(3)	(4)	(5)
Included	all	males	males	females	females
Regr. Model	Logit mfx	Logit mfx	Logit mfx	Logit mfx	Logit mfx
Jewish	0.18***	0.16***	0.11***	0.20***	0.16***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	184,055	89,767	89,767	94,288	94,287
Pseudo R2	0.000786	0.00117	0.120	0.000817	0.164

Notes: Dependent variable: Literacy. Reference religion-category is Catholic. The age range included is 26 to 50 years. Birth decades are included in all models and other social variables included in Models 3 and 5 are married, urban, migrant, persons living in the household, residence state in Mexico and number of children per woman (only Model 5). P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.1

Table 4.13: Secondary Schooling in Mexico

	(1)	(2)	(3)	(4)	(5)
Included	all	males	males	females	females
Regr. Model	Logit mfx	Logit mfx	Logit mfx	Logit mfx	Logit mfx
Jewish	0.17***	0.24***	0.06***	0.11***	0.02*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.07)
Observations	184,060	89,770	89,767	94,290	94,287
Pseudo R2	0.00701	0.00777	0.139	0.00513	0.0958

Notes: Dependent variable: Secondary school track completed. The age range included is 26 to 50 years. Reference religion-category is Catholic. Birth decades are included in all models and other social variables included in Models 3 and 5 are married, urban, migrant, persons living in the household, residence state in Mexico and number of children per woman (only Model 5).
Robust P-Values in parentheses: ***p<0.01, **p<0.05, *p<0.1

4.8 Appendix

Table A.1: Definitions of Variables

Variables	Definition
yrschl	Years of schooling
edattand	Educational attendance (ranging from 0 to 9)
second	Secondary school track completed
literacy	Literacy
female	female
child. no.	Number of children per woman
persons in HH	Number of persons living in the household
urban	Living in an urban region (takes the value of 0 or 1)
married	Being married
age	Individuals' age
yrschl mother	Years of schooling of the mother
migrant	Being a migrant
mig highed	Migrants from high education countries in northwestern and central Europe
mig lowed	Migrants from low education countries

5. Who left the "Fortunate Islands"? Skill Selectivity of Migrants to Latin America. Evidence from 19th Century Canary Islanders

Abstract:

One potential determinant of long-run human capital formation in a country is selective immigration. Chiswick argued that 20th century immigrants to the U.S. were positively selected and contributed with their human capital to economic growth. Whether this is also the case for Latin America during the 19th century is not clear yet. This paper aims to shed more light on the skill selectivity of migrants to Cuba, Uruguay and Argentina during the age of mass migration as well as on their comparative human capital advantage relative to natives in those destination countries. Using new micro data compiled from passenger lists and census data I analyze the numeracy of immigrants with the use of the age heaping technique. I find that the numeracy level of immigrants relative to natives in the destination countries varies across destinations. In a second step I quantify the human capital of emigrants from Spain and from the Canary Islands in particular and compare it to the origin region's population. The Canary Islands was one of the Spanish provinces with the highest emigration rates and emigration flows started earlier than in the rest of Spain. The selectivity with respect to the sending region is negative for emigrants from the Canary Islands, whereas emigrants from the Spanish peninsula to Argentina were negatively selected and no particular selection applied for Spanish emigrants to Cuba.

5.1 Introduction

Analyzing the skill selectivity of migrants is an interesting topic in economics which affects both origin and destination countries. On the one hand, the characteristics of those who leave can affect dramatically the destiny of those who stay behind. In today's developing countries, the out-migration of highly educated – the so called "brain drain" – deprives the home economy of important human capital resources, which may hinder future progress.[1] On the other hand, countries attracting higher skilled immigrants could improve the average human capital level and thus foster subsequent economic development, both directly and through its effect on the quality of institutions.[2] The entrance of relatively low skilled migrants may however difficult their integration and lead destination countries to implement immigration restrictions. Abramitzky and Braggion (2006) argue that positively selected immigration should be considered as a driving factor for human capital development and growth in receiving countries. In a similar vein as Glaeser et al. (2004), they suggest that the distinct human capital brought by immigrants to the United States in the past could have been an important determinant of the evolvement of good institutions and therefore a strong predictor of economic growth. Within an empirical model these authors find that the better skilled servants[3] who left for the U.S. in the 18th century decided to serve in the Mainland, whereas the less talented opted for the West Indies as their destination. Their guess is that the relative educational quality of immigrants to the New World could have influenced subsequent economic performance of American regions in distinct directions. Whereas the West Indies enjoyed the higher initial

[1] For an extensive discussion on the "brain-drain" effect, see e.g. Doquier (2006), Stark (1991).
[2] That human capital has a direct effect on growth, has been discussed by Mankiw, Romer and Weil (1992) and Galor and Weil (2000).
[3] European servitude was a common form of migration until the beginning of the nineteenth century, when passages to the New World were very expensive. A work contract for a number of years in the destination country that was sold after arrival to merchants or farmers who needed labor, served to finance the trip.

income, the mainland attracted better skilled immigrants and embraced a higher growth rate. There is no consensus on the impact of selective immigration in Latin America. Based on the composition of sending countries – mainly Italy, Spain and Portugal – it has been claimed that immigration to South America and the Caribbean was poor and illiterate. However, this assumption rests on the comparison between Southern European and Northern European migrants to the U.S. (O'Rourke and Williamson 1999). Of course, Southern Europe had lower educational investments than Northern Europe and consequently migrants from the latter region were relatively richer and more literate. Sánchez-Alonso (2007) on the other hand claims that emigrants from the Iberian Peninsula to Latin America were positively selected from their countries of origin and had higher literacy levels than natives in the target countries. What is true? Were (Southern) European migrants to Latin America actually better or worse skilled than natives in the destination countries? And how were the emigrants selected with respect to those they left behind in their country or region of origin? Latin America in the 19th century offers a good scenario to analyze migrant self-selection, because no entry restrictions in terms of human capital were present yet.[4]

This study also puts a special focus on migrants from the Canary Islands. There are good reasons to assume that they differ from migrants from the Spanish peninsula (though, of course, regional differences exist there too) and other Mediterranean countries. Canary Islanders in Cuban censuses are the only Spanish group that is identified as such and not as "Spanish". The Canary Islands was one of the provinces with the highest emigration rates within Spain (see Figure 5.1), although they were relatively illiterate and poor.[5] High

[4] Argentina was the first Latin American country to impose some restrictions on immigration after the 1890s (Timmer and Williams 1998), but this analysis is based on evidence before that.

[5] Sánchez-Alonso (2000) argues that Spanish emigration in the 19th century was income and literacy constrained. In provinces with higher literacy and income people had better access to information on foreign

emigration from the Islands to the Spanish colonies in America started long before the age of mass migration and was characterized by its family structure and its permanent nature. Uruguay and Cuba were two of the three most important destination countries for Canary Islanders in the 19th century (Macías Hernández 1992).[6] A number of new settlements that became cities were founded by Canary migrant families, who made a significant demographic and economic contribution to those countries.[7] In contrast to Castilian settlers in the colonies (like Cuba), who were mostly traders or office holders, or to Iberian immigrants in the new republics (like Uruguay or Argentina), who mostly lived in urban centers, Canary Islanders were generally farmers. They had usually been engaged in agricultural activities at home, and probably belonged to the most vulnerable section of the population that fled the Islands in times of crises and unemployment.[8]

This study contributes to the assessment of emigration patterns to Latin America in the age of mass migration. It especially addresses the question of whether migrants from the Canary Islands and the Spanish peninsula to Latin America were positively or negatively selected with respect to the home region and to the destination country. As suggested above, both "selections" are interesting to analyze, because selective migration has consequences for sending and receiving economies. The human capital of migrants accounts for formal schooling, but also motivation and entrepreneurial spirit. The age heaping technique, now widely used in early human capital research, allows me to quantify

labor markets and were more likely able to finance the move; therefore those provinces display higher emigration rates. However, the Canary Islands do not fit into this picture, since the population was illiterate and poor and still left in high numbers.

[6] The third one is Venezuela.

[7] There is a large number of settlements founded by Canary immigrants in Latin America, such as Montevideo and Maldonado in Uruguay; Matanzas, Santiago de las Vegas, Regla, Guantanamo, Caiba Mocha in Cuba or Campeche and Vagalar in Mexico.

[8] However, there is evidence that after settling in the New World, some of them also enlisted in trade and even in politics (Barreto 2008).

basic numerical skills. The ability to deal with numbers is mainly determined by schooling investments (A'Hearn, Crayen and Baten 2010) and is crucial in a society that is mostly composed of farmers, but especially for entrepreneurial activities. I use various sources to study the selectivity of migrants from the Canary Islands to Cuba and Uruguay: (1) passenger lists to Cuba and Uruguay; (2) censuses from the Canary Islands; (3) census lists from Cuba and Uruguay.[9] In Section 2 I summarize the literature on the determinants of migration decisions and on migrant selectivity. Section 3 offers an insight on the patterns of 19th century emigration from the Canary Islands. In section 4 I proceed with the methodology used and the description of the data sources, section 5 depicts the main results and section 6 concludes.

5.2 Literature Overview

5.2.1 Determinants of Migration

The determinants of migration decisions have been widely discussed in the literature and assessed empirically in some studies. Hatton and Williamson (1998, 2008) and O'Rourke and Williamson (1999) assess the patterns of free migration to the New World during the first mass migration wave (1820s to 1914) and find five key driving factors:[10] (1) Malthusian pressure, measured as population growth 20 or 30 years before in the home country; (2) the degree of industrialization and urbanization at home, which could absorb the demographic pressure; (3) the difference in real wages between the home country and the destination country; (4) the quantity of previous migrants from one particular country

[9] Another source that has been used to study emigration from the Canary Islands consists of the boarding licenses that were expelled by the town councils (Hernández Garcia 1977). It was necessary to request such a license to emigrate by legal ways. However, assessing the sources is difficult.

[10] Mass emigration started from the 1820s with the British Islanders as pioneers, but Southern European migration followed only in the 1880s, most probably due to income-constraints and to the demand for unskilled labor at home (Sánchez-Alonso 2000).

or region (chain migration)[11] and (5) economic and political conditions in the destination country. Other factors can be added, such as agricultural crises and famines in Europe, the movement from sail to steam which reduced migration costs dramatically (over the course of the 1850s, steamships began to account for a larger, though still minority, percentage of immigrant travel), and the presence or absence of immigration restrictions.

As to migration decisions to the Southern continent, Williamson (1999) suggests that the Latin American wage advantage – especially in labor scarce areas like Cuba, Uruguay and Argentina – over the western and central Mediterranean (Italy, Portugal and Spain) in the 19th century was a major factor of attraction for immigrants. This explains also why immigrants to Latin America came mostly from Southern Europe. A former incentive for European migrants were the generous subsidies that especially Brazil and Argentina, but also Cuba and Uruguay, started adopting in the 1880s to countervail the relative labor scarcity[12] and to be able to compete for immigrants with North America. Moreover, in the most important sending countries for migrants to Latin America, Spain and Italy, the rapid population growth was not absorbed by industrialization and urbanization such as in Britain in the 18th century, but it pushed external migration (see Sánchez Alonso 2000).

5.2.2 Migrant Selectivity

That emigrants are not a random sample of the population has long been uncontested (Borjas 1987, Roy 1951). However, the self-selectivity of migrants (particularly

[11] Simone Wegge (2002) proves that migration networks have an effect on the future quantity of migrants (but also on the type of migrants) in a study on German emigration to the U.S. Yannay Spitzer (2013) confirms the importance of chain migration for Jewish migrants from Eastern Europe to the U.S.
[12] Labor scarcity, however, declined dramatically in the beginning of the 20th century, given that mass migration by itself probably explained from 50 to 70 percent of the real wage convergence in the late 19th century Atlantic economy (Taylor and Williamson 1997, O'Rourke and Williamson 1999).

educational selectivity) has only been assessed empirically for a number of sending countries and mostly for migrants to the U.S. (for example, by Hatton and Williamson 1998, Feliciano 2005, Wegge 2002, Abramitzky et al. 2012).[13] The selectivity of migrants in the economics literature commonly refers to how migrants perform compared to the home country population. That definition will also apply here.

Feliciano (2005) looks at the educational level of immigrants from 32 countries in the U.S. and concludes that migrant selectivity depends a lot on the distance from the source country to the destination country. Due to the higher costs of migrating long distances, migrants from countries that are farther from the destination country are more highly selected.[14]

Most scholars corroborate the theory of Massey et al. (1993) that, although migrants tend to be positively selected, they become less highly selected over time after successive migration waves from a particular country. Hatton and Williamson (2008) empirically assess migrant selectivity to the U.S. during both globalization periods – one before the First World War and the other one since the Second World War – and confirm that in both time spans the same phenomenon of declining educational quality occurred (their analysis is based on occupational data and some evidence on literacy rates). Friends-and-family-ties in the U.S. could be bond after the first migration wave. With each new act of migration networks expanded, reducing migration costs for the source country population, because information on the U.S. labor market was more easily accessible.

[13] Some exceptions are Sánchez-Alonso (2000, 2007) for 19th century Spanish emigration and Stolz and Baten (2012) for migrants to Brazil.

[14] On the other hand, the theory provided by Borjas (1987) on the basis of Roy's (1951) *self-selection model* claims that migrants are not always positively selected. It holds that only a higher relative inequality (and higher relative rewards to skill premia) in the destination country attracts highly skilled migrants. This theory arises from the observation of the opposite case of Mexican migration to the U.S. High-skilled Mexicans supposedly preferred to stay in their home country, where they were rewarded relatively well under the more unequal Mexican system, and only low-skilled Mexicans chose to emigrate because they were the most relatively disadvantaged in Mexico.

Moreover, the received remittances might have helped to overcome poverty constrains. As a consequence, over time, persons who were not relatively well educated or skilled began to migrate.

The most common view is that migrants tend to be positively selected for labor market success. Hatton and Williamson (1998, 2002, 2004, 2008) argue that due to the entrepreneurial spirit and ambition necessary to start a new life in a foreign country, migrants are usually positively selected. Moreover, poverty constraints for most of the history did not allow the poorest section of the population to emigrate. Chiswick (1978, 2000) shows that, after controlling for years of schooling and ethnic origin, foreign-born had higher earnings than natives – after a period of adjustment – in the 1970s United States. This is consistent with the view that immigrants tend to be favorably self-selected in terms of ability or human capital and do well in the host labor market.[15]

What characterized the selectivity of migrants for 19th century Latin America?[16] Based on the composition of sending countries – mainly Italy, Spain and Portugal – it has been claimed that immigration to South America was poor and illiterate, for example by O'Rourke and Williamson (1999). However, this assumption rests on the comparison between migrants from Southern Europe and Northern Europe to the U.S. But how were (Southern) European migrants to Latin America actually selected with respect to the country of origin? And, did they contribute to raise the average human capital in the destination countries? Sánchez-Alonso (2007) analyzes literacy rates of Italian, Spanish and Portuguese immigrants in Argentina in the census of 1895 and finds higher literacy rates for them than for the population of origin, especially for the Iberian Peninsula. That

---

[15] This selectivity applies only to economic migrants and those who migrate to join family members or as political refugees are presumably not as highly selected (Chiswick 2000).

[16] One study that found a positive impact is Stolz, Baten and Botelho (2013). They showed that in the case of Brazil, the human capital brought in by European immigrants in the 19th century had a long-term positive effect on GDP per capita.

means that migrants were positively selected. Her guess is that literacy was an important determinant of the propensity of migration at the time, because being able to read eased the access to information. Moreover, she finds that migrants were more literate than the native population in her sample and, thus, Latin America profited from immigration. This study differs from Sánchez-Alonso's in a number of aspects. Her findings are based mostly on the Argentinean case. I shed light on migrants' relative skill levels in two additional countries, Cuba and Uruguay. Moreover, I do not compare literacy rates, but a measure of basic numerical skills explained in section 4. This measure, calculated from age statements available in censuses and passenger lists, makes it possible to avoid the data scarcity problem of other human capital approximating variables.

5.3 The Genesis of Canarian Migration

The Canary Islands was one of the provinces with the highest emigration rates throughout the 19th century (see Figure 5.1). Migration from the "Fortunate Islands" was notoriously focused on Latin America, and especially on Cuba, Uruguay and Venezuela.[17] It was not so much concentrated in urban regions in the destination countries, such as migrants from other Spanish provinces, but rather in the rural parts. The demographic pressure that arose due to the rapid population increase of the 18th century was a precondition for the high emigration rates later (Hernández García 1977). Other factors driving emigration pointed out in the literature are the importance of droughts and agricultural crises in the Canary Islands, wage differentials, migration chains and immigration policies of receiving countries.

[17] During the age of mass migration, Latin America was the second major destination region after the United States. It received about 13 million Europeans from 1879 to 1930, mostly from Italy, Spain and Portugal (Sánchez-Alonso 2000).

The 19th century was characterized by important crises in the agrarian export sector, first in viticulture (1830) and later in the grain and cochineal exports (1870). At the beginning of the century, protectionist policies that had been introduced in Spain seriously harmed the economy of the Canary Islands, which was traditionally based on trade with non-Castilian regions, especially England and Holland, but also the American colonies (a short recovery of the agrarian export sector followed the enactment of the "law of open ports" in 1852, by which international commerce was liberalized for the Canaries, until the next crisis hit the grain and cochineal exports).

The crisis in viticulture coincided with the relatively calmed situation in Latin America after the wars of independence, and encouraged emigration.[18] Between 1835 and 1850, up to 35,000 people out of a total population of 235,000[19] left the Canary Islands to Cuba, Uruguay and Venezuela by legal ways, according to Macías Hernández (1992).[20] Probably, another 10,000 left the Canary Islands between 1850 and 1859 (Yánez Gallardo 1994). Canarians found their way to the new republics of Latin America despite that the Spanish government had generally banned emigration without an official boarding license since 1836 (except to the colonies Cuba and Puerto Rico which remained still faithful to the Spanish Crown).[21] In 1853, a real circular order abolished the ban of emigration pointing out especially the exit of the abundant unemployed residents of the Canary Islands, following the great fall of the agrarian export sector. Emigration statistics since

[18] The first quarter of the 19th century has experienced a slow decrease in emigration due to the wars of independence in the South American continent, which lasted until 1830 (Macías Hernández 1992).

[19] The census of 1857 accounts for 233,784 inhabitants in the Canary Islands (source: http://www.ine.es/inebaseweb/)

[20] It has been estimated that 20 percent of migrants left by illegal ways, being the major reason the wish to evade the military service (Sánchez-Alonso 2000).

[21] Cuba and Puerto Rico were the last Spanish colonies; they only became independent in 1898. In the first half of the century, relations between Spain and the future Republics of America, which were at that time fighting for their independence were difficult, but also later due to the political instability in those regions. Furthermore, there was low demand for free labor, given that the Americas enjoyed either a large native population (in countries like Mexico) or they had imported slave labor (such as, for example, Brazil). In addition to this, the kingdom of Spain followed the policy of keeping the peasant population in the country since the 18th century (Sánchez-Alonso 2000).

1860 show that the Canary Islands accounted for among the highest emigration rates of Spain, in total and out of a thousand residents (see Figure 5.1 and Yánez Gallardo 1994). Emigration reached its peak levels in the 1880s – a mean of 18 per thousand residents per year – coinciding with a price crash of the export goods grain and cochineal, as natural ink was substituted by artificial ink produced elsewhere (Macías Hernández 1992, p. 134).

What made Latin America attractive to Canary Islanders? Sánchez-Alonso (2000) points out that sharing language, religion and cultural identity are factors that were probably decisive for the destination choice of all Spanish emigrants. Those cultural ties existed with Latin America because it had been under Spanish colonial rule for most of the time since the 16th century. The crown had encouraged Spanish emigration to the colonies already in the 17th and 18th centuries, partly to countervail the population surplus in some areas of the motherland and partly to occupy vast lands in conflict areas to which it had legal titles but needed to be defended against other European powers (Bethell 1984). In Cuba, European (and Asian) immigration was encouraged to balance out the massive importation of African slaves and "whiten" the population (Bethell 1984). After their independence, a number of Latin American countries – especially the ones that suffered from labor scarcity such as Uruguay and Argentina – also adopted receiving attitudes to colonize the unpopulated areas, for example by subsidizing passages. Some republican governments especially promoted the immigration of North-Europeans, who were thought to be culturally "superior" and therefore to contribute to social and economic development (Sánchez-Alonso 2007). The preference of Canary Islanders for Cuba and Uruguay, but also Puerto Rico and Venezuela, can probably be attributed to the geographical proximity. The fare price and the opportunity costs of lost working days were lower when migration was directed to countries on the Atlantic coast and also when leaving from the Canary Islands instead of another Spanish port. Emigration of Canary Islanders in the 19th century was most presumably facilitated by friends-and-family ties, since the presence of "isleños" in some Latin American regions was notable since the 17th century (Macías Hernández 1992). The demand for labor in the booming coffee and tobacco plantations in Cuba[22] and the higher wages for agricultural laborers in Cuba and Uruguay must have attracted Canarian laborers. Williamson (1999) shows that real wages in the agricultural sector, compared to the mean of Mediterranean countries, were about 3.4 times higher in Uruguay in the 1890s and about 1.6 times higher in Cuba in the first decade of the 20th century. Macías Hernández (1992) estimates that compared to the Canary Islands wages were even three to five times higher in Cuba in the 1910s. All those factors help us understand the numerous movements from Canary Islanders to Latin America.

How significant was the Canarian contribution to the destination countries? In the case of Cuba and Uruguay, Canary Islanders made up the largest immigrant group. The demographic contribution is clearly significant in both destination countries. According to the Cuban census of 1860, 47.3 percent of the Spanish born migrants living in Cuba came from the Canary Islands and the Canarian migrants represented 6.3 percent of the white population on the Caribbean island including white natives (Macías Hernández 1992).[23] Agricultural laborers from the Archipelago were mostly recruited by plantation owners through the system of indentured servitude ("contratas") for the transition to a non-slavery production.[24] In Cuba, the majority of them were engaged either in the sugar or the tobacco industry.[25] Immigrants of Canarian origin are supposed to have been mostly permanent in nature during the first part of the century, given that they were predominantly not engaged

[22] Tobacco was originally introduced by "isleños" in Cuba in the early 18th century and since then Canarian chain migration to Cuba was in place (Macías Hernández 1992).
[23] They account for 7.12% of the white population in our data from based on various censuses between 1862 and 1887.
[24] Within this system, passage costs were assumed by the contractor and, in compensation, immigrants were bound to work for a determined number of years for him. The conditions for the laborers seem to have been relatively poor and slave-like.
[25] In this context, the importance of human capital brought in terms of know-how on the cultivation of tobacco should not be disregarded (Suárez Bosa and Ojeda Déniz 2006).

in commerce but in agriculture, the passage was expensive (about 80 days of work) and the number of women belonging to this group was relatively high. However, in the late 19th century, the share of seasonal migrants increased, who tended to stay only for the sugar harvest period, which turned shorter every year (Macías Hernández 1992). Plantation owners in Cuba in the 19th century were interested in a mass immigration directed to the rural sector to combat labor scarcity and dump wages of free laborers. Canarian rural laborers surely contributed to the rise of the export sector, and consequently, to GDP growth. In Uruguay, immigration from the Canary Islands had a long tradition. Already in 1726, an equal number of families from the Canary Islands and Galicia founded the city of Montevideo. Canary families were also involved in the foundation of the cities of Canelones and Pando in the second half of the 18th century (Fernández 1964). However, the mass immigration of Canarians took place the 1830s, when the system of "contratas" was popularized in Uruguay. Those immigrants made a significant contribution to a high population increase in the rural parts of the country. In our census sample, Canary Islanders make up 24% of all immigrants and 9% of the whole population. The new settlers changed the structure of the economy in that they introduced the plantation of cereals (maize and grain) in a region that was only dedicated to cattle raising activities until then (Barreto 2008).

How are Canarian emigrants selected in terms of human capital with respect to the home population? The characteristics of Canarian emigration lead me to expect that, on average, emigrants were not positively selected. Since the most important driving factors of emigration were the crises in the agricultural export sector, probably the most vulnerable section of the population, those who became unemployed, left for America. Moreover, the literature points out that chain migration, which played a very important role in this context, reduces migration costs and consequently diminishes the positive selectivity of migrants.[26]

How do Canarians perform in terms of human capital compared to immigrants from other countries or regions and to the destination country? In comparison to the Spanish average, the educational level in the Canary Islands was always much lower. Literacy rates from 1860 until the end of the century were 10 to 16 percent lower than the Spanish average and numeracy levels from 1840 about 5 percent lower (see Figure 5.2 and 5.3). Thus, it is expected that Canary immigrants in Latin America perform worse in terms of education than immigrants from the rest of Spain, and from other countries in Europe. The question of whether "isleños" still brought a higher human capital in terms of numeracy relative to natives will also be analyzed in the following chapters.

5.4 Data Sources and Method

5.4.1 Source Description

I use census data from the Canary Islands and the Spanish peninsula as well as passenger lists to Uruguay and Cuba and census lists from Cuba, Uruguay and Argentina. These sources allow me to assess the skill selectivity of Canary and peninsular emigrants, as well as to compare immigrants of different origins with one another and with the native population. Table 5.1 displays a summary of the datasets. The Uruguayan censuses of the 19th century in various provinces (Minas, Canelones, Paysandu, Montevideo, Maldonado and Soriano) are kept in the national archive of Uruguay and can be found in excel-format on www.pueblosynumeros.fcs.edu.uy. This dataset includes information on 7,563 individuals of ages between 23 and 62 years, of which 35.8% are foreign born. Almost

[26] We cannot be sure which mechanism is more significant: that the agricultural laborers who migrated in times of crises were not the best in their profession and therefore became unemployed or that those who left were the more entrepreneurial and with higher aspirations, thus positively selected.

nine percent of the total population in the sample can be identified as born in the Canary Islands.[27] Cuban censuses of the second half of the 19th century were provided by Linda Twrdek, who compiled a data set out of the original census documents that are kept at the national archive of Cuba, in La Habana (see Twrdek 2011). Those census lists comprise the provinces of Pinar del Rio, Matanzas and La Habana and thus cover the western part of the island. The sample includes 14,453 individuals, of whom 55.8 percent are foreign born and 8.9 percent of total originate from the Canary Islands (see Appendix Table A.3). The Uruguayan passenger ship lists of between 1833 and 1863 are downloaded from the website of Spanish archives (www.pares.es) and the original documents can be found in the national archive of Uruguay under the fond of the Police of Montevideo. This data base includes only Canary Islanders, as I only searched for them. In the 19th century the police of Montevideo took tedious note of almost all foreigners' place of birth, profession, residence place in Uruguay and age at their arrival (Martínez Díaz 1982). A passenger list from the Cuban port of Matanzas that consists mostly of Canary Islanders ("Registro pasajeros que desembarcan en el puerto en 1854") was collected in the Provincial Archive of Las Palmas, as well as a census list of the city of Las Palmas from 1874. Moreover, I use the Spanish national censuses of 1877 and 1900, downloaded from the website of the national statistics office (www.ine.es). The sample of the Argentinean census of 1895 was provided by Somoza and Lattes (1967). In this census 42% are foreigners. No Canary Islanders can be identified here.

5.4.2 Representativeness of the Sources

[27] I cannot rule out the possibility that within those who reported to be born in Spain, naming no city or province, one part was born in the Canary Islands.

The representativeness of the sources was carefully assessed within my possibilities. The Cuban censuses are representative for rural and urban shares in the three historical provinces of Pinar del Rio, La Habana and Matanzas. These provinces accounted for 50% of the population in 1899. Among the population of La Habana 87% lives in an urban region, 10% in Pinar del Rio and 51% in Matanzas. This is in accordance with the census of 1899 that was carried out under U.S. rule (Report on the census of Cuba 1899). The sample represents all social classes and ethnic groups. 6,388 individuals are born in Cuba and 8,065 have a migrant origin. Among the migrants, 48% were Asians[28], 34.5% were born in Spain (half of them in the Canary Islands), 16.4% in Africa and only 0.6% came from the rest of Europe or other places. Moreover, only 10% of migrants are female, but we know that migrants in fact were in majority males. The foreign-born group should be representative in its composition of origin and gender.[29] However, there is an overall bias towards the western part of the island (today's provinces of Pinar del Rio, Artemisa, La Habana, Mayabeque and Matanzas) which was the most densely populated area, even among the rural population, and the residence place of the major part of the island's elite that also owned the sugar industry (Report on the census of Cuba 1899 and Zeuske 2002). Due to this, I should take into account that the numeracy in the Cuban sample might be slightly higher than the national average. When looking at other human capital measures, the census of 1899 indicates that literacy rates in the province of La Habana, Pinar del Rio

[28] Asian workers, so-called "coolies" were imported in large numbers as "contract laborers" to work in the sugar plantations in Cuba in semi-slavery conditions, especially from 1848 (Zeuske 2002). It is estimated that 150,000 Asians arrived in Cuba between 1853 and 1872 (Bethell 1993). Large landowners wanted to replace African slaves by Asian contract workers for their fear of another "Haitian" slave revolution and because prices for slaves were rising high.

[29] It is interesting to state that numeracy among natives was equal for males and females, but there was a gender gap of 11 ABCC points among migrants.

and Matanzas, which comprise most of the sample used here, are respectively 53.1%, 18.9% and 34.8%, together four percent above the national average of 36%.[30]

The Uruguayan census sample includes the departments of Paysandu, Soriano, Canelones, Maldonado, Minas and Montevideo. In those departments resides 64% of the national population in 1852.[31] The sample is nationwide representative of Uruguay, since the capital is included as well as rural and urban areas. Also here, 35.8% are foreign born and migrants are mostly male (73%), which roughly coincides with the shares indicated in the summary statistics of the censuses of 1852 and 1860. The largest group here is the Canary Islanders who account for 23.7% of all migrants; 17.6% of migrant come from other regions in Spain, 19.3% come from other European countries, mostly France and Italy, 16.9% originate from other Latin American countries and 17.2% from Africa (see Appendix Table A.1). The group of foreigners should also be representative by place of origin and by gender.

The census sample of Argentina in 1895, which includes also many foreign born since mass immigration had taken place in the decade before the census, was collected by Somoza and Lattes (1967) and constitutes a representative nationwide sample of the population (Manzel, Baten and Stolz 2012).

The census of Las Palmas de Gran Canaria of 1974 only includes the population of the city and its surroundings (mainly urban). However, the numeracy level was counterchecked with that of the Canary Islands province within the national census of 1877 (which includes only aggregate numbers by age) and turned to be almost equal.

5.4.3 The Method

[30] Weighted by the population of every province, the average literacy of the three provinces is 40.3%.
[31] Anuario estadístico de la Republica Oriental del Uruguay, Libro XVIII del "Anuario" y XXXII de las Publicaciones de la Dirección General de Estadística, Tomo 1, p. 44.

The method used here to measure human capital is the age heaping technique. It allows me to analyze and compare skill levels by using population lists that include age statements. Age heaping measures an important component of human capital, which is numeracy, or the ability to deal with numbers. It makes use of the phenomenon given in historical times – and in developing countries even today – that the poorly educated people tend to round their age on multiples of five when they are asked to state it for census enumerations or other records. Age heaping is most commonly measured by the Whipple index, which displays the ratio between the preferred ages and the rest.[32] This index measures the proportion of people who state an age ending in a five or zero, assuming that each terminal digit should appear with the same frequency in the "true" age distribution (or the degree to which the distribution of age statements approaches an equal distribution).[33]

$$(1)\ \ Wh = \left(\frac{(Age25 + Age30 + Age35 + ... + Age60)}{1/5 * (Age23 + Age24 + Age25 + ... + Age62)} \right) \times 100$$

Only the age range of 23 to 62 is included (for Argentina also the ages 63 to 72 are included), as younger persons usually can guess their age more accurately or the household head could have reported it for them. The older ages are excluded because of selective mortality bias and the tendency of the elder generation to exaggerate their age.

For an easier interpretation, A'Hearn, Baten, and Crayen (2010) suggested another index, which we call the ABCC index.[34] This is a simple linear transformation of the

[32] A'Hearn, Baten and Crayen (2010) found that this index is the only one that fulfills the desired properties of scale independence (a linear response to the degree of heaping), and that it ranks samples with different degrees of heaping reliably. It is also used by the UN Statistics Division.

[33] A value of 500 means an age distribution with ages ending only on multiples of five, whereas 100 indicates no heaping patterns on multiples of five, that is exactly 20 percent of the population reported an age ending in a multiple of five.

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

Whipple index and yields an estimate of the share of individuals who correctly report their age:

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

The share of persons able to report an exact age has been shown to be highly correlated with other measures of human capital, such as literacy and schooling, across countries and individuals and over time (Mokyr 1983; A'Hearn, Baten, and Crayen 2010).[35]

5.5 Main Results

5.5.1 Skill Selectivity of Those who left the Canary Islands

As mentioned above, those who left the Canary Islands for Cuba and Uruguay in the 19th century were not expected to be positively selected with respect to those who stayed behind. Since the most important driving factors for emigration were the crises in the agricultural export sector, probably the most vulnerable section of the population left. Many became unemployed with the fall of agricultural exports and therefore took their way to America, where they could make use of their skills and experience in agriculture in the booming plantation economies. Moreover, chain migration played a very important role in the movements from the archipelago to America. The first generations of migrants usually need a special ambition or entrepreneurial spirit to make the move. However, the literature has pointed out that second and higher generations of migrants are less ambitioned and less positively selected, because friends-and-family ties reduce the information costs and financial costs of migrating considerably. Previous emigrants provided latecomers with remittances to finance the passage or even arranged prepaid tickets for them (O'Rourke and Williamson 1999). With the age heaping technique I can

[35] This paragraph is based on Juif and Baten (2013).

compare directly the numeracy levels of residents of the Canary Islands from 1820 until 1850 with residents in Cuba born in the Canary Islands. The result is a gap of between fifteen numeracy points in 1820 and eight numeracy points in 1840 (Figure 5.4).

I also test the difference in numeracy levels between the population residing in the Canary Islands and Canary Islanders in Cuba and in Uruguay in logistic regressions (Table 5.2). The dependent dummy variable (numeracy) is 0 when the age stated was a multiple of 5 and 1 if it wasn't. After controlling for the share of females (since migrants are more often males) and for the age group "23 to 33" (given that the youngest age group generally displays higher numeracy levels) the coefficient of the explanatory dummy variable "all Canary migrants" is negative and significant (Model 1). Thus emigrants were negatively selected. Canary emigrants to Cuba were 7.50% less numerate than non migrants, whereas the coefficient for Canary emigrants to Uruguay is also negative but insignificant (Model 2).

Canary Islanders in Cuba and Uruguay are also compared with other immigrants and the destination country population. Figure 5.5 shows that Canary migrants in Cuba perform worse in terms of numerical skills than Cuban non migrants (by around 10% ABCC points) and also than migrants from the Spanish peninsula. They can be found on an ABCC level comparable to that of immigrants from Africa. A logistic regression in Table 5.3 Model 1 confirms this finding. The result is consistent with the fact that the Spanish population in Cuba consisted mostly of civil servants who held important posts in the administration, but this did not account for Canary Islanders. As already discussed, Canarian migrants were mostly engaged in agriculture. When looking at the occupations of all immigrants of Canary origin, only 19% in Cuba and 32.4% in Uruguay were not engaged in an occupation related to farming (see Appendix Table A.4 and A.5). Most of the latter group can be found in the passenger entry lists of the police of Montevideo. It is well-known that migrants often indicated a false profession to the authorities when they thought that it would provide them better possibilities to immigrate into a country (Cohn 2009). Moreover, the passenger entry lists include both temporal and permanent migrants, whereas the census lists include probably only permanent migrants. In census lists where Canarians are recorded probably years after immigration, almost all are farmers. However, Canarian migrants to Uruguay seem to be less negatively selected with respect to the receiving population (see Figure 5.6 and Table 5.3 Model 2). If comparing the skill level of immigrants from the Canary Islands in Cuba and in Uruguay, Canarians in Uruguay seem to have a higher skill level than those Cuba, since "isleños" in the republic reach an ABCC level of 82 already in 1800, whereas the numeracy level of the Caribbean island is still 69 in 1820. This makes sense given that Uruguay was not founded until the late 18th century, and some Canary Islanders included in the Uruguayan samples probably had still been offered large pieces of land to occupy the area (others of course came as contract workers). Cuba on the other hand was a plantation economy that attracted high numbers of low skilled laborers, who mostly came as contract workers in quasi-coerced-labor conditions. Moreover, migrating to Uruguay was more expensive due to the longer distance and therefore the poorest could maybe not afford to migrate to Uruguay, but to Cuba.

5.5.2 The Other Immigrants

For Argentina I can confirm that immigration made an important contribution to raise the average human capital level. Migrants to Argentina were on average 13.75% more numerate than natives (Table 5.4 Model 5). The Spanish immigrants were 12.5% more numerate (Table 5.5 Model 2), even though they were negatively selected out of the sending country population (see Figure 5.9). European migrants from all nationalities in Argentina seem to be better skilled than the native population, but especially those from Germany, France and Great Britain (Table 5.4 Model 6 and Figure 5.7). These Western Europeans account for 14.7% percent of all immigrants here, which is much more than in

Cuba and Uruguay. Regarding immigrants from Latin America, Brazilians and Uruguayans display higher levels of numeracy than natives, whereas Chileans are moderately less numerate (Figure 5.8).

My findings for Cuba are different. Migrant numeracy seems to be on average 7.5% lower than native numeracy (see Table 5.4). The same is true if we control for migrants from Africa, Latin America and Asia. When leaving Canary Islanders aside, Spanish numeracy is only 5% lower than Cuban (Table 5.3 Model 1). Spanish emigrants to Cuba are not selected in terms of skills from the sending country population (see Figure 5.9). In fact, ABCC levels of Spanish (non-Canarian) migrants to Cuba and the average of the Spanish peninsula population are the same. However, native numeracy levels for Cuba at the time seem to have been considerably higher than Argentinean and even higher than Spanish. Although still a colony of Spain, the Island of Cuba is claimed to have been more developed than its Metropolis, especially as an international exporter (Zeuske 2002, p. 93). This could have had positive effects on numeracy levels. Argentina could catch up in the late 19th century when government started adopting investments in public schooling to attract immigrants from Europe (Stolz, Baten and Botelho 2013; Manzel, Baten and Stolz 2012). These policy measures and the human capital brought in by European migrants from the 1880s helped to raise the average educational levels in the republic and become a leading country in Latin America in terms of human capital (Manzel, Baten and Stolz 2012). In Argentina the younger natives caught up with immigrants' numeracy levels as can be seen in Figure 5.7 and Table 5.5. The section of the population born in the 1850s and 1860s already had a good access to education. Later birth cohorts of natives probably overtook new immigrants in terms of human capital, since in 1884 primary education became compulsory and other measures like literacy indicate that Argeninean human capital levels increased sharply (Sánchez-Alonso 2013). In Cuba on the other hand, immigrant numeracy increased faster than native numeracy and this way also here the numeracy gap between natives and immigrants decreased over time (Figure 5.5).

In Uruguay, human capital levels of migrants from Europe do not differ significantly from those of the native population, but migrants from Africa – who came as slaves and are therefore expected to be unskilled – and migrants from South America perform worse (see Table 5.3 Model 2). Also here among European immigrants the least come from Northern Europe. Almost all (at least 82%) were born in Spain, Italy or Portugal.

5.6 Conclusions

I analyzed migrants' contribution to Latin America's numeracy and found different results depending on the destination country. The composition of immigrant source countries and the initial human capital levels of natives played a major role. While the majority of immigrants in Argentina came from Italy and the second largest group was the Spanish, considerable numbers of Northern Europeans settled here. Most migrants to Cuba came from Asia, Africa or Spain. This is one reason for the higher human capital advantage of all immigrants over natives in Argentina than in Cuba. Moreover, the average numeracy level of natives was higher Cuba than in Argentina. Therefore I can confirm that in Argentina the arrival of the higher educated immigrants at the end of the 19th century – and probably the policy of investing into public schooling to attract Europeans – helped to improve the overall skill level, as documented by Sánchez-Alonso (2007), and overtake the Caribbean later (Manzel, Baten and Stolz 2012). But for Cuba and Uruguay, I did not find such a human capital contribution of immigrants. On the contrary: non European and Southern European immigrants performed worse than natives in terms of numerical skills.

I also found a slight negative selectivity of Spanish emigrants to Argentina compared to the sending country population, but no particular selectivity of Spaniards in

Cuba. This might be true because the Argentinean government subsidized passages for European migrants for a couple of years in the 1880s. The poorer and less numerate section of the population might then have been able to afford the move. Cuba might be a special case since it still had colonial ties to Spain throughout the 19th century and the crown sent well-paid civil servants to the Caribbean Island.

A special focus was set on migrants from the Canary Islands. They were negatively selected compared to the population of the Archipelago. Relative to the destination countries, they display lower numeracy levels than Cubans and no significant difference from Uruguayans. I argue that the driving factors for emigration were mainly the crises in the agricultural export sector and thus Canarians emigrated to escape poverty in their home region. Many had become unemployed with the fall of agricultural exports and therefore took their way to America to work in the booming plantation economies. Even the poorest Canary Islanders could probably finance the move due to the system of "contratas" that was very common in both destinations and therefore almost no income constraints were present. Moreover, chain migration is known to have a negative effect on migrant skill selection and it played a very important role in the movements from the Archipelago to America.

5.7 References

Abramitzky, R., and Braggion, F. (2006). Migration and Human Capital: Self-Selection of Indentured Servants to the Americas. *Journal of Economic History*, 66(4): 882-905.

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2012). Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, 102(5): 1832-56.

A'Hearn, B., Baten, J., and Crayen, D. (2010). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. *Journal of Economic History*, 47(1): 82-99

Barreto, I. (2008). *Canarios de Ayer. La Inmigración Canaria al Uruguay*. Montevideo: Imprenta Don Bosco.

Bethell, L (1984). *The Cambridge History of Latin America*. Vol. 2: Colonial America. Cambridge: Cambridge University Press.

Bethell, L. (1993). *Cuba: A Short History*. Cambridge: Cambridge University Press.

Borjas, G.J. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, 77:531-53.

Chiswick, B. R. (1978). The Effect of Americanization on the Earnings of Foreign-born Men. *Journal of Political Economy*, 86(5): 897-921.

Chiswick, B. R. (2000). Are Immigrants Favorably Self-Selected? An Economic Analysis. In *Migration Theory: Talking Across Disciplines*, edited by C. D. Brettell, and J. F. Hollifield. New York: Routledge.

Cohn, R.L. (2009). *Mass Migration under Sail: European Immigration to the Antebellum United States*. Cambridge: Cambridge University Press.

Docquier, F., (2006). Brain Drain and Inequality Across Nations. IZA Discussion Paper Series. IZA DP 2440.

Family Search Genealogy Catalogue [https://familysearch.org/]

Feliciano, C., (2005). Educational Selectivity in U.S. Immigration: How Do Immigrants Compare to those left behind? *Demography*, 42(1): 131-152.

Fernández. D. W. (1964). Los Canarios en Uruguay. *Revista de Historia Canaria*, 141-148: 56-69.

Galor, O., and Weil, D. N. (2000). Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review*, 90(4): 806-828.

Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004). Do Institutions Cause Growth?. *Journal of economic Growth*, 9(3): 271-303.

Hatton, T. J., and Williamson, J. G. (1998). *The Age of Mass Migration: Causes and Economic Impact.* New York: Oxford University Press.

Hatton, T. J., and Williamson, J. G. (2002).What Fundamentals Drive World Migration?. CEPR Discussion Papers 458, Centre for Economic Policy Research, Research School of Economics, Australian National University.

Hatton, T. J., and Williamson, J. G. (2004). International Migration in the Long-Run: Positive Selection, Negative Selection and Policy. IZA Discussion Papers 1304, Institute for the Study of Labor (IZA).

Hatton, T. J., and Williamson, J. G. (2008). *Global Migration and the World Economy: Two Centuries of Policy and Performance*. Cambridge, MA: MIT Press Books.

Hernández García, J. (1977). *La Emigración de las Islas Canarias en el siglo XIX*. Las Palmas: Cabildo Insular de Gran Canaria.

Juif, D., and Baten, J. (2013). On the Human Capital of 'Inca' Indios before and after the Spanish Conquest: Was there a 'Pre-Colonial Legacy'?. *Explorations in Economic History*, 50 (1): 227–241.

López, C. B. (2005). *Estadísticas históricas de España: siglos XIX-XX* (Vol. 1). A. Carreras, and X. Tafunell (eds.). Fundación BBVA.

Macías Hernández, A. M. (1992). *La Migración Canaria 1500-1980*. Gijón: Ediciones Júcar.

Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2): 407-437.

Manzel, K., and Baten, J. (2010). Gender Equality and Inequality in Numeracy: The Case of Latin America and the Caribbean, 1880-1949. *Revista de Historia Económica*, 27(1): 37-73.

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

Martínez Díaz, N. (1982). *La Emigración Clandestina desde las Islas Canarias al Uruguay. Formas de Incorporación Social. Ensayo de Estudio Cuantitativo*. V Coloquio de Historia Canario-Americana. Las Palmas de Gran Canaria, Francisco Morales Padrón (ed.), Vol. 1: 249-314.

Massey, D. S., Arrango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J. E. (1993). Theories of International Migration: A Review and Appraisal. *Population and Development Review*, 19 (3): 431–66.

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

O'Rourke, K. H., and Williamson, J. G. (1999). *Globalization and History*. Cambridge, MA: MIT Press Books.

Portal de Archivos Españoles, Ministerio de Cultura y Deporte [http://pares.mcu.es/MovimientosMigratorios]

Report on the census of Cuba 1899. University Libraries, University of South Carolina [http://digital.tcl.sc.edu/cdm/compoundobject/collection/CCC/id/1683/rec/1]

Roy, A. (1951). Some Thoughts on the Distribution of Earnings. Oxford Economic Papers, 3: 135–46.

Sánchez-Alonso, B. (2000). Those Who Left and Those Who Stayed Behind: Explaining Emigration From the Regions of Spain, 1880–1914. *The Journal of Economic History*, 60(3): 730-755.

Sánchez-Alonso, B. (2007). The other Europeans: Immigration into Latin America and the International Labour Market (1870-1930). *Journal of Iberian and Latin American Economic History*, 12(3): 395-426.

Sánchez-Alonso, B. (2013). Making Sense of Immigration Policy: Argentina, 1870-1930. *Economic History Review*, 66(2): 601-627.

Somoza, J., and Lattes, A. (1967). Muestras de los dos Primeros Censos Nacionales de Población, 1869 y 1895. Buenos Aires: Instituto DiTella working paper 46.

Spitzer, Y. (2013). Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914. Working paper presented at the Economic History Association Annual Meeting in Washington, 2013.

Stark, O., and J. E. Taylor (1991). Migration Incentives, Migration Types: the Role of Relative Deprivation. *Economic Journal*, 101(408): 1163–78.

Stolz, Y., and Baten, J. (2012). Brain Drain in the Age of Mass Migration: Does Relative Inequality Explain Migrant Selectivity?. *Explorations in Economic History*, 49(2): 205-220.

Stolz, Y., Baten, J., and Botelho, T. (2013). Growth Effects of Nineteenth-century Mass Migrations:"Fome Zero" for Brazil?. *European Review of Economic History*, 17(1): 95-121.

Suárez Bosa, M., and Ojeda Déniz, P. (2006). Los Trabajadores Canarios en Latinoamérica. *Cuadernos Americanos*, 117: 11-46.

Taylor, A. M., and Williamson, J. G. (1997). Convergence in the Age of Mass Migration. *European Review of Economic History*, 1(1): 27-63.

Timmer, A. S., and Williams, J. G. (1998). Immigration Policy prior to the 1930s: Labor Markets, Policy Interactions, and Globalization Backlash. *Population and Development Review*, 24(4): 739-771.

Twrdek, L. (2011). Essays on the Biological Standard of Living in Latin America and the Caribbean. Dissertation, University of Tübingen.

Wegge, S. A. (2002). Occupational Self-selection of Nineteenth-Century German Emigrants: Evidence from the Principality of Hesse-Cassel. *European Review of Economic History*, 6(3): 365-394.

Williamson, J. (1999). Real Wages Inequality and Globalization in Latin America before 1940. *Revista de Historia Economica-Journal of Iberian and Latin American Economic History*, 17(1): 101-142.

Yánez Gallardo, C. (1994). La Emigración Espanola a América (Siglos XIX y XX). Dimensión y Características Cuantitativas. Sevilla: Colección "Cruzar el Charco", Archivo de Indianos.

Zeuske, M. (2002). *Kleine Geschichte Kubas*. München: Verlag C.H. Beck.

5.8 Tables

Table 5.1: Descriptive Statistics of the Sources

Country/ type of register	Province/ department	Year	Obs. No.	Individual data?	Bias relative to total population
Cuba various censuses	Habana, Pinar del Rio, Matanzas	1862-1887	14,453	yes	regional
Uruguay various censuses	Paysandu, Soriano, Canelones, Minas, Maldonado, Montevideo	1827-1858	7,563	yes	no
Argentina national census	All provinces	1895	45,493	yes	no
Las Palmas census	Canary Islands	1874	1,338	Yes	urban
Spain national census	all	1877	115,321	No	no
Spain national census	all	1900	159,353	No	no
Uruguay, passenger list (only Canary Islanders)	Montevideo	1833-1863	254	yes	Only migrants
Cuba, passenger list (only Canary Islanders)	Matanzas	1854	904	yes	Only migrants

Table 5.2: Selectivity of Canarian Migrants in Logistic Regressions

	(1)	(2)
Incl. samples	Canary Islands census and Canarian emigrants	Canary Islands census and Canarian emigrants
Model	Logit mfx	Logit mfx
Female	-2.50	-2.50
	(0.20)	(0.21)
Age 23-32	5.00	3.75
	(0.14)	(0.46)
Migrants to Uruguay		-1.25
		(0.80)
Migrants to Cuba		-7.50**
		(0.03)
All Canary migrants	-6.25**	
	(0.03)	
Time Dummies incl.	yes	yes
Observations	3,492	3,492
Pseudo R2	0.00685	0.00692

Notes: Reference category is a non migrant male, aged 33 to 72, born and living in Canary Islands. The dependent variable is numeracy. Time dummy variables are included as decades of birth. All coefficients are multiplied by 125. Robust P-Values in parentheses: ***p<0.01, **p<0.05,*p<0.10

Table 5.3: Numeracy of Canary Islanders in Cuba and Uruguay Relative to the Native Population in the Target Country

	(1)	(2)
Included samples	Cuba	Urguguay
Model	Logit mfx	Logit mfx
Female	-1.25	-0.00
	(0.58)	(0.71)
Age 23-32	12.5***	3.75
	(0.00)	(0.14)
Canaries	-10.00***	2.50
	(0.00)	(0.33)
Criollos	1.25	
	(0.53)	
Spain	-5.00***	-7.50**
	(0.01)	(0.03)
Europe	-3.75	-0.00
	(0.65)	(0.96)
Africa	-13.75***	-30.00***
	(0.00)	(0.00)
Asia	0.00	
	(0.93)	
South America	-3.75	-12.50***
	(0.67)	(0.00)
Other Nationality	-15.00*	-37.50***
	(0.09)	(0.00)
Portugal		-5.00
		(0.37)
Time Dummies incl.	yes	yes
Observations	15,389	7,399
Pseudo R2	0.0134	0.0156

Notes: Reference category in Model (1) are Cuban natives, reference category in Model (2) are Uruguayan natives. The dependent variable is numeracy. Time dummy variables are included as decades of birth. All coefficients are multiplied by 125.
Robust P-Values in parentheses: ***p<0.01, **p<0.05,*p<0.10

Table 5.4: Logistic Regressions of the Numeracy of Cuba, Uruguay and Argentina's Census Population on Migrant Status and non-European Migrant Origin

	(1)	(2)	(3)	(4)	(5)	(6)
Included samples	Cuba	Cuba	Uruguay	Uruguay	Argentina	Argentina
Model	Logit mfx	Logit mfx	Logit mfx	Logit mfx	Logit mfx	Logit mfx
Female	-2.50**	-1.25	-1.25	0.00	-3.75***	-3.75***
	(0.05)	(0.44)	(0.58)	(0.82)	(0.00)	(0.00)
Age 23 to 32	16.25***	11.25***	1.25	3.75	13.75***	13.75***
	(0.00)	(0.00)	(0.45)	(0.14)	(0.00)	(0.00)
Migrant status	-7.50***	-7.50***	-8.75***	-1.25	13.75***	13.75***
	(0.00)	(0.00)	(0.00)	(0.47)	(0.00)	(0.00)
African		-7.50***		-28.75***		
		(0.00)		(0.00)		
Asian		7.50***				
		(0.00)				
Latin American		3.75		-11.25***		10.00***
		(0.69)		(0.00)		(0.00)
Other nationality		-6.25		-37.50***		2.00***
		(0.44)		(0.00)		(0.00)
North America						21.25***
						(0.00)
Time Dummies incl.	yes	yes	yes	yes	no	no
Observations	15,390	15,390	7,399	7,399	45,493	45,493
Pseudo R2	0.00978	0.0129	0.0061	0.0149	0.0177	0.0184

Notes: Reference category in Models (1) and (2) are natives in Cuba, reference categroy in Models (3) and (4) are natives in Uruguay and reference category in Models (5) and (6) are natives in Argentina. In Models (2), (4) and (6) the "migrant status" dummy refers to European migrants. Time dummy variables are included as decades of birth. All coefficients are multiplied by 125. Robust P-Values in parentheses: ***p<0.01, **p<0.05,*p<0.10

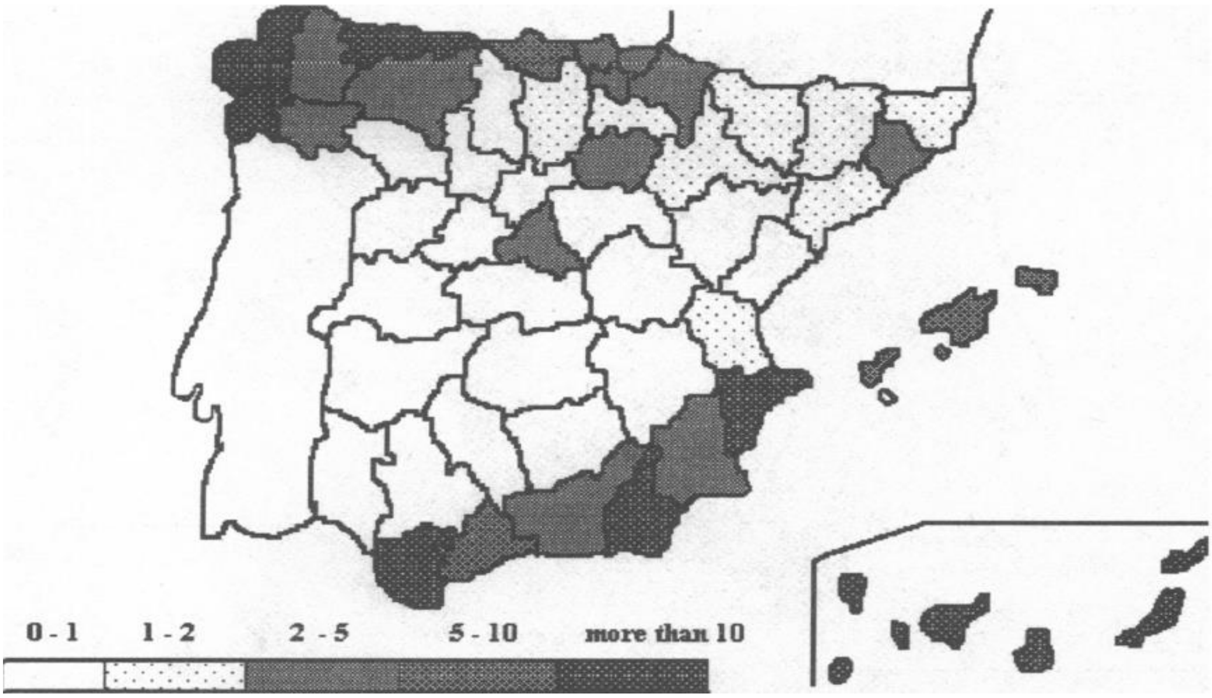
Table 5.5: Numeracy of Argentina's Population on Migrant Origin and Migrant Numeracy Advantage over time in Logistic Regressions

	(1)	(2)
Included samples	Argentina 1895	Argentina 1895
Model	Logit mfx	Logit mfx
Female	-3.75***	-3.75***
	(0.00)	(0.00)
Age 23-32	13.75***	
	(0.00)	
Spain	12.50***	
	(0.00)	
Italy	12.50***	
	(0.00)	
Europe	17.50***	
	(0.00)	
South America	5.00***	
	(0.00)	
Other Nationality	15.00***	
	(0.00)	
North America	30.00***	
	(0.00)	
Birth decade 1820-1829		-0.25***
		(0.00)
Birth decade 1830-1839		-0.19***
		(0.00)
Birth decade 1840-1849		-0.17***
		(0.00)
Birth decade 1850-1859		-0.09***
		(0.00)
Migrant*bdec 1820-1829		20.00***
		(0.00)
Migrant*bdec 1830-1839		18.75***
		(0.00)
Migrant*bdec 1840-1849		16.25***
		(0.00)
Migrant*bdec 1850-1859		12.5***
		(0.00)
Migrant*bdec 1860-1869		8.75***
		(0.00)
Observations	45,493	45,493
Pseudo R2	0.0187	0.0203

Notes: In Model (1) the reference category is an Argentinean native, 33 to 72 years old. Reference category in Model (2) is an Argentinean born native. All coefficients are multiplied by 125. Robust P-Values in parentheses: ***p<0.01, **p<0.05,*p<0.10

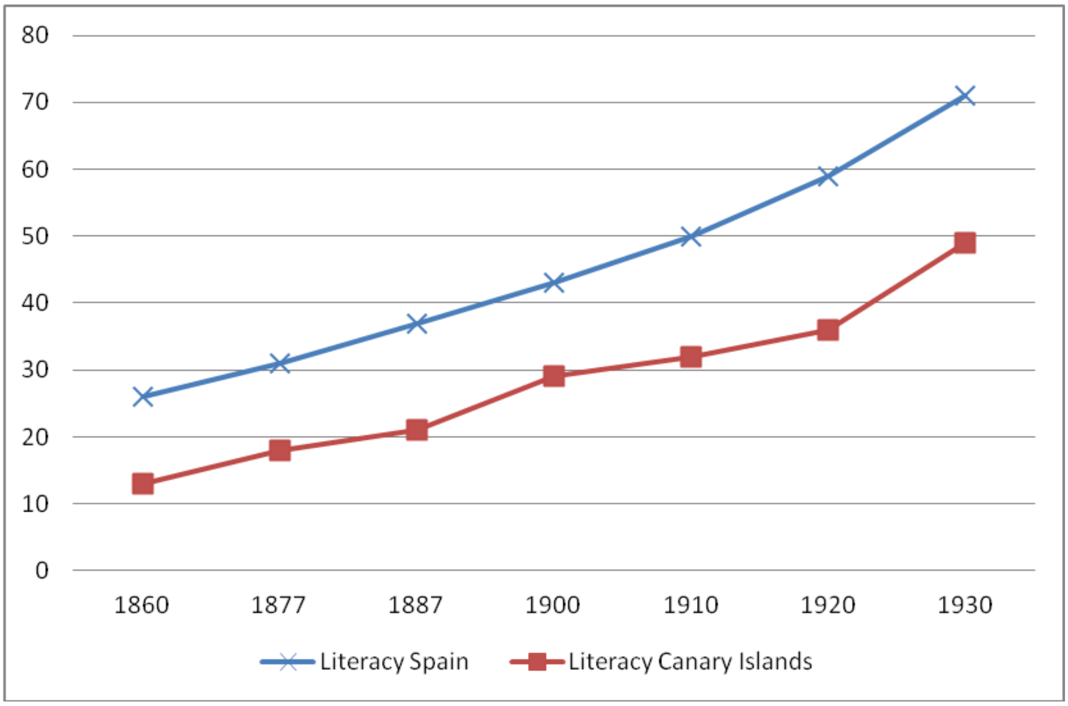
5.9 Figures

Figure 5.1: Gross Emigration Rates per Thousand Population in each Province, 1887



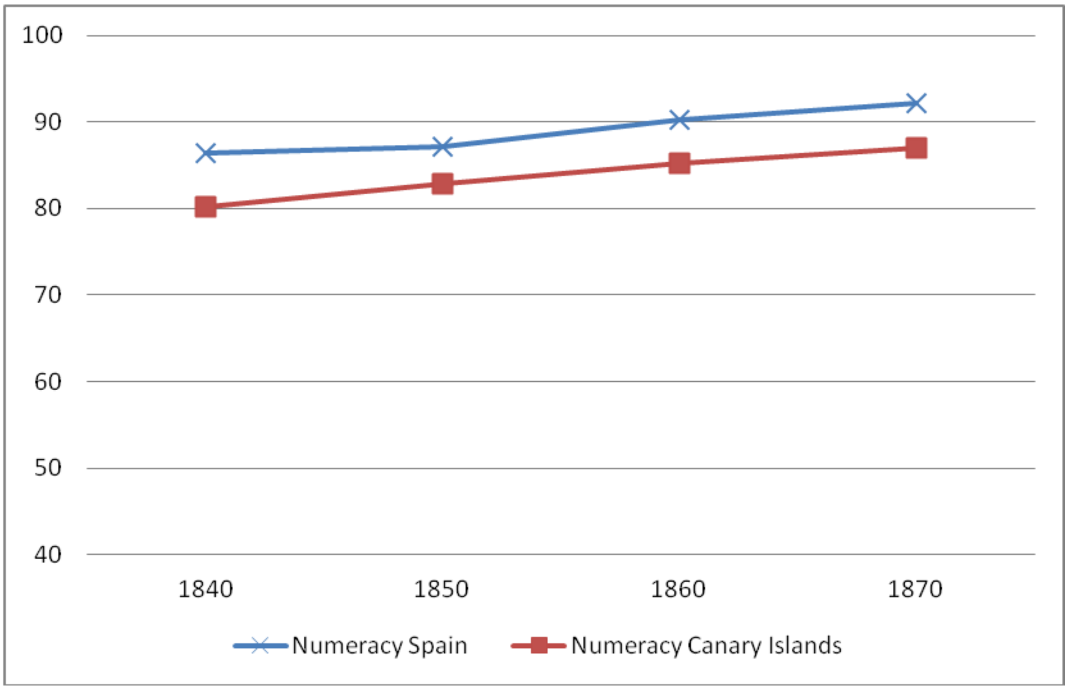
Source: Sánchez-Alonso (2000).

Figure 5.2: Literacy Rates in the Canary Islands and the Total Spanish Average



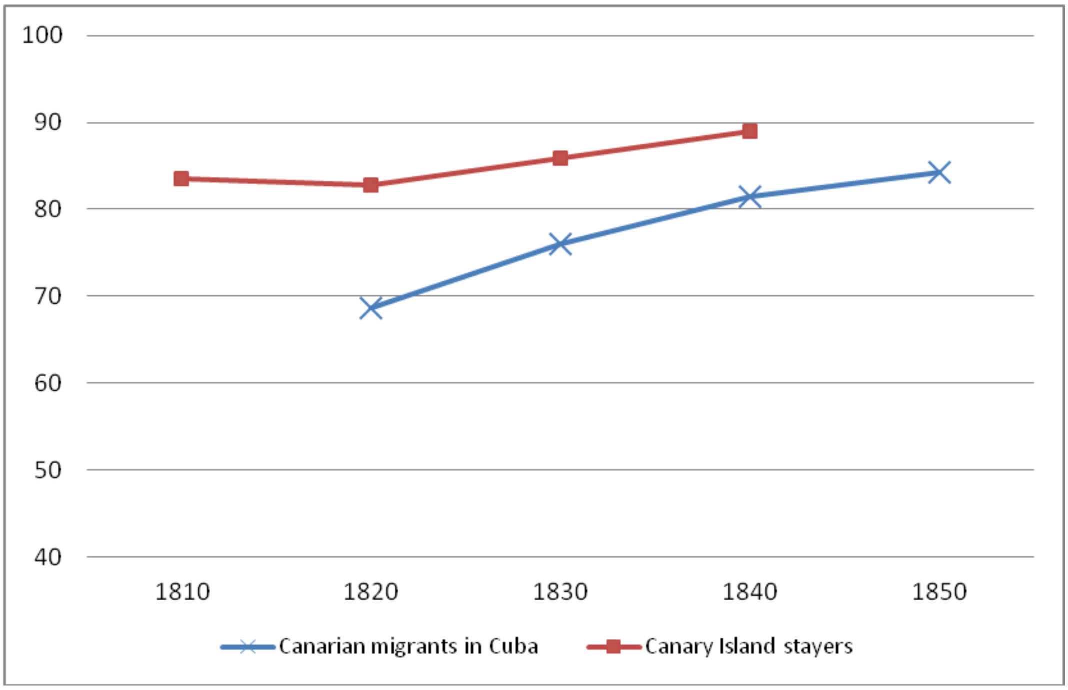
Source: Estadísticas Históricas de España Siglo XIX y XX (López 2005)

Figure 5.3: ABCC Index of Canary Islanders and the Total National



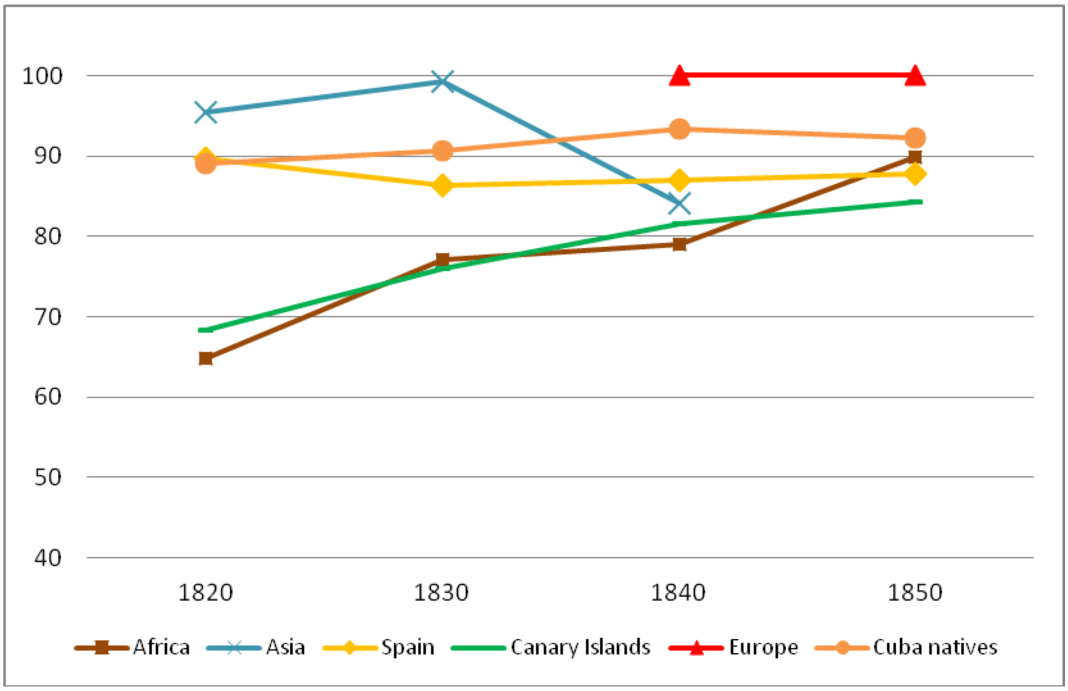
Source: Spanish National Census of 1900 (http://www.ine.es/inebaseweb)

Figure 5.4: ABCC Index of Canarian Migrants to Cuba



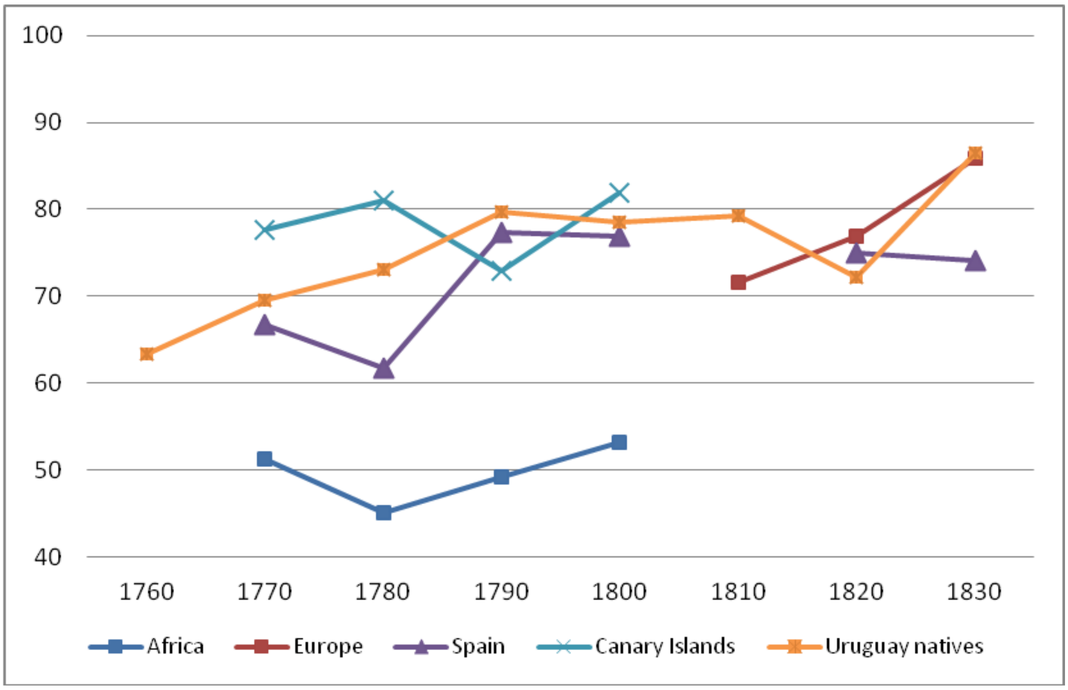
Sources: Spanish Census of 1877 and Cuba various censuses 1862 to 1887

Figure 5.5: ABCC Index by Birth Decades of Migrants to Cuba and Cuban Natives



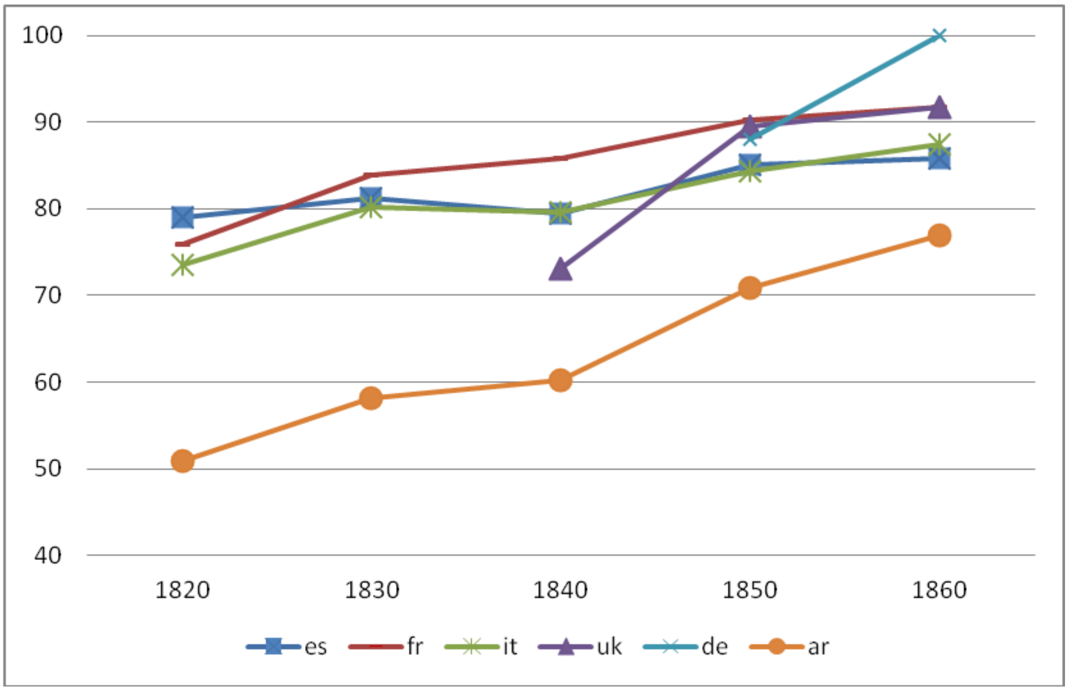
Source: Censuses of 1862 to 1887 provided by Linda Twrdeck

Figure 5.6: ABCC Index of Migrants to Uruguay and Uruguayan Natives by Birth Decades



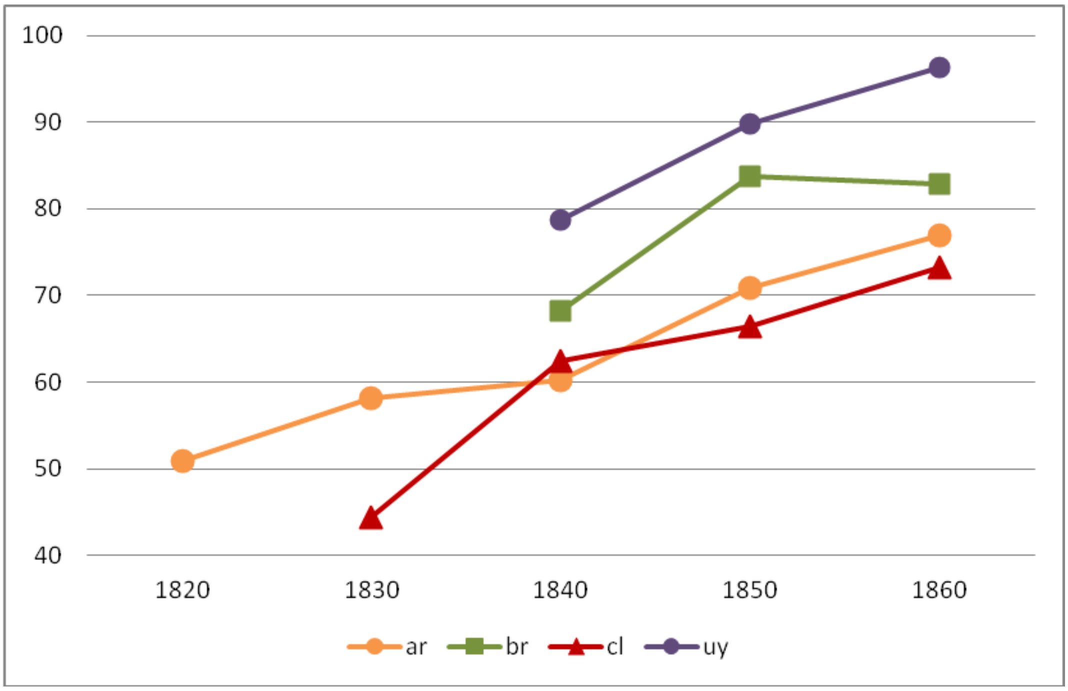
Source: Various censuses in Uruguay

Figure 5.7: ABCC Index by Birth Decades of Migrants from Europe to Argentina and Native Argentineans



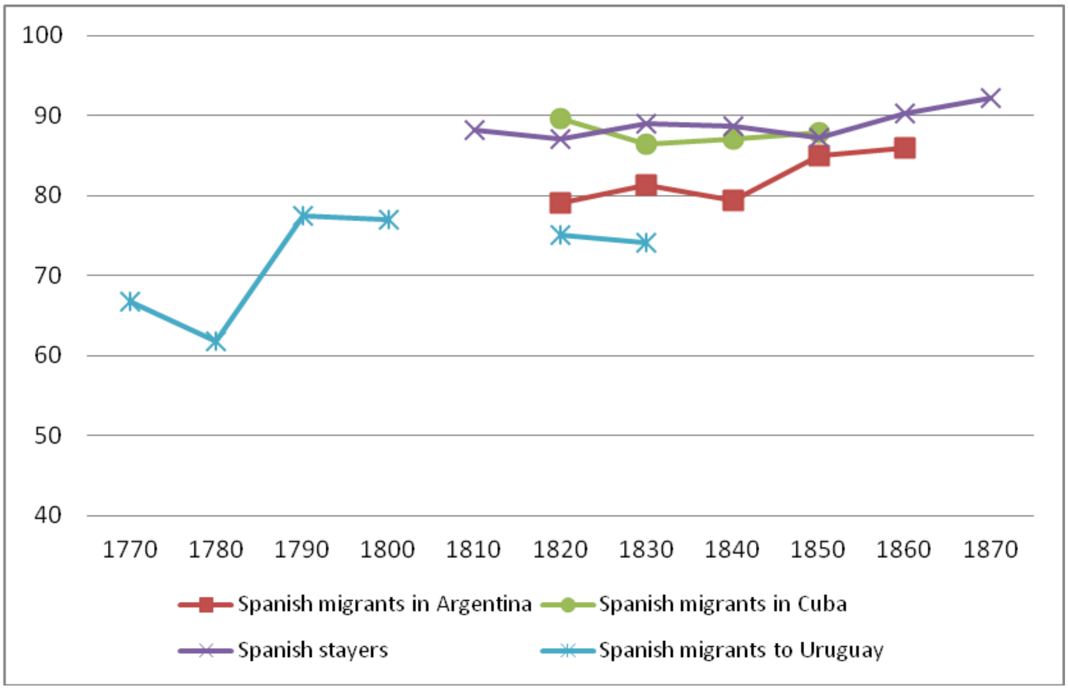
Source: National Census of Argentina 1895 (Somoza and Lattes 1967)

Figure 5.8: ABCC Index by Birth Decades of Migrants from South America to Argentina and Argeninean Natives



Source: National Census of Argentina 1895 (Somoza and Lattes 1967)

Figure 5.9: ABCC Index of Spanish Migrants to Cuba, Uruguay and Argentina



Source: census of Argentina 1895 (Somoza and Lattes 1967), various census of Cuba (1862 to 1887), Spanish census of 1877 and 1895

5.10 Appendix:

A.1. Summary Statistics: Migrant Origin in Uruguayan censuses

	Total population		Only migrants	
Origin	Obs. No	%	Obs. No.	%
Natives	4,856	64.21		
Africa	465	6.15	465	17.18
Latin America	457	6.04	457	16.88
Canary Islands	641	8.48	641	23.68
Europe (not Portugal or Spain)	523	6.92	523	19.32
Portugal	126	1.67	126	4.65
Spain (not Canary Islands)	476	6.29	476	17.58
Other nationality	19	0.25	19	0.70
Total	7,563	100.00	2,707	100.00

A.2. Summary Statistics: Migrant Origin in Argentina – Total National and City of Buenos Aires

	Argentina 1895				City of Buenos Aires 1895	
	Total population		Only migrants		Total population	
Origin	Obs. No.	%	Ob. No.	%	Ob. No.	%
Natives	26,419	58.07			1,792	22.26
Latin America	1,727	3.80	1,727	9.05	278	3.45
Northern Europe	2,810	6.18	2,810	14.73	963	11.96
Italy	9,431	20.73	9,431	49.44	3,277	40.70
Other nationality	1,156	2.54	1,156	6.06	247	3.07
North America	36	0.08	36	0.19	22	0.27
Spain	3,914	8.60	3,914	20.52	1,473	18.29
Total	45,493	100.00	19,074	100.00	8,052	100.00

A.3. Summary Statistics: Migrant Origin in Cuban censuses

	Total population		Only migrants	
Origin	Obs. No.	%	Obs. No.	%
Natives	6,388	44.20		
Africa	1,326	9.17	1,326	16.44
Latin America	17	0.12	17	0.21
Asia	3,872	26.79	3,872	48.01
Canary Islands	1,286	8.90	1,286	15.95
Northern Europe	48	0.33	48	0.60
Portugal	5	0.03	5	0.06
Spain	1,494	10.34	1,494	18.52
Other nationality	17	0.12	17	0.21
Total	14,453	100.00	8,065	100.00

A.4. Occupations of Canary Islanders in Uruguay

	Percentage	Total	In passenger list (percentage)	In censuses (percentage)
Occ. related to farming	67.6	259	51.0	97.8
Occ. not related to farming	32.4	124	49.0	2.2

A.5. Occupations of Canary Islanders in Cuba

	Percentage	Total
Occ. related to farming	81.0	128
Occ. not related to farming	19.0	30

6. A Story of Large Landowners and Math Skills: Inequality and Human Capital Formation in Long-Run Development, 1820-2000

Abstract:

We create a new dataset to test the influence of land inequality on long-run human capital formation in a global cross-country study and assess the importance of land inequality relative to income inequality. Our results show that early land inequality has a detrimental influence on math and science skills even a century later. We find that this influence is causal, using an instrumental variable (IV) approach with geological, climatic and other variables that are intrinsically exogenous. A second major contribution of our study is our assessment of the persistence of numerical cognitive skills, which are an important component of modern human capital measures. Early numeracy around 1820 is estimated using the age-heaping strategy. We argue that countries with early investments in numerical education entered a path-dependency of human capital-intensive industries, including skill-intensive agriculture and services. The combined long-run effects of land inequality and human capital path-dependence are assessed for the first time in this article.

This chapter is based on a paper published in *Journal of Comparative Economics*, co-authored with Jörg Baten (University of Tübingen). The concept of the paper was developed jointly, analyses and writing were equally shared.

6.1 Introduction

"How can they understand politics, if they do not even know where they are standing? They might vote for the communists, like the mining workers in the North…"

(Isabel Allende, 1982).[1]

In Isabel Allende's novel "La Casa de los Espíritus", the large landowner Esteban Trueba was sure that the day laborers on his estate were not capable of making wise political decisions and should therefore not acquire voting rights. However, he was most likely an exceptional landowner in Chile at the time, as he built a school for his workers during the early 20th century.

Recently, Galor, Moav and Vollrath (2009) (henceforth GMV) argued that two different elite groups strongly influenced political decision-making about educational reforms: large landowners and industrialists. "Latifundia" (large estate) owners were typically not interested in tax-financed mass-schooling programs that would remove unskilled day laborers from the agricultural workforce on their estates or make them rebellious. Industrial capitalists, in contrast, typically favored a workforce that had at least some basic skills. In countries that displayed a large amount of land inequality, such as some Latin American countries, the elite landowner group gained political power because their large estates provided financial and social backing to influence the ruling groups. In the process of industrialization, countries with a more equal distribution of land overtook countries with high inequality and performed better in terms of modern income growth. According to Wegenast (2009, 2010), plantation owners also had political reasons to neglect educational expansion, for instance, by monopolizing the decision-making process for literacy requirements for suffrage. GMV provided some evidence regarding regional

[1] Own English translation from the Spanish original version of "Ghost House".

land inequality and school investment in the U.S. from 1880-1940 as well as some qualitative evidence on land reforms in East Asia and Russia. Hippe and Baten (2012) recently tested and confirmed the link between land inequality and human capital development for European regions.

However, the implications of these theories have not been tested for a large cross-section of countries over the long run, partly because evidence of early inequality for poorer countries was scarce. We create a new dataset to test the influence of land inequality on long-run human capital formation. The methodological idea is to (1) use evidence of 20th century land inequality, (2) assess the influence of land reforms and (3) estimate late 19th century land inequality on the basis of backward projection.

In this study, we assess the influence of both within-country land and income inequality on human capital formation for the period from 1964 to 2003. As a measure of income inequality, we use a recently created global dataset by Van Zanden, Baten, Földvari and Van Leeuwen (2012) and Blum and Baten (2011). This dataset documents within-country income inequality, employing anthropometric inequality measures as well as income distribution (and GDP per capita divided by wages of unskilled workers, so-called "Williamson Ginis", and other indicators).

Which measure of human capital should be employed as a dependent variable? We argue that a measure should be used that is most conducive to economic growth. For example, school years have often been criticized because the productivity of a school hour differs between countries and cultures and, thus, is not a perfect growth predictor. Recently, the leading human capital economists Hanushek and Woessmann (2012a, 2012b) argued that cognitive skill test results related to math and science abilities are the strongest correlates of economic growth. They extended the famous PISA results from the 2000s into the period 1964-2003 by recalibrating a large number of international math and science tests; they also developed a comprehensive index of those core skills that will be our dependent variable.

A second major contribution of our study is our assessment of the long-run persistence of numerical cognitive skills, which are an important component of the Hanushek-Woessmann measure. We include a new measure of numeracy around 1820 that is constructed on the basis of age-heaping estimates. We argue that countries with early investments in numerical education (and perhaps cultures that promoted numerical skills) entered a path-dependency of human capital-intensive industries, including skill-intensive agriculture and services. Those countries that took the numerical lead (but not necessarily the lead in living standards) in the 1820s were situated in Scandinavia, Central and Western Europe, and East Asia. In the second wave, North American, Southern and East-Central European countries followed. Other world regions lagged behind.

The next section provides a literature overview on the determinants of human capital formation, such as institutional quality, geography, fertility choice and physical capital. In section 3, we introduce the data and sources. Section 4 and 5 describe the OLS and instrumental variable test regression results of our empirical study. Section 6 concludes.

6.2 Components of Human Capital Formation

A number of theories have been proposed to explain the great divergence in education that took place over the past two centuries and had a strong influence on income divergence. Some influential studies have emphasized the detrimental effect that early inequality had on the emergence of human capital and growth-promoting historical institutions (Engerman and Sokoloff 1997, 2000; Acemoglu et al. 2001, 2002).[2]

Sokoloff and Engerman (2000) hypothesized that in more unequal societies, the elites gained power to influence the choice of legal and economic institutions. In those countries that were unequal, a small number of elites restricted the rights of the majority of people, such as in education and voting, to perpetuate the existing social structures and maintain or even reinforce their elite status.[3]

GMV (2009) set up a theoretical model showing that an unequal land distribution in a country negatively affects per capita income in the long-run by delaying the implementation of human capital-promoting institutions. These authors argued that two different elite groups strongly influenced political decision-making about educational reforms: large landowners and industrialists. In contrast to “latifundia” owners – who were typically not interested in tax-financed mass-schooling programs – industrial capitalists preferred a workforce with at least some basic skills because education increased productivity in the industrial sector more than in agricultural production. Especially numerical skills can only be developed to a higher level if public funds are spent on well-educated math teachers. Rich landowners often did not send their own children to public schools. And even if they would have done that, they benefited less than the poor masses from taxes spent on schooling funds. As they paid the largest share of taxes in countries with high land inequality, their self interest stood against higher school spending on math

---

[2] Institutions are normally divided into growth-promoting and growth-retarding types of institutional set-ups. For example, if the risk of expropriating wealth is high due to “bad institutions”, incentives to invest are lacking that otherwise could have promoted economic and technological advances (North 1981, Acemoglu et al. 2001, 2002).

[3] According to Lindert (2004), the burden of educating laborers had more weight than its benefits (preventing crime and sedition) from the landlords’ point of view. The premise was that as educated workers sought better-paid jobs outside the agricultural sector, the political supremacy of the landed elite could be threatened, and their taxes would be raised to subsidize the masses.

teachers. For example, Cinnirella and Hornung (2011) describe the interesting example of East German (today Polish) regions in which the large land owners opposed increased tax-based spending on higher school investments (for the micro-founded theoretical model see GMV, p. 146 and 155). Thus, concentration of landownership hampered the emergence of institutions such as public schools and child labor regulations, as well as the process of industrialization. This was the case, for example, in some 19th century Latin American countries in which this elite group had particularly strong political power because large estates provided financial and social backing to influence the ruling groups, and the institutions were set up to maintain inequality. In contrast, countries with lower land inequality or countries that had installed land redistribution programs were more often governed by capitalist interests; hence, tax-funded mass-schooling was frequently extended. Galor et al. (2009) provided some evidence on regional land inequality and school investment in the U.S. from 1880-1940, as well as some qualitative evidence on land reforms in East Asia and Russia. Ramcharan (2010) assessed the impact of educational spending in U.S. counties and states from 1890 to 1930.

According to Wegenast (2009, 2010), plantation owners had economic and *political* reasons to neglect educational expansion, for instance, by impeding the mobilization of rural workers so that a cheap supply of labor was secured and by monopolizing political participation, given the literacy requirements for suffrage.

Another theory exploring this link is the "credit market imperfection approach", which holds that land inequality, usually implying the landlessness of many, results in credit rationing, in which most individuals cannot undertake investments in human capital (Galor and Zeira[4] 1993, Deininger and Squire 1998).[5] Supporting this view and adding a

[4] In their study, Galor and Zeira show that in a cross-country comparison, wealth and income inequality has a detrimental effect on subsequent economic growth by hampering investments in human capital.

labor market-power component, Deininger (2003) noted that unequal land distribution removes incentives of rural workers to accumulate human capital. When landlords have a monopsony over the labor market and keep all wages at subsistence levels, investing in human capital seems much less rewarding. As an example, the author described the growing gap in the development of literacy rates between 19th century El Salvador and Guatemala on the one hand and Colombia and Costa Rica on the other hand. While in El Salvador and Guatemala a few large landowners held a monopoly over the labor market and people had no incentive to invest in the accumulation of human capital, in Colombia and Costa Rica the coffee boom led to a small landholder coffee economy where literacy rates increased (Deininger 2003, p. 20).

Another part of the literature emphasizes the effect of land inequality on the access to the public goods of health and nutrition for the poor, because they observed empirically that land unequal areas performed worse in terms of public investments.[6] For instance, Banerjee and Iyer (2005) find that in India, areas where proprietary rights in land were historically given to landlords, today invest less in public infrastructure such as health centers and hospitals. As large land-owners paid a high share of taxes in those areas, but did not benefit as much from public hospitals and nutrition targeted education programs, they opposed spending of public funds on such programs. Health and the quality of nutrition – especially during the first years of life – are important determinants of cognitive abilities (and numeracy), and hence this could be an alternative channel of causal impact

---

[5] For an extensive literature review on how inequality affects growth, see also Galor (2011).

[6] One possible reason is that the conflict between stake-holders (here: landlords and the masses) could hamper unanimous policy decisions (on this body of literature, see also Bates 1981 and Easterly and Levine 1997).

(Currie 2009, Case and Paxson 2008). Using a natural experiment of a military blockade, Baten, Crayen and Voth (2013) recently showed the effect of health on numeracy.

Hippe and Baten (2012) tested and confirmed the link between land inequality and human capital formation by regressing the share of large landholdings (above 100 hectares) on numeracy levels for several hundred regions of ten European countries (see also Cinnirella and Hornung 2011 for Prussia). In contrast to those authors, we use a global sample of Gini coefficients of land inequality during the late 19th century to assess this theory both in industrialized countries and in developing world regions.[7]

Various exogenous factors have been suggested to determine inequality, including initial endowments of geographical conditions, soil quality and surplus labor. Engerman and Sokoloff (1997, 2000, 2002) argued that countries in tropical zones that are suited for the cultivation of cash crops (such as sugar cane, tobacco, coffee, rubber or cocoa) were more likely to develop into unequal societies than countries in temperate zones, which were more suitable for the production of food crops such as wheat, maize or rice, because the production of cash crops was subject to substantial economies of scale (as substantial fixed costs were unavoidable, but fell per unit of production at higher scale). Thus, cash crop production favored large-scale plantations and attracted large inflows of mainly slave labor. Geography has also been alleged to have had a direct productivity-retarding effect through climate and climate-induced diseases in some world regions (Montesquieu 1748, Jones 1981, Sachs and Warner 1997, Diamond 1997, Sachs 2001).[8]

Fertility choice is another potential determinant of human capital accumulation, especially during the 20th century: The famous quantity-quality trade-off might determine

[7] Easterly (2007) uses agricultural endowments as an instrument for income inequality as well, but not specifically for land distribution.
[8] However, Easterly and Levine (2003) show in a cross-country study that climatic conditions ("tropics, germs and crops") affect growth only through institutions and not directly.

the level of educational investment because if a family decides to have more children, the human capital investment per child might be lower (see Becker 1960, Easterlin 1980). However, Clark (2007) recently presented historical evidence that this result does not necessarily hold for early societies in which richer individuals had more surviving children (Galor and Moav (2002) proposed a theory of natural selection earlier than Clark).

Human capital clearly has endogenous as well as exogenous components. Exogenous factors (not influenced by economic variables) include, for example, religious educational values (Botticini and Eckstein 2007, Becker and Woessmann 2009). Selective migration (in terms of skills) can affect the human capital development of a country, both positively and negatively (Glaeser et al. 2004). For example, when we speak of a "brain drain" phenomenon, the home country's average human capital is negatively affected by emigration of high skilled workers, while the host country benefits from positively selected immigration (and sometimes vice versa).[9]

Our study contributes to this research stream by providing empirical evidence on the causal link between inequality more than one hundred years ago and recent human capital levels for a large cross-section of countries.

6.3 Data

The data on income inequality was recently compiled by Van Zanden, Baten, Földvari and Van Leeuwen (2012) and Blum and Baten (2011). This global dataset documents within-country inequality for a large number of countries, employing anthropometric inequality measures as well as income distributions. This effort allows us to cover income inequality in a substantial number of countries (see also Data Appendix D).

[9] Stolz and Baten (2012) find the opposite – arithmetic "brain gain" for the home countries which "lost" many their less skilled inhabitants -- for a number of countries during the 19th century era of mass migration.

How can we estimate land inequality? Unfortunately, the number of available estimates of land inequality for the 19th century is much more limited than even the number of income Ginis. Frankema (2010) presented 199 Gini coefficients of inequality in land holdings[10] for 111 countries in the 20th century. His tables are based on the World Census of Agriculture published by the International Institute of Agriculture (IIA) and its subsequent incarnation, the Food and Agriculture Organization (FAO). Before the 1920s, only 27 land Ginis are available, and for the 1960s, 63 are available. In examining those countries for which two or more land Ginis are available, it becomes clear how little land inequality has changed, especially in developing and threshold economies. For example, Argentina's Gini coefficient of land inequality was 0.803 in the 1910s. In the 1940s, it was 0.806, and in the 1960s and 1980s, it was around 0.81.

Over nearly a century, there was barely any change. The same applies to Brazil in the 1920s through the 1960s, despite land reform initiatives during the 1930s that did not have much effect. In Chile, there was a small increase from an already high inequality level of 83.7 in the 1920s to 86.5 in the 1960s, and the results are similar in Uruguay (77.5 to 79.1 in the 1930s and 1960s). Because we know that all inequality measures contain some measurement error, the hypothesis of no change can most likely not be rejected. The stability of land inequality can be observed for a large number of countries. Given the overwhelming stability of land inequality values, we estimate a cross-sectional dataset for the late 19th century by assuming that the earliest value of land inequality is a good proxy for the initial conditions in each country if there was no land reform.[11]

---

[10] The size distribution of the land holdings reflects the access to land as a production factor, rather than only ownership. It is usually more equally distributed than land ownerships. However, this data includes land holdings used for all purposes – arable land, land used for permanent crops, pastures or woods – and do not take into consideration the quality or the location of the land (Frankema 2010).

[11] Our *early land inequality* variable contains the land Gini values of the earliest year available for each country from Frankema's (2010) dataset.

While land inequality typically did not change much over several decades, change in land inequality is likely if a land reform took place in the early 20th century.[12] Because we are concerned that land reforms implemented between 1890 and the earliest year for which we have information on land concentration could have had a substantial effect on land distribution, we need to control for the average effect of land reform on the land Gini coefficient. To test whether land reforms might explain the change in land inequality significantly, we collected evidence on the implementation and success of land reforms (listed in Appendix B). With this information, we assessed the average reduction effect of a land reform on the change of the Gini coefficient of land concentration in a regression (Table 6.1).[13] The coefficient of the explanatory variable 'land reform' (a dummy variable that takes the value of 1 if a land reform was carried out in the previous period) in both specifications implies that land reform will reduce the Gini coefficient in the following period by six Gini points. We estimated a second specification in which we controlled for "Western Europe and offshoots", a variable that includes developed economies in which land concentration declined during the twentieth century, typically without land reforms. The results were robust when including this additional variable. According to the results, we correct our values of early land Ginis – using those that are the closest to 1890 contained in the Frankema data set – by adding six points when land reforms were performed successfully. This tactic allows us to back-project early land inequality estimates to the level around 1890. In Appendix C, we list the sources from which we derive information on land reforms for the various countries. Of course, the ideal variable would have measured the extent of land reforms, which our binary variable does not

[12] One could imagine that also a change in crop production might alter the distribution of land holdings. However, for the time period and the countries observed here, crop production hardly changed. All Latin American countries, for example, had already specialized in cash crop production.
[13] We included only those countries for which Gini coefficients of land holdings for more than one point in time were available.

measure. However, adjusting for the existence of land reforms already improves the measurement of our land inequality variable substantially.

We should consider the question, whether changes in population or in crop production did impact on the change in land inequality, especially between the 1890s and the 1950/60s. Clearly, substantial changes in the crop composition or the share of land used for livestock production could have had an influence on landholding patterns. Especially a switch from small-scale subsistence farming to cash crop commodity production in large plantations could lead to higher land inequality. Expropriation of native small land holders to introduce large scale primary goods production was a typical effect of the colonial 'land-grab'. For example, in Malaysia and Sri Lanka colonialists produced rubber on a large scale, a cause of the high land Gini coefficient in the second half of the 20th century (Frankema 2010). Moreover, the adoption of cattle-raising activities for leather production or beef export could have led to a higher concentration of land holdings. We will thus have a careful look at the changes in the composition of crops grown and the time of adoption of cattle-raising activities in former colonies included in our data set. We will then draw the link to the evolution of land inequality.

In Latin America already the Spanish colonizers of the 16th to 18th century introduced cash crops and livestock production on large farms (Bulmer-Thomas 1994, p. 9). This was the case for the cultivation of sugar in Cuba, the Caribbean and Brazil, which began in the 16th century. Production and export of tropical crops such as coffee[14], cocoa, tobacco and bananas boomed in Latin America since the middle of the 19th century, to satisfy a rising demand in Europe and North America after the Industrial Revolution (Bulmer-Thomas 1994, p. 36). Even rubber production in Brazil, a commodity that was

[14] However, coffee seems to be an exception of a crop that is more effective to be planted in small plantations. There seem to be no scale economies and cultivation is labor intensive, thus more profitable for smallholders (Nugent and Robinson 2010)

exploited later in Asia, reached its peak already in the middle of the 19th century (Resor 1977). Although small variations in the composition of the cash crop cultivation may have occurred later, mostly landholding patterns were already established by 1890. Hence the Latin American development of the early 20th century does not affect our back-projection of land inequality much.

What about African and Asian colonies? In some more recent colonies, the British rule between the end of the 19th century and the first half of the 20th century might have had an increasing impact on land inequality. Frankema and van Waijenburg (2012) described how British colonies in sub-Saharan Africa ("non-settler colonies") mostly adopted the existing land tenure system. However, commodities were cultivated by small native landholders before and after colonization. In French colonies of the same region land expropriations were also rare. In other sub-Saharan colonies, such as Sierra Leone, the native population did not switch to commercial agriculture. Land continued to be held by many smallholding subsistent farmers (Frankema 2010). Yet in South Africa – a "settler colony" – the native population was expropriated in large scale and land was redistributed to the white population at the turn of the century, resulting in a more unequal land distribution (Binswanger and Deininger 1993). The situation was similar in Malaysia, where British colonizers introduced large rubber plantations at the end of the 19th century and land inequality probably intensified (Frankema 2010). In the regression analyses below, however, South Africa, Malaysia and Sri Lanka do not play a role because early numeracy could not be estimated. One missing value in one variable causes the observation to be dropped. In sum, crop change probably influenced our back projection of land inequality mildly but not substantially. The same is probably true for immigration and population change.[15]

Comparing land and income inequality around 1890, we find that a general correlation exists (Figure 6.1). Chile and Peru ("cl" and "pe") both had high land and income inequalities, whereas land and income inequalities were low in Finland and Estonia ("fi" and "ee"). On the other hand, some countries had high land inequality but not high income inequality (for example, Egypt, "eg", and Australia, "au"), whereas other countries deviated to the other side of the regression line. An example of the latter case of low land inequality and elevated income inequality is Sweden in 1890. In this period, the country experienced its industrialization boom, which temporarily shifted income inequality. Inequality was not structural and permanent in Sweden and moved downward later. Looking at modern math and science skill levels, Egypt (high land and low income inequality) had low values, whereas Australia was characterized by high values. Therefore, we need to perform a regression analysis to cope with potential outliers.

Our equation for human capital determinants also includes the ABCC-Index values for 1820 – the earliest year for which we obtained a sufficiently large set of country data – to assess the path-dependence of education reflected in numeracy skills (see also Appendix E). A considerable number of recent studies have used this innovative proxy of basic numeracy, which was introduced by A'Hearn, Crayen and Baten (2009) and is based on

---

[15] Did changes of population over time, for example due to high immigration, affect land inequality? We find that land inequality was not substantially affected by migration. First, a large share of migrants went to urban centers (for example in Argentina) or to work as employees in already existent crop plantations. If immigrants arriving had no access to land in the destination country, this population increase would not affect our land inequality measure, because it does not account for the landless population (only for land holders). Second, those immigrants who were brought to settle in rural communities as small landowners in Chile and Argentina obtained land pieces of middle size, and overall distribution was not affected. Argentina was one of the most immigration intensive countries between 1870 and 1930. However, the early land Gini measures for Argentina by Frankema (2010) show that land inequality remained stable. Summerhill (2010) provides further evidence for an unchanged land Gini in Sao Paolo state (Brazil) between 1905 and 1995 (0.688 for 1905 and 0.677 for 1995).

the "age-heaping" technique. The age-heaping phenomenon applies to historical populations (as well as people in the poorest countries today) when a substantial share of the people are not able to state their exact age and therefore reported a rounded age, such as "I am 30", when they are in fact 29 or 31. The ABCC-Index reflects the share of people who were able to state an exact age (or the degree to which the distribution of age statements approaches an equal distribution). Crayen and Baten (2010) showed that this proxy of early human capital has a strong correlation with other measures such as literacy and schooling. The calculation of the ABCC Index is shown here as a derivation of the Whipple Index:[16]

$$(1)\ Wh = \left(\frac{(Age25 + Age30 + Age35 + ... + Age60)}{1/5 * (Age23 + Age24 + Age25 + ... + Age62)} \right) \times 100$$

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

As mentioned in the introduction, we argue that countries that made early investments in numerical education (and perhaps cultures that promoted numerical skills) entered a path-dependency of human capital-intensive industries, including skill-intensive agriculture (such as dairy farming in Denmark) and services (Figure 6.2, Appendix H). After these skill-intensive industries became important during the second Industrial Revolution of the 19th century, the following generations again invested in skill formation, both through formal schooling and cognitive training within families. For example, they used basic numerical training devices such as calendars and games that required basic numerical skills (Baten et al. 2009).

---

[16] A'Hearn, Baten and Crayen (2009) found that this index is the only one that fulfils the desired properties of scale independence (a linear response to the degree of heaping), and that it ranks samples with different degrees of heaping reliably. For an easier interpretation, A'Hearn et al. (2009) introduced the ABCC index. It is a simple linear transformation of the Whipple index and yields an estimate of the share of individuals who correctly report their age

Crayen and Baten (2010) and Manzel et al. (2012) – as well as recently Hippe and Baten (2012) – have already carefully discussed the evidence on early numeracy. However, to evaluate its contribution to modern differences of math and science skills, a short summary of the institutional sources of the evidence and the discussion of the degree of (sometimes not entirely avoidable) bias can be quite helpful. For Western Europe, Crayen and Baten could rely on census material which reported the age distribution repeatedly in many different censuses. For the UK and the United States, even individual level census data was available. For Eastern Europe, they could rely on the Russian Imperial Census of 1897, which was taken in quite disaggregated regional units, and hence could be re-aggregated in order to make the evidence comparable with modern data and math and science skills which are organized by modern state boundaries. In other words, for the numeracy of Armenia in 1820 only those districts were included which fall into the modern boundaries of Armenia. Similar procedures could be applied to the Austro-Hungarian (Habsburg) Empire evidence.

For Asia and Africa, the basic sources of evidence were more challenging. For example, for China, Baten et al. (2010) used a large variety of different data sets in order to arrive at the most likely estimate for the early 19th century. For Southeast Asia, census records were unavailable before the 1870s and had to be replaced by court records (which required adjustments due to the fact that persons appearing in court mostly for property crimes are not representative for the whole population). In contrast, for India, census material of quite high quality was available thanks to the intensive colonial administration of the British. The evidence on Latin America for the 19th century is quite comprehensive, and Manzel et al (2012) took care that any unrepresentativeness of census collection activity was taken into account (see Appendix G on the sources of the ABCC 1820 estimates in detail).

Figure 6.3 illustrates the strong correlation between the 1820s measure of numerical skills and the Hanushek-Woessmann measure for the late 20th century. The dependent variable *cognitive* is our measure for the cognitive skills of human capital. This output measure of human capital reflects the knowledge and abilities that are most favorable for subsequent success. Hanushek and Woessmann (2012a) constructed this variable from test scores in mathematics and science for 77 countries between 1964 and 2003. More specifically, they calculated the simple average of all math and science scores on International Student Achievement Tests conducted by the Organization for Economic Co-Operation and Development (OECD) and the International Association for the Evaluation of Educational Achievement (IEA) during that period. With linear regressions, they found that by adding cognitive skills to a growth model with school attainment as a dependent variable, the model explains three quarters of the variance in growth rates (instead of one quarter if only school years are included). Further, the coefficient for school years turns statistically insignificant in the presence of cognitive skills. When testing the correlation between test score improvements and growth rates by world regions – conditional on the initial level of real GDP per capita in 1960 – they find an R^2 of 0.98 (see Figure 6.4, albeit with N=5). To control for possible endogeneity of cognitive skills, which might be present if the factors leading to growth are also related to high cognitive skills and have been omitted from the estimation, they instrument cognitive skills with institutional measures of schooling and confirm that schooling-induced differences in cognitive skills are significantly related to economic growth.

This measure of human capital is said to best explain economic growth because, unlike other measurements such as schooling, it controls for differences in the quality of education across countries. A second test for the impact of the cognitive skills measure on economic performance was carried out in a case study on US-migrants. The authors looked at the performance in the labor market of US-migrants educated in the US and US-migrants educated at home, holding constant cultural and other country-of-origin fixed effects. The study again confirmed the growth-inducing effect of the *cognitive skills* measurement.[17]

Further variables included in our dataset are as follows: initial GDP in 1910, population density (logarithm), a measure of institutional quality, fertility in 1950 (the earliest estimates available for a sufficient number of countries), the share of the population living in the tropical zone in 1995, ethnic fractionalization, and a measure of physical capital that could proxy industrial development, constructed by Enflo and Baten (2006). Population density was included because visiting schools is less costly in countries with high population density.

6.4 Base Regression Results

Our model for the estimation of human capital has the following form:

$$\textit{Cognitive}_i = \alpha + \beta_1 \textit{ land ineq}_i + \beta_2 \textit{ income ineq}_i + \beta_3 \textit{ early ABCC}_i + \gamma X_i + \varepsilon_i$$

The dependent variable *cognitive* is our measure for cognitive skills. This output measure of human capital best reflects the knowledge and abilities that are favorable for subsequent success. Hanushek and Woessmann (2012a) constructed this variable from test scores in mathematics and science for 50 countries between 1964 and 2003 (Figure 6.5).

We test both income inequality around 1890 and early land inequality as determinants of cognitive skills. We add numerical abilities ("early ABCC") to assess the persistence of numerical skills. This measure also proxies to a certain extent the GMV factor of industrialists promoting basic education because countries that were numerate in 1820 also tended to be industrial.

[17] The finding that numerical cognitive skills are most conductive to economic growth further corroborates the argument of Schumpeter for the use of "numeracy" as an important proxy of human capital (see also Baten 2010).

In the econometric specification above, the additional vector of explanatory variables *X* captures other factors that could have an effect on human capital for all countries *i*: initial GDP in 1910, population density (logarithm), institutional quality, fertility in 1950, the share of the population living in the tropical zone in 1995, ethnic fractionalization and physical capital in 1925. Table 6.2 provides an overview of the summary statistics for the variables included in the model. The Skewness/Kurtosis tests for normality indicate that physical capital in 1925 and population density were right skewed; hence, we chose to take the natural logarithm of those variables.

Based on the basic model of human capital introduced in the last section, we perform cross-country analyses of the effect of inequality and early human capital on cognitive skills (Table 6.3). A large part of the variation in cognitive skills in the period 1964-2003 – as much as 54 percent – can be explained by early numeracy from the ABCC-Index in 1820 (see specification (1)).

In specifications (2) to (7), we assess the additional long-run impact of the Gini coefficient of early land inequality in OLS regressions for all the countries for which we had data. Land inequality always has a significantly negative impact on the dependent variable. When examining the coefficients, we can observe that a one standard deviation increase in land concentration leads to a 0.35 standard deviation decrease in human capital. We can best illustrate the impact of the land Gini coefficient with an example: a one standard deviation increase in land inequality reduces the level of cognitive skills from Serbia and Montenegro to the level of India. The same is true if we instead include income inequality (columns 5 and 7). In the case of income inequality, a one standard deviation increase of the Gini leads to a 0.26 to 0.31 standard deviation decrease of cognitive skills.

We also include a large set of other variables that could exert an influence besides inequality and early human capital. First, we control for early income (GDP in 1910) because the differences in early human capital could arise from the fact that richer economies spend more on education. However, all except one of the income coefficients are insignificant, and they do not lead to insignificance of early numeracy or inequality. We further insert controls for institutions, geography, population density, fertility, ethnic fractionalization and physical capital, but these variables are often insignificant (except tropic share).

Moreover, if we run a horse race between the two concepts of inequality, land inequality has the greater explanatory power (columns 12-14). Land inequality remains significant, whereas income inequality is no longer statistically significant. This result is consistent with Easterly's (2007) argument that land inequality is an important component of "structural inequality" because income inequality can temporarily rise, especially in growth and transformation processes. Structural inequality is mostly caused by the long-run effect of colonial land grabbing. Structural inequality has a much more depressing effect on human capital and development because educational efforts may not be rewarded in a situation of permanent, structural inequality.

The negative effect of land inequality, with other factors held equal, ranges from minus 0.81 to minus 1.60 and remains strongly significant through all specifications. These results confirm the GMV theory of a persistent negative effect of land inequality and early numeracy on the average quality of education during the second half of the 20th century.

In all specifications of Table 6.3 we include robust standard errors in order to cope with heteroskedasticity. We decided to cluster the standard errors by region (using the `stata`-option, `vce(cluster)`), given that standard errors could be correlated within world regions.

We additionally use alternative techniques like regressions with weighted least squares by population numbers in the countries assessed (columns 10 and 11) and robust against outliers together with population weights (columns 8 and 9).[18] Assigning a greater weight to more populated countries does not alter the significance of the coefficients for land inequality and their sizes stay almost unchanged. However, when running a robust regression (which gives less weight to outliers), the estimated effect of land inequality remains statistically significant, but the coefficient of income inequality turns insignificant. As a further robustness check, we run all specifications shown in Table 6.3, this time always using the same set of countries (number of observations equal to 30). The results are very similar and are shown in the Table in Appendix I.

6.5 Instrumental Variable Models

Next, we need to consider endogeneity. The results of the ordinary least squares regressions could be affected by reverse causality. For example, apart from the direction of causation running from the inequality of land via the political economy of landlords who oppose primary schooling (and the tax burden that comes with it), one can also imagine that in the long run, regions with relatively good education even for small landholders could reach a lower level of inequality of land distribution as those peasants are able to buy more and more land. The peasants might also influence political activity in favor of land reforms, as Cinnirella and Hornung (2011) have noted for the historical German kingdom of Prussia. On the other hand, educated small landholders might decide to sell their plots to obtain the return on their human capital investment in other industries – in nearby cities, for example. Instrumental variable estimation allows us to circumvent these issues of endogeneity.

[18] For running regressions with weighted least squares by population, we use the Stata option `[aweight=popsqrt]`. We used the R-statistics- package command `lmRob` to compute robust regressions and the option `weights=popsqrt` to include weights by countries' population.

We base our first stage of the two stage-least-square estimate on the following equation:

$$Landineq_i = \beta_1 + \beta_2\ Sugar/wheat\text{-}rice_i + \beta_3\ lowpop1500\text{-}Southern_i + \beta_4\ X_i + \varepsilon_i$$

where *sugar/wheat-rice* is an Easterly-type instrumental variable of relative soil suitability, *lowpop1500 –Southern* is the interaction of inverse population density in 1500 with "Southerness" of latitude and *X* is a vector of other exogenous variables.

Easterly (2007) and GMV recently advocated for the use of climatic, geological and similar variables that allow types of agriculture that correlate either with higher or lower efficient sizes of scale. Sugar plantation and cattle-raising for hide and beef exports are clear examples of agricultural production types in large-scale economies. On the other hand, wheat and rice production is already highly productive on much smaller farm units, as has been amply demonstrated in the agricultural economics literature. The specialization of a country on the cultivation of large scale cash crops is negatively associated to the distribution of landholdings, whereas food crops such as wheat or rice are not scale intensive and were historically planted in smallholdings. Sugar, wheat and rice production requires relatively clear-cut climatic and soil characteristics. Based on this premise, the UN- Food and Agriculture Organization quantified the share of a country's area that is suitable for the production of each of those crops. In the spirit of Easterly's (2007) we use the ratio of the share of the land suitable for the cultivation of the "inequality crop" (sugar) to the share of the land suitable for the cultivation of the "equality crops" (wheat and rice).

The cattle-raising activity associated with inequality is unfortunately less concentrated on specific soils (Figure 6.6). However, there is a clear correlation with southern locations and initially low population density. Argentina became the prototype of this type of land use because the indigenous Indian population on its great plains was always very thin, and in contrast to the United States, European immigration was relatively limited in numbers until the late 19th century. Therefore, the U.S. type of grain farming associated with lower inequality developed on a smaller scale. Thus, we instrument this type of agricultural inequality in southern latitudes in interaction with population density in 1500. In Figure 6.6, we show the level of land inequality around 1890. It is clear that the sparsely populated countries in the most southern locations showed the largest land inequality. Apart from Argentina and Australia, South Africa, New Zealand and Uruguay also had very unequal distributions following the colonial land grab (Eastwood et al. 2010). In contrast, the equality crops of wheat and rice led to low land inequality in East Asia and some of the European countries.

The advantage of the ratio between the climatic and geological suitability ratio of sugar and wheat/rice is its intrinsically exogenous nature, whereas actual crop use could be influenced by educational levels. Similarly, population density around 1500 is a very popular instrumental variable because it captures human development in a very early time period (Acemoglu et al. 2002). The autocorrelation of early population densities and those of the 20th century is quite limited due to the unequal population increase in some world regions and massive migration movements. Similarly, southern latitude is most likely very exogenous.

The results of the two-stage-least-squares regressions confirm that both variables fulfill the necessary requirements to be good measures of land inequality: First, they mostly correlate with land inequality, as is documented by the "first stage" section of Table 6.4. The F-Test is mostly below 10, requiring a Limited Information Maximum Likelihood (LIML) estimation, but it is still in the range of 5, which means that the instruments have some strength (see Stock and Yogo 2005). In the case of Model 1, in which we use both instrumental variables, the F-Statistic is larger than 10. Hence, this model should be preferred from the criterion "strength of instruments". However, because the instrumental variable component of "Southerness" (southern latitude) might be correlated with other factors, we find it important to show that the suitability variables alone yield similar results in Models (2) to (6). Second, we argue that the instruments influence the dependent variable only through the potentially endogenous variable, land inequality (see Figure 6.6, Table 6.5 and the discussion on the exclusion restriction in Appendix F). The p-value of the Durbin-Wu-Hausman test indicates only once (column 2) that the IV approach is necessary, as the results are significantly different from OLS estimates. As a result, the significant impact of land inequality and early numeracy remains a consistent determinant of today's human capital. Most of the other controls are insignificant, except for early numeracy, which is always positive and significant, and in one specification the coefficient of income per capita in 1910 becomes significant. In Appendix A, we show the results of income inequality as a determinant of human capital in the long run. When using the same instruments for income inequality, TSLS regression results become far less robust and usually insignificant.

6.6 Conclusion

The combined long-run effects of land inequality and human capital path-dependence were assessed in this article for the first time. We found that initial income inequality in a cross-section of countries might play a major role in determining the development of human capital formation in the long run on the first view. However, a more important predictor of recent human capital formation was early land concentration. This finding supports the GMV (2009) theory that the behavior of landowners is crucial to development relative to the influence of industrialists. The industrialist component of the model was to some extent proxied by early numeracy because the top performers in early numeracy were also quite industrial. We tested the robustness of our results by including various control variables and running a series of alternative regression models. We find that this influence is actually causal, using an Instrumental Variable (IV) approach with geological, climatic and other variables that are intrinsically exogenous.

These results are consistent with Easterly's (2007) argument that land inequality is an important component of "structural inequality" because income inequality can temporarily rise, especially in growth and transformation processes. Structural inequality is mostly caused by the long-run effects of colonial land grabbing. Structural inequality has a much more depressing effect on human capital and development because it seems that educational efforts might not be rewarded in a situation of permanent, structural inequality.

A second major contribution of our study was to assess the persistence of numerical cognitive skills, which are an important component of the Hanushek-Woessmann measurement. Our results confirm that countries with early investments in numerical education (and perhaps cultures that promote numerical skills) entered a path-dependency of human capital-intensive industries, including skill-intensive agriculture and industries. Early numeracy explained a considerable portion of recent cognitive skills.

6.7 References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91 (5): 1369-1401.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, 117(4): 1231-1294.

A'Hearn, B., Baten, J., and Crayen, D. (2009). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. *Journal of Economic History*, 68 (3): 783‑808.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2): 155-194.

Allende, I. (1982). *La Casa de los Espiritus*. Barcelona: Plaza and Janés.

Auty, R. M. (1993). *Sustaining Development in Mineral Economies: The Resource Curse Thesis*. London: Routledge.

Banerjee, A., and Iyer, L. (2005): History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *American Economic Review*, 95(4): 1190-1213.

Barro, R. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5(1): 5-32.

Baten, J. (1999). *Ernaehrung und wirtschaftliche Entwicklung in Bayern 1730-1880*. Stuttgart: Steiner.

Baten, J., Földvari, P., van Leeuwen, B., and van Zanden, J. L. (2012). World Income Inequality 1820-2000. Working Paper.

Baten, J., Crayen, D., and Voth, H. J. (2013). Numeracy and the Impact of High Food Prices in Industrializing Britain, 1780-1850. *The Review of Economics and Statistics* (forthcoming).

Bates, R. (1981). *States and Markets in Tropical Africa: The Political Basis of Agricultural Policy*. Berkeley: University of California Press, Series on Social Choice and Political Economy.

Becker, G. S. (1960). An Economic Analysis of Fertility. In: Demographic and Economic Change in Developed Countries, edited by G.S. Becker. Princeton, NJ: Princeton University Press, 209–240.

Becker, S., and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, 124(2): 531-596.

Binswanger, H. P., and Deininger, K. (1993). South African Land Policy: The Legacy of History and Current Options. *World Development*, 21(9): 1451-1475.

Blum, M., and Baten, J. (2011). Anthropometric within-country Inequality and the Estimation of Skill Premia with Anthropometric Indicators. *Zeitschrift fuer Wirtschaftspolitik*, 62(2): 107-138.

Bodenhorn, H., Guinnane, T.W., and Mroz, T. (2013). Problems of Sample-selection Bias in the Historical Heights Literature: A Theoretical and Econometric Analysis. SSRN Working Paper.

Botticini, M., and Eckstein, Z. (2007). From Farmers to Merchants, Conversions and Diaspora: Human Capital and Jewish History. *Journal of the European Economic Association*, 5(5): 885-926.

Bulmer-Thomas, V. (1994). *The Economic History of Latin America since Independence*. Cambridge: Cambridge University Press.

Case, A., and Paxson, C. (2008). Stature and Status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy*, 116(3): 499-532.

Cinnirella, F., and Hornung, E. (2011). Landownership Concentration and the Expansion of Education (September 30, 2011). CESifo Working Paper Series No. 3603.

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

Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History*, 47 (1): 82–99.

Currie, J. (2009). Healthy, Wealthy, and Wise: Is there a Causal Relationship Between Child Health and Human Capital Development?. *Journal of Economic Literature*, 47(1): 87-122.

Currie, J., and Hyson, R. (1999). Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birth Weight. *American Economic Review*, 89 (2): 245–50.

Deininger, K. (2003). Land Policies for Growth and Poverty Reduction. Oxford: Oxford University Press.

Deininger , K., and Squire, L. (1998). New Ways of Looking at Old Issues: Inequality and Growth. *Journal of Development Economics*, 57(2): 259–287.

Diamond, J. (1997). *Guns, Germs and Steel: The Fates of Human Societies*. New York: W.W. Norton.

Easterlin, R. A. (1980). Towards a more General Economic Model of Fertility Determination: Endogenous Preferences and Natural Fertility. In*: Population and Economic Change in Developing Countries*, edited by R. A. Easterlin. Chicago: University of Chicago Press for National Bureau of Educational Research.

Easterly, W. (2007). Inequality Does Cause Underdevelopment. *Journal of Development Economics*, 84(2): 755-776.

Easterly, W., and Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Division. *Quarterly Journal of Economics*, 152(4): 1203-1250.

Easterly, W., and Levine, R. (2003). Tropics, Germs, and Crops: How Endowments influence Economic Development. *Journal of Monetary Economics*, 50(1): 3-39.

Eastwood, R., Lipton, M., and Newell, A. (2010). Farm Size. In Handbook of Agricultural Economics, edited by R. Evenson and P. Pingali. Amsterdam: Elsevier, 3323-3397.

Enflo, K., and Baten, J. (2006). Estimates of Early Capital Stock Series. Working Paper Univ. Tuebingen/Lund.

Engerman, S., and Sokoloff, K. (1997). Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States. In: How Latin America Fell Behind, edited by S. Haber. Stanford, CA: Stanford University Press.

Engerman, S., and Sokoloff, K. (2002). Factor Endowments, Inequality and Paths of Development among New World Economies. *Economia*, 3(1): 41-109.

Frankel, J. (2010). The Natural Resource Curse: A Survey. Harvard Kennedy School Working Paper 10-005.

Frankema, E. (2010). The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions?. *Economic History Review*, 63(2): 418-451.

Frankema, E., and Van Waijenburg, M. (2012). Structural Impediments to African Growth? New Evidence from Real Wages in British Africa, 1880–1965. *The Journal of Economic History*, 72(4): 895-926.

Galor, O. (2011). Inequality, Human Capital Formation and the Process of Development. NBER Working Paper Series No. 17058.

Galor, O., and Moav, O. (2002). Natural Selection And The Origin Of Economic Growth. *The Quarterly Journal of Economics*, 117(4): 1133-1191.

Galor, O., and Weil, D. (2000). Population, Technology and Growth: from the Malthusian Regime to the Demographic Transition and Beyond. *American Economic Review*, 90(4): 806–828.

Galor, O., and Zeira, J. (1993). Income Distribution and Macroeconomics. *Review of Economic Studies*, 60: 35-52.

Galor, O., Moav, O., and Vollrath, D. (2009). Inequality in Landownership, the Emergence of Human- Capital Promoting Institutions, and the Great Divergence. *Review of Economic Studies*, 76(1): 143-179.

Gelman, J., and Santilli, D. (2010). Rising Inequality. Land Property in Buenos Aires Between 1839 and 1855. Investigaciones de Historia Económica (IHE) *Journal of the Spanish Economic History Association*, 18: 11-33.

Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004). Do Institutions Cause Growth?. *Journal of Economic Growth*, 9(3): 271-303.

Hanushek, E.A., and Woessmann, L. (2012a). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *Journal of Economic Growth*, 17(4): 267-321.

Hanushek, E.A., and Woessmann, L. (2012b). Schooling, Educational Achievement, and the Latin American Growth Puzzle. *Journal of Development Economics*, 99(2): 497–512.

Hippe, R., and Baten, J. (2012). 'Keep Them Ignorant.' Did Inequality in Land Distribution delay Regional Numeracy Development?. University of Tuebingen Working Papers.

Isham, J., Woolcock, M., Pritchett, L., and Busby, G. (2005). The Varieties of Resource Experience: Natural Resource Export Structures and the Political Economy of Economic Growth. *The World Bank Economic Review*, 19:141–174.

Jones, E. L. (1981). *The European Miracle*. Cambridge: Cambridge University Press.

Lindert, P. (2004): *Growing Public. Social Spending and Economic Growth Since the Eighteenth Centruy*. Cambridge: Cambridge University Press.

Lindert, P. H., and Williamson, J. G. (2003). Does Globalization Make the World More Unequal? In: *Globalization in Historical Perspective*, edited by M.D. Bordo, A. M. Taylor and J.G. Williamson. Chicago: University of Chicago Press.

Lipton, M. (2009). *Land Reform in Developing Countries: Property Rights and Property Wrongs*. London: Routledge.

Manzel, K. , Stolz, Y., and Baten, J. (2012). Convergence and Divergence of Numeracy: the Development of Age Heaping in Latin America from the Seventeenth to the Twentieth Century. *The Economic History Review*, 65(3): 932-960.

Montesquieu, Charles de Secondat (1748). *The Spirit of the Laws*. New York: Cambridge University Press, 1989.

Moradi, A., and Baten, J. (2005). Inequality in Sub-Saharan Africa 1950-80: New Estimates and New Results. *World Development*, 33(8): 1233-1265.

North, D. (1981). *Growth and Structural Change*. New York: W. W. Norton.

Nugent, J. B., and Robinson, J. A. (2010). Are Factor Endowments Fate?. *Revista de Historia Economica – Journal of Iberian and Latin American Economic History*, 28(1): 45-82.

Ramcharan, R. (2010). Inequality and Redistribution: Evidence from U.S. Counties and States, 1890-1930. *Review of Economics and Statistics*, 92(4): 729–744.

Resor, R. R. (1977). Rubber in Brazil: Dominance and Collapse, 1876-1945. *The Business History Review*, 51(3): 341-366.

Sachs, J. D. (2001). Tropical Underdevelopment. NBER Working Paper No. 8119.

Sachs, J.D., and Warner, A.M. (1997). Sources of Slow Growth in African Economies. *Journal of African Economies*, 6(3): 335-376.

Sachs, J. D., and Warner, A. M. (1995). Natural Resource Abundance and Economic Growth. NBER Working Paper No. 5398.

Sokoloff, K. L., and Engerman, S. L. (2000). History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives*, 14 (3): 217–232.

Solimano, A. (2001). International Migration and the Global Economic Order: An Overview. Vol. 2720. World Bank Publications.

Stock, J.H., and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In: *Identification and Infernce for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by D.W.K. Andrews and J.H. Stock. Cambridge: Cambridge University Press, 30-108.

Stolz, Y., and Baten, J. (2012). Brain Drain, Numeracy and Skill Premia during the Era of Mass Migration: Reassessing the Roy-Borjas Model. *Explorations in Economic History*, 49: 205-20.

Summerhill, W. R. (2010). Colonial Institutions, Slavery, Inequality, and Development: Evidence from São Paulo, Brazil. MPRA Paper, No. 22162.

Wegenast, T. (2009). The Legacy of Landlords: Educational Distribution and Development in a Comparative Perspective. *Comparative Governance and Politics*, 3(1): 81-107.

Wegenast, T. (2010). Cana, Café, Cacau: Agrarian Structure and Educational Inequalities in Brazil. *Revista de Historia Económica – Journal of Iberian and Latin American Economic History*, 28(1): 103-137.

6.8 Tables

Table 6.1: Average Influence of a Land Reform on the Change in Land Inequality

	(1)	(2)
Method	LSDV	LSDV
Land reform	-5.57*	-5.57*
	(0.05)	(0.06)
Western Europe and Offshoots		-11.31***
		(0.00)
Country fixed effects	Yes	Yes
Constant	0.37	0.37
	(0.24)	(0.25)
Observations	60	60
R-squared	0.65	0.80

Notes: Dependent variable: Change in land inequality. Only those countries were included for which the Gini coefficient was available in more than one year. “Change in land inequality” is constructed by the difference in the Gini from one point in time to another. “Land reform” is a dummy variable that takes the value of 1 if a land reform took place between two Gini estimates. “Western Europe and Offshoots” refers to countries with high GDP per capita (over 25000 U.S. Dollar, Maddison 2001), namely Austria, Australia, Finland, France, Italy, Netherlands, New Zealand and Sweden. The estimation technique is Least Squares estimation with country dummy variables.
Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Table 6.2: Summary Statistics of Data included in Regressions

Variable name	Obs. No.	Mean value	Standard deviation	Minimum value	Maximum value
Cognitive skills	69	4.62	0.55	3.09	5.45
Income inequality (Gini 1890)	42	0.38	0.07	0.28	0.62
Early land inequality (Gini coefficient)	52	0.61	0.14	0.29	0.91
Early numeracy (abcc 1820)	54	0.74	0.25	0.08	1.00
Tropic share	69	0.11	0.28	0.00	1.00
Population density (ln)	61	2.22	1.93	-3.14	4.65
Institutional quality (Polity 2)	50	3.25	7.46	-9.00	10.00
Fertility 1950	57	4.14	1.80	1.98	7.34
Ethnic fractionalization	74	0.33	0.23	0.00	0.85
Student enrolment 1930s	58	0.45	0.27	0.03	0.97
Natural resource exports	51	0.33	0.29	0.02	0.99
Physical capital (ln)	38	7.76	1.18	4.78	9.99

Notes: Cognitive Skills is a measure of human capital introduced by Hanushek and Woessmann (2012a) that is constructed out of international test scores in math and sciences.

The measure of land inequality refers to the size distribution of the land holdings. Landholders were those who produced on their own land or those who or rented land. It reflects the control over land as a production factor (at least in the medium run). However, this data includes land holdings used for all purposes – arable land, land used for permanent crops, or pastures. The "quality" of land has not been reflected in most land inequality estimations.

Tropic share' refers to the share of a country's population living in a tropical zone (1995). Source: Center for international development, Geography datasets (download at http://www.cid.harvard.edu/ciddata/geographydata.htm, General Measures of Geography).

Fertility data are constructed as an average of United Nations country data from 1950-1955 and from 1955-1960. See also the Data Appendix.

Table 6.3: Inequality and Human Capital in OLS Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Robust regression	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No
WLS	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes
Early numeracy	1.48***	1.47***	1.30***	0.91**	0.93**	1.16**	1.05***	1.34***	1.03***	0.91***	0.95***	1.27***	1.16***	0.81***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Early land inequality (around 1890)		-1.37*	-1.50*	-1.39**		-1.44**		-0.81***		-1.19**		-1.43*	-1.22*	-1.10***
		(0.07)	(0.07)	(0.02)		(0.04)		(0.00)		(0.02)		(0.05)	(0.05)	(0.00)
Income inequality (1890)					-1.93**		-2.28*		-0.88		-1.85**	-0.93	-0.37	-0.83
					(0.04)		(0.06)		(0.27)		(0.04)	(0.16)	(0.59)	(0.41)
GDP/c 1910			0.16	0.21	0.08	0.46	0.28	-0.07	-0.10	0.01	-0.08			-0.02
			(0.16)	(0.14)	(0.60)	(0.13)	(0.39)	(0.35)	(0.407)	(0.80)	(0.40)			(0.68)
Tropic share				-0.27	0.10	-0.42*	0.00	-0.85***	-0.59***	-0.49***	-0.15	-0.46***	-0.52***	-0.44**
				(0.21)	(0.68)	(0.06)	(0.98)	(0.00)	(0.007)	(0.00)	(0.38)	(0.00)	(0.00)	(0.02)
Population density				0.03	0.05	0.04	0.04	0.03	0.07*	0.05	0.06			0.05
				(0.38)	(0.30)	(0.48)	(0.51)	(0.15)	(0.078)	(0.21)	(0.22)			(0.24)
Institutional quality				-0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00			-0.00
				(0.70)	(0.81)	(0.68)	(0.86)	(0.95)	(0.626)	(0.81)	(0.98)			(0.43)
Fertility				-0.02	-0.09	0.03	-0.04	0.03	-0.05	-0.04	-0.10*			-0.09**
				(0.75)	(0.10)	(0.62)	(0.60)	(0.53)	(0.372)	(0.27)	(0.09)			(0.04)
Ethnic Fractionalization				-0.07	-0.34	-0.14	-0.78	0.30*	0.03	0.09	-0.22			0.13
				(0.81)	(0.52)	(0.75)	(0.28)	(0.07)	(0.919)	(0.75)	(0.54)			(0.65)
Physical capital						-0.17	-0.16							
						(0.30)	(0.29)							
Constant	3.57***	4.39***	3.36***	3.24***	4.30***	2.36*	4.10**	4.14**	4.49***	4.68***	5.57***	4.97***	4.74***	5.43***
	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.089)	(0.015)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Observations	54	44	37	35	32	31	27	35	32	35	32	34	34	30
Adj. R-squared	0.54	0.68	0.71	0.71	0.71	0.69	0.71	0.61	0.82	0.81	0.81	0.76	0.82	0.85

Notes: Dependent variable: Cognitive Skills, average test scores 1964 to 2003. Early numeracy: ABCC 1820; Early land inequality: Gini coefficient; Institutional quality: Polity 2; Fertility: refers to 1950; Fractionalization: Index from Alesina et al. (2003); Population density and GDP/c are in logs; Specifications (8) to (14) are weighted by the square root of population numbers of each country. Models 8 and 9 were estimated with R-command mlrobust and the option weights=popsqrt to include weights by population. Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.1. See Appendix I and J for the same estimations using (i) always the same set of countries and (j) not using income inequality estimates based on volunteer army and prison data sets (see the note to Appendix J for a discussion of potential sample selection bias).

Table 6.4: Regressions of Cognitive Abilities on Land Inequality and other Factors: Instrumental Variable Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation	TSLS	TSLS	LIML	TSLS	TSLS	TSLS
Instruments	Sugar/ (WheatRice) Def. 1 + LowPop* Southern	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3
Second Stage						
Early numeracy	1.57***	1.41***	1.41***	1.00***	1.09***	1.01***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Early land inequality (around 1890)	-1.28*	-2.75**	-2.75**	-2.24*	-2.41**	-2.24*
	(0.07)	(0.01)	(0.01)	(0.08)	(0.02)	(0.08)
GDP/c 1910				0.23	0.25*	0.24
				(0.14)	(0.08)	(0.19)
Tropic share				-0.20	-0.29	-0.18
				(0.43)	(0.22)	(0.45)
Institutional quality					-0.01	-0.00
					(0.60)	(0.77)
Fertility				-0.04		-0.03
				(0.59)		(0.61)
Constant	4.25***	5.28***	5.28***	3.71***	3.44***	3.55***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	44	47	47	37	40	37
Second Stage R-squared	0.71	0.58	0.58	0.74	0.72	0.74
First Stage						
Sugar/(Wheat Rice) Def. 1	0.73**					
	(0.01)					
LowPop*Sou-thern	0.01***					
	(0.00)					
Sugar/(Wheat Rice) Def. 2		0.54**	0.54**			
		(0.01)	(0.01)			
Sugar/(Wheat Rice) Def. 3				0.56**	0.66**	0.56**
				(0.02)	(0.03)	(0.03)
F-Statistic	11.24	6.78	LIML	5.63	5.51	5.29
Durbin-Wu-Hausman p-value	0.99	0.79	LIML	0.41	0.31	0.37
Sargan p-value	0.17					

Notes: Dependent variable (of the second stage): Cognitive Skills. Control variables: ABCC 1820, tropic share, population density (ln), institutional quality, GDP per capita in 1910 (ln) and physical capital (ln). Instrument "Sugar/ (wheat-rice) Def. 1" includes the share of the land that is very suitable and suitable for the plantation of the three crops (sugar, wheat and rice), Instrument "Sugar/ (wheat-rice) Def. 2" the share being very suitable, suitable and moderately suitable and, last, "Sugar/ (wheat-rice) Def. 3" includes the share of the land very suitable, suitable, moderately suitable and marginally suitable for crop plantation.
Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

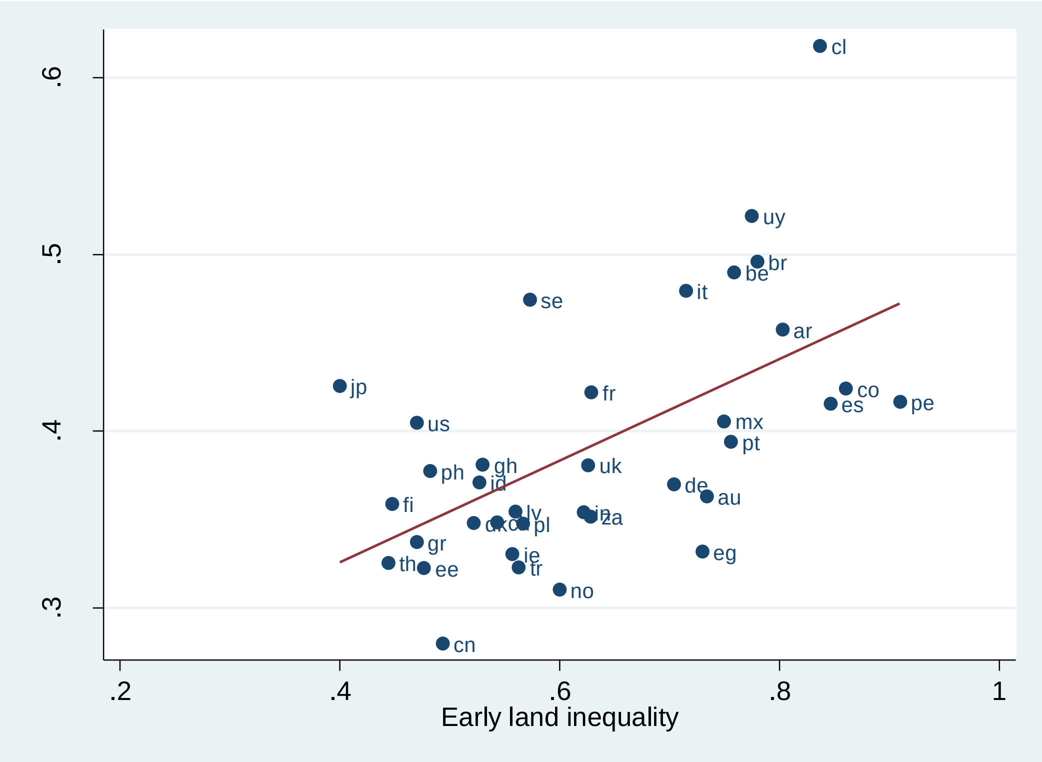
Table 6.5: Instrumental Variable Regression controlling for Primary Goods Exports

	(1)	(2)	(3)
Method	TSLS	TSLS	TSLS
Instruments	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 1 + LowPop* Southern	Sugar/ (WheatRice) Def. 2 + LowPop* Southern
Second Stage			
Land inequality (Gini 1890)	-2.67**	-1.10*	-1.42*
	(0.04)	(0.09)	(0.07)
Early numeracy (abcc 1820)	1.21***	1.21***	1.18***
	(0.00)	(0.00)	(0.00)
Natural resource exports 1980	-0.05		
	(0.85)		
Tropic share	-0.55**	-0.35	-0.36
	(0.03)	(0.16)	(0.13)
School enrollment		0.34	0.36*
		(0.12)	(0.08)
Constant	5.47***	4.26***	4.47***
	(0.00)	(0.00)	(0.00)
Observations	40	41	41
Second Stage R-squared	0.65	0.75	0.76
First Stage			
Sugar/(WheatRice) Def. 1		0.71**	
		(0.03)	
Sugar/(WheatRice) Def. 2	0.66**		0.63**
	(0.04)		(0.01)
LowPop*Southern		0.01***	0.01***
		(0.00)	(0.00)
F-Statistic	4.63	8.68	6.69

Notes: Dependent variable (of the second stage): Cognitive Skills. Control variables: ABCC 1820, tropic share, primary goods exports and school enrollment in the 1930s. Instrument "Sugar/ (wheat-rice) Def. 3" includes the share of the land that is very suitable, suitable and moderately for crop plantation. Natural resource exports: raw material and mining exports per GDP. Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

6.9 Figures

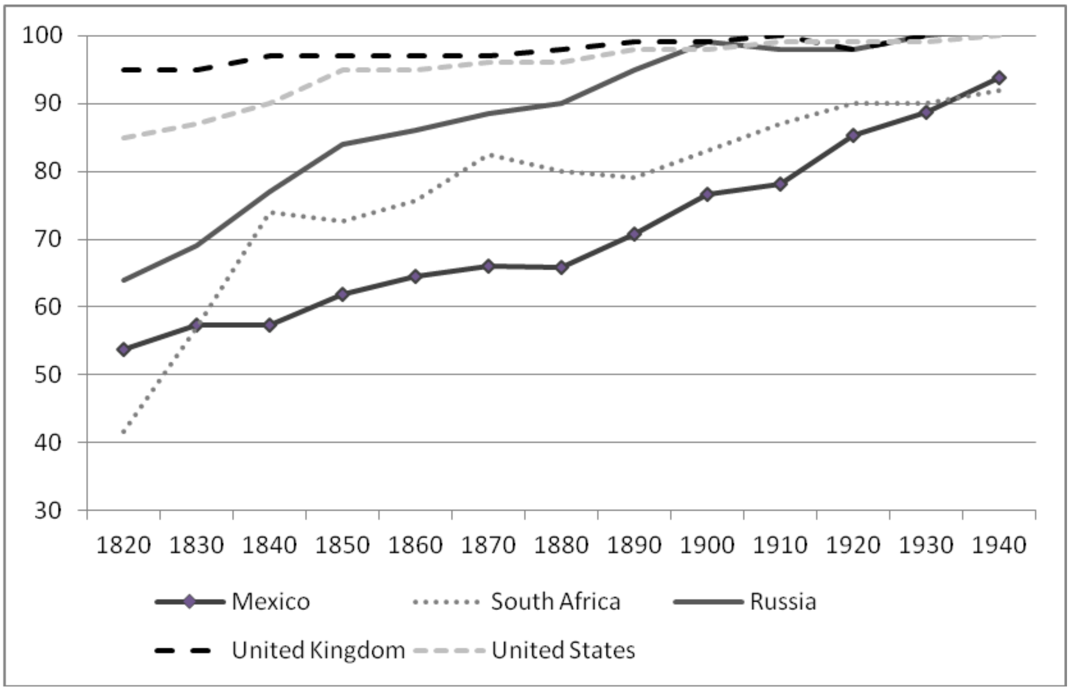
Figure 6.1: Land Inequality versus Income Inequality



Note: Abbreviatios follow ISO codes. See also Appendix C

Figure 6.2: How Path-Dependent is Numeracy over Time? (ABCC Index of selected countries)

Panel A



Panel B

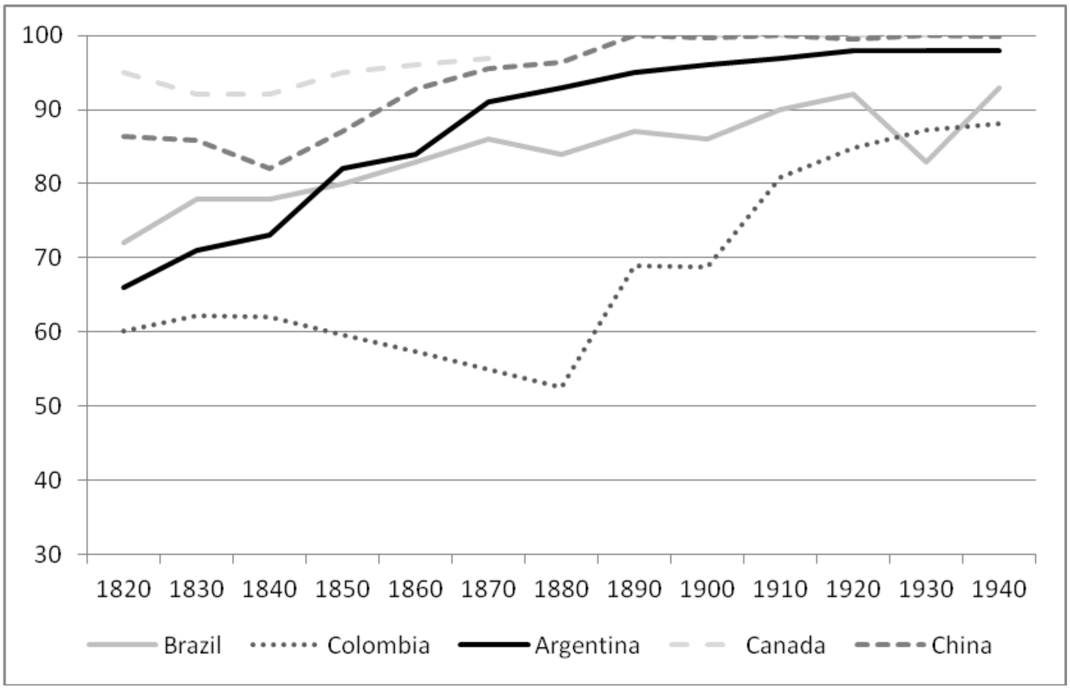
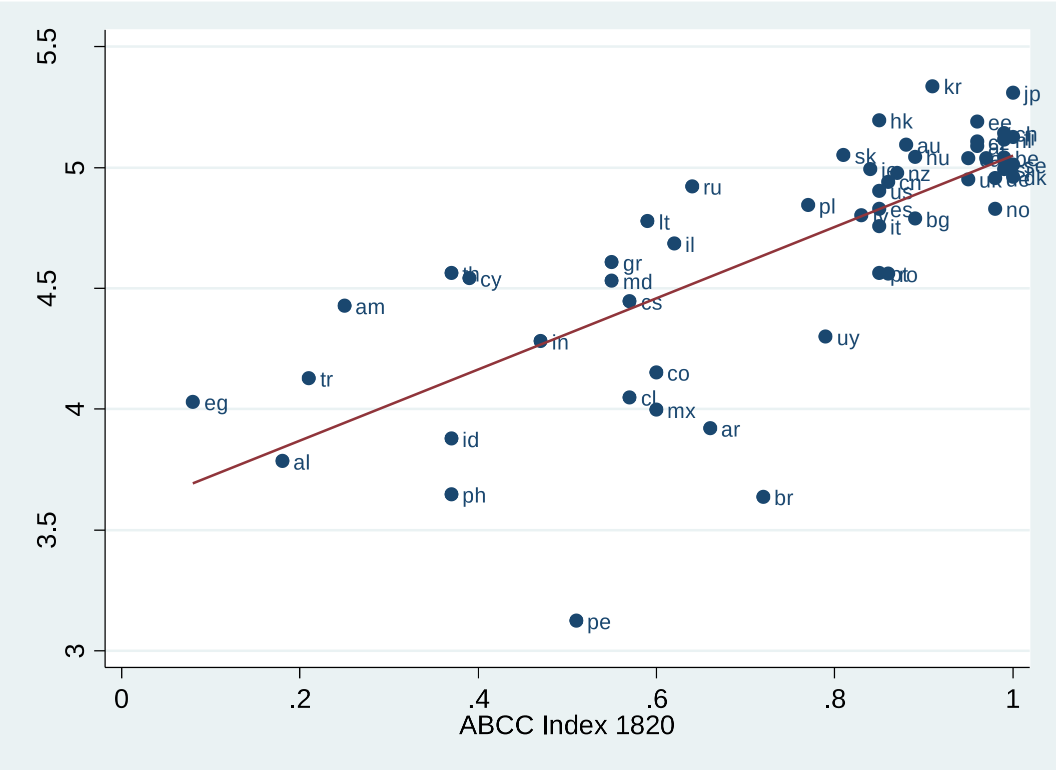
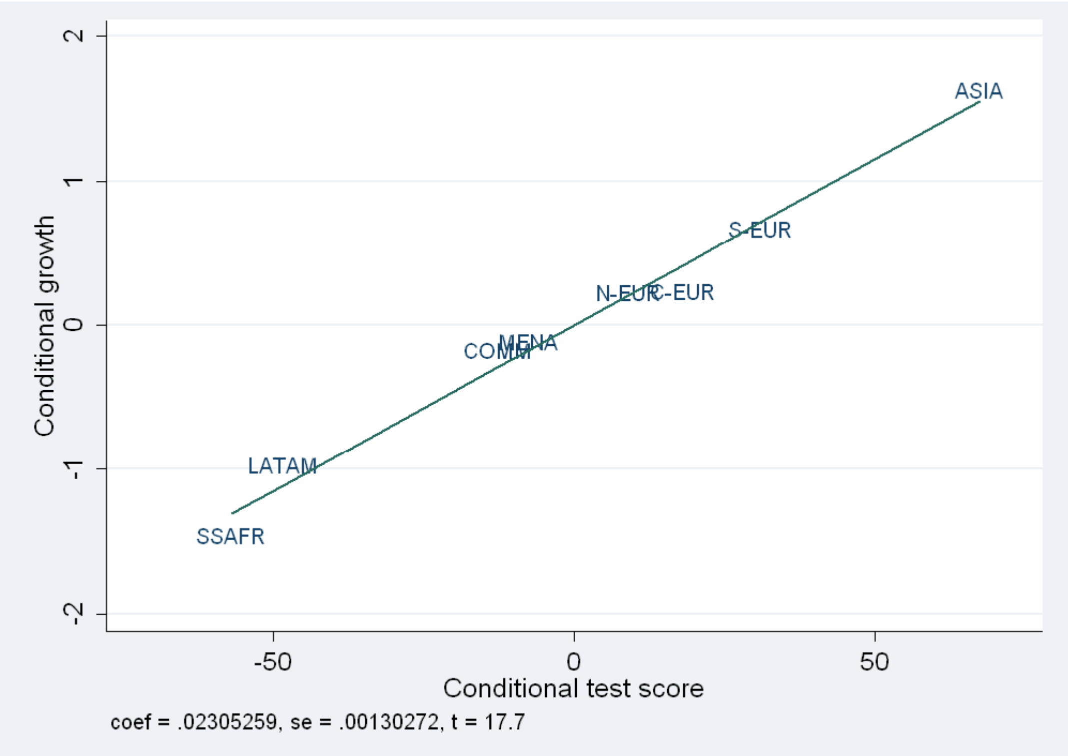


Figure 6.3: Math and Science Cognitive Skills 1964-2003 and Numeracy 1820 (ABCC Index)



Note: Abbreviatios follow ISO codes. See also Appendix C

Figure 6.4: Cognitive Skills and Growth Across World Regions



Source: Hanushek and Woessmann (2012a)
Notes: Region codes are East Asia and India (ASIA), Central Europe (C-EUR), Commonwealth OECD members (COMM), Latin America (LATAM), Middle East and North Africa (MENA), Northern Europe (N-EUR), Southern Europe (S-EUR), Sub-Saharan Africa (SSAFR).

Figure 6.5: Cognitive Skills, 1964-2003

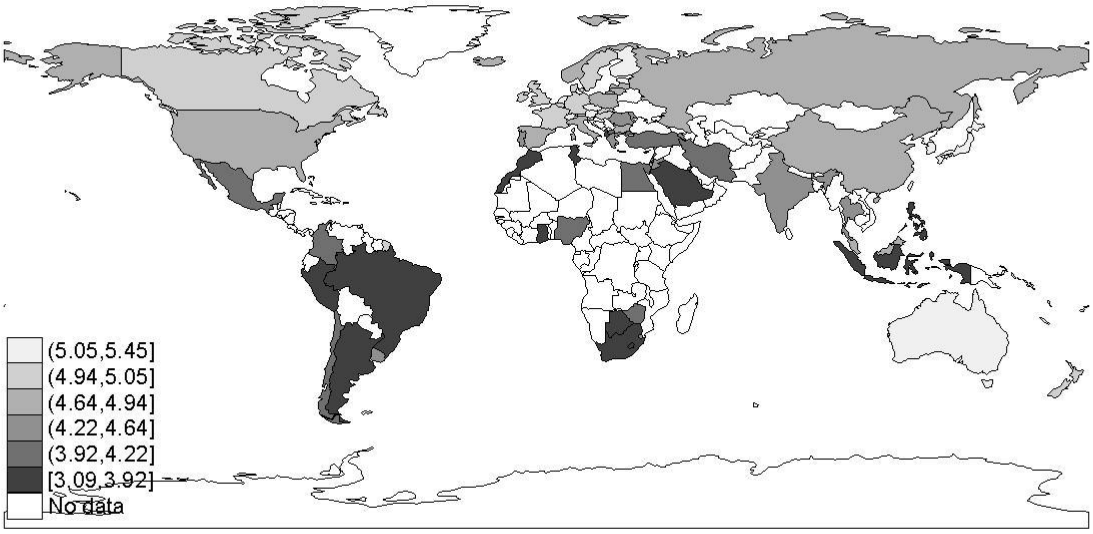


Figure 6.6: Early Land Inequality (around 1890, Gini coefficients)

6.10 Appendix

Appendix A: Regressions of Cognitive Abilities on Income Inequality and other Factors: Instrumental Variable Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation	TSLS	TSLS	LIML	TSLS	TSLS	TSLS
	Sugar/ (WheatRice) Def. 1 + LowPop* Southern	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 2	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3	Sugar/ (WheatRice) Def. 3
Second Stage						
Early numeracy	1.75***	1.80***	1.40*	1.25**	1.30***	1.35**
	(0.00)	(0.00)	(0.06)	(0.03)	(0.00)	(0.01)
Early income inequality (around 1890)	-5.04	-7.06**	-11.13	-7.97	-9.37	-10.11
	(0.17)	(0.04)	(0.46)	(0.45)	(0.13)	(0.38)
GDP/c 1910			0.47	0.36	0.43	0.41
			(0.60)	(0.61)	(0.14)	(0.57)
Tropic share			0.23	0.19	0.06	0.21
			(0.60)	(0.56)	(0.86)	(0.57)
Institutional quality			0.01		0.01	0.01
			(0.43)		(0.60)	(0.35)
Fertility			-0.02	-0.06		-0.03
			(0.94)	(0.73)		(0.87)
Constant	5.24***	6.00***	4.37	4.29	4.01***	4.48
	(0.00)	(0.00)	(0.17)	(0.11)	(0.01)	(0.12)
Observations	37	39	33	34	35	33
Second Stage R-squared	0.61	0.38			0.11	0.02
First Stage						
Sugar/(WheatRice) Def. 1	0.21					
	(0.13)					
LowPop*Southern	0.00					
	(0.94)					
Sugar/(WheatRice) Def. 2		0.20*	0.07			
		(0.07)	(0.57)			
Sugar/(WheatRice) Def. 3				0.72	0.13	0.08
				(0.48)	(0.18)	(0.49)
F-Statistic	1.31	5.02	0.32	0.52	1.91	0.48
Durbin-Wu-Hausman p-val.	0.40	0.07	0.33	0.50	0.14	0.33
Sargan p-value	0.88					

Notes: Dependent variable (of the second stage): Cognitive Skills. Control variables: ABCC 1820, tropic share, population density (ln), institutional quality, GDP per capita in 1910 (ln) and physical capital (ln). Instrument "Sugar/(WheatRice) Def. 1" includes the share of the land that is very suitable and suitable for the plantation of the three crops (sugar, wheat and rice), Instrument "Sugar/(WheatRice) Def. 2" the share being very suitable, suitable and moderately suitable and, last, "Sugar/(WheatRice) Def. 3" includes the share of the land very suitable, suitable, moderately suitable and marginally suitable for crop plantation.
Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Appendix B: Land reforms with some potential impact on land inequality. Abbreviations: See Appendix C

co	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990
al							1	1				1
am												
ar												
at												
au												1
be												
bg	1				1							
bh												
bo								1				1
br												1
bw												
ca	1											
ch												
cl									1	1		
cn							1	1		1		
co						1			1			
cs												
cu									1			
cy												
cz					1		1					
de					1							
dk					1							
dz										1		
ec									1	1		
ee				1								
eg								1				
es						1						
et										1		
fi				1								
fr												
gh												
gr												
gt								1				
hk												
hn										1		
hr					1		1					
hu							1	1				1
id									1			
ie	1											
il												

co	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990
in										1		
iq										1		
ir									1	1		
is												
it												
jo												
jp							1					
ke									1			
kr							1					
kw												
la										1		
lb												
li					1							
lk										1		
lt												
lu												
lv				1								
ma												
md												
mk												
mo												
mx				1		1	1					
my												
na												1
ng												
ni											1	
nl												
no					1							
nz												
pe								1	1		1	
ph									1		1	
pl					1		1					
pr							1	1				
ps												
pt												
ro					1		1					
ru				1								1
sa												

Appendix C: Sources for land reforms

Co. abbr.	Country	Source
AL	Albania 1946, 1950, 1991	Statistical Yearbook of the Popular Republic of Albania 1963. Tirana: Department of Statistics, 1964 Lerman, Z., Csaki, C., and Feder, G. (2004). Agriculture in Transition: Land Policies and Evolving Farm Structures in Post-Soviet Countries. Landham, MD: Lexington Books.
DZ	Algeria 1971	http://www.economywatch.com/agrarian/land/algeria.html
AU	Australia 1993	United Nations Economic and Social Council (2007). Indigenous Land Reform in Australia.
BO	Bolivia 1953, 1996	Tai, H. C. (1974). Land Reform and Politics: A comparative Analysis. Berkeley: University of California Press. Baranyi, S., Deere, C. D., and Morales, M. (2004). Land and Development in Latin America: Openings for Policy Research. National Library of Canada Cataloguing in Publication.
BG	Bulgaria, end of 19th c., 1920-1923	Cochrane, N. J. (1993): Central European Agrarian Reforms in a Historical Perspective. *American Journal of Agricultural Economics*, 75(3) : 851-856
BR	Brazil 1990s	Deininger, K. (1999). Making Negotiated Land Reform Work: Initial Experience from Colombia, Brazil and South Africa. *World Development,* 27: 651-672.
CA	Canada 1873	Smith, P. J. (1987). The Ideological Origins of Canadian Confederation. *Canadian Journal of Political Science*: 3-30
CL	Chile 1964-1973	Collier, S., and Sater, W. F. (1996). A History of Chile: 1808-1994. Cambridge: Cambridge University Press.
CN	China 1948, 1950s, 1970s	Twitchett, D., Fairbank, J. K., and MacFarquhar, R. (1992). The Cambridge history of China. Cambridge: Cambridge University Press.
CO	Colombia 1936, 1958-1970	Deininger, K. (1999). Making Negotiated Land Reform Work: Initial Experience from Colombia, Brazil and South Africa. *World Development,* 27: 651-672.
CS	Serbia 1919, 1945	Z. Lerman, C. Csaki, and G. Feder (2004). Agriculture in Transition: Land Policies and Evolving Farm Structures in Post-Soviet Countries. Lanham, MD: Lexington Books.
CU	Cuba 1959-1963	Kellner, D. (1988). Ernesto “Che” Guevara - World Leaders Past and Present. Chelsea House Publishers.
CZ	Czech Rep. 1919, 1945, 1948	Cornwall, M. (1997). 'National Reparation'?: The Czech Land Reform and the Sudeten Germans 1918-38. *The Slavonic and East European Review*, 75(2): 259-280.
DE	Germany	Brockhaus Encyclopedia (1929), 15th edition, Article “Bodenreform”, Leipzig: Brockhaus.
DK	Denmark 1919	Encyclopedia Britannica [http://www.britannica.com/]
EC	Ecuador 1964, 1973	Blankstein, C.S., and Zuvekas, C. (1973). Agrarian Reform in Ecuador: An Evaluation of past Efforts and the Development of a New Approach. *Economic Development and Cultural Change*, 22(1): 73-94
EG	Egypt 1952	Abdel Malek, A.(1968). Egypt: Military Society. New York: Random House.
ES	Spain 1936-1939	Bernecker, W. L. (1991). Krieg in Spanien 1936-1939. Darmstadt: Wissenschaftliche Buchgesellschaft.

EE	Estonia 1919	Meyers, W. H., and Kazlauskiene, N. (2005). Land reform in Estonia, Latvia, and Lithuania: A comparative analysis. In: Land Reform in the Former Soviet Union and Eastern Europe. Edited by Wegren, S. K. London: Routledge.
FI	Finland 1922	Encyclopedia Britannica [http://www.britannica.com/]
GT	Guatemala 1944-1954	Gleijeses, P. (1989). The Agrarian Reform of Jacobo Arbenz. *Journal of Latin American Studies*, 21(3): 453-480.
HN	Honduras 1973-1977	http://www.fao.org/righttofood/KC/downloads/vl/docs/AH265.pdf
HU	Hungary 1945	Wegren, S. K. (2005). Land Reform in the Former Soviet Union and Eastern Europe. London: Routledge.
ID	Indonesia 1960	A rather ineffective landreform: http://countrystudies.us/indonesia/65.htm
IN	India 1977	Deininger, K. (2003). Land Policies for Growth and Poverty Reduction. Oxford: World Bank and Oxford University Press.
IR	Iran 1962-1971	Amid, M. (1990). Agriculture, Poverty and Reform in Iran. London: Routledge.
IQ	Iraq 1970	Rodinson, M. (1979). Marxism and the Muslim world. Zed Press.
JP	Japan 1947	Kawagoe, T. (1999). Agricultural Land Reform in Postwar Japan: Experiences and Issues. World Bank Policy Research Working Paper.
KE	Kenya 1963	Kanyinga, K. (2009): Land Redistribution in Kenya, In: *Agricultural Land Redistribution: Toward Greater Concesus*, edited by H. Binswanger-Mkhize et al. Washington: The World Bank Press.
KR	Korea, Rep. 1945-1950	Jeon, Y.D., and Kim, Y.Y. (2000):. Land Reform, Income Redistribution, and Agricultural Production in Korea. *Economic Development and Cultural Change*, 48(2): 253-68.
LA	Laos 1975	Stuart-Fox, M: (1997). A History of Laos. Cambridge: Cambridge University Press.
LK	Sri Lanka 1972-1975	Abt Associates (1999). *The Land Tenure System in Sri Lanka*. Bethesda, Md.
LV	Latvia 1918-1919	Meyers, W. H. and Kazlauskiene, N. (2005). Land reform in Estonia, Latvia, and Lithuania: A comparative analysis. In: *Land Reform in the Former Soviet Union and Eastern Europe*, edited by S. K. Wegren, S. K.. London: Routledge.
MX	Mexico 1910, 1934	Dunn, M. (2000). Privatization, Land Reform, and Property Rights: The Mexican Experience. *Constitutional Political Economy*, 11
NA	Namibia 1990	Weidlich, B. (2010). Land ministry tests new farm acquisition model. *The Namibian*. Tapscott, C. (1994). Land reform in Namibia: Why not?. *Southern Africa Report*, 9(3): 12
NI	Nicaragua 1979-1990	Deininger, K., Zegarra, E., and Lavadenz, I. (2003). Determinants and Impacts of Rural Land Market Activity: Evidence from Nicaragua. *World Development, Elsevier*, 31(8): 1385-1404.
NO	Norway	Norway (1979). Lov om jordskifte o.a. (jordskifteloven). Oslo.
PE	Peru 1950s, 1968, 1985-1988	Deininger, K. (2003). Land Policies for Growth and Poverty Reduction. World Bank and Oxford University Press. Zegarra Méndez, E. (1999). El Mercado de Tierras Rurales en el Perú. Productive Development Series no. 63. Santiago de Chile: Economic Commission for Latin America and Caribbean.
PH	Philippines 1960, 1985	Borras, Saturnino M. Jr. (2006): The Philippine Land Reform in Comparative Perspective: Some conceptual and Methodological Implications. Journal of Agrarian Change. 6,(1): 69-101

PL	Poland 1919-1928, 1944	Swinnen, J. F. M. (1999). The political economy of land reform choices in Central and Eastern Europe. *Economics of Transition*, 7: 637-664. Wolf, N. (2005). Path dependent border effects: the case of Poland's reunification (1918-1939). *Explorations in Economic History*, 42: 414-438
RO	Romania 1921, 1945, 1991	Sabates, R. (2005). Cooperation in the Romanian Countryside: An Insight into Post-Soviet Agriculture. New York: Lexington Books.
RU	Russian Fed. 1906-1917,	Service, R. (1998): *A history of twentieth-century Russia*. Cambridge, Massachusetts: Harvard University Press. http://www.nytimes.com/2002/07/09/opinion/russian-land-reform.html
SN	Senegal 1964	Ubink, J. M., Hoekema, A.J., and Assies, W.J. (2009). Legalising Land Rights: Local Practices, State Respones and Tenure Security in Africa, Asia and Latin America. Leiden University Press.
SV	El Salvador 1980s	Arnson, C. J. (2003). El Salvador's Democratic Transition Ten Years After the Peace Accord. Washington, D.C: Woodrow Wilson International Center for Scholars.
SY	Syria 1958-1961	Heydemann , S. (1999). Authoritarianism in Syria: Institutions and Social Conflict 1946-1970. Ithaca: Cornell University Press.
TN	Tunisia 1963	http://perspective.usherbrooke.ca/bilan/servlet/BMEve?codeEve=916
TW	Taiwan (Chinese Taipei) 1950s	Yager, J.A. (1988). Transforming Agriculture in Taiwan: The Experience of the Joint Commission on Rural Reconstruction. Ithaca: Cornell University Press.
VN	Viet Nam 1953-1956	Fall, B. (1975). The Viet Minh Regime, Government and Administration in the Democratic Republic of Vietnam. Connecticut: Greenwood Press.
ZA	South Africa 1994	Deininger, K. (1999). Making Negotiated Land Reform Work: Initial Experience from Colombia, Brazil and South Africa. *World Development*, 27: 651-672
ZW	Zimbabwe	Moyo, S. (2000). Land Reform Under Structural Adjustment in Zimbabwe: Land Use Change in the Mashonaland Provinces. Uppsala: Nordiska Afrikainstitutet.

Appendix D: Data sources

Variable name	Source
Cognitive skills (average scores of international tests in math and science 1964-2003)	Hanushek, E.A., and Woessmann, L. (2012). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *Journal of Economic Growth* , 17(4): 267-321.
Income inequality (Gini 1890)	- Baten, J., Földvari, P., van Leeuwen, B., and van Zanden, J.L. (2012). World Income Inequality 1820-2000. Working Paper. - Blum, M., and Baten, J. (2011). Anthropometric within-country Inequality and the Estimation of Skill Premia with Anthropometric Indicators. *Review of Economics - Jahrbuch fuer Wirtschaftswissenschaften,* 62(2): 107-138.
Early land inequality (Gini coefficient)	[adapted from] Frankema, E. (2010). The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions?. *Economic History Review*, 63(2): 418-451.
Early numeracy (abcc index 1820)	Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History,* 47 (1): 82–99. [Complemented with new evidence based on Manzel, K., Baten, J., Stolz, Y. (2012). Convergence and divergence of numeracy: the development of age heaping in Latin America, 17th to 20th centuries. *Economic History Review*, 65 (3): 932–960.]. For details see Appendix G.
Tropic share (the share of a country's population living in a tropical zone, 1995)	Center for international development at Harvard University, General Measures of Geography [http://www.cid.harvard.edu/ciddata/geographydata.htm]
Population density (ln) (total area population / land area in square kilometers)	Total population numbers: Maddison, A. (2001). The World Economy**:** A Millennial Perspective. Paris: OECD Publishing. Country area: CIA World Fact Book [https://www.cia.gov/library/publications/the-world-factbook/]
Institutional quality (Polity II)	Gurr, T. R. (1990). Polity II: Political Structures and Regime Change, 1800-1986. Boulder, CO: Center for Comparative Politics. [http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/9263]
Fertility 1950 (Births per woman)	World Population Prospects: The 2010 Revision [Downloaded at http://data.un.org/Default.aspx]
Ethnic fractionalization	Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth,* 8(2): 155-194. [http://www.anderson.ucla.edu/]
Student enrolment 1930s (Primary Enrollment*100 / Children of School Age)	Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History,* 47 (1): 82–99. [Based on: Benavot, A., and Riddle, P. (1988). The expansion of primary education, 1870-1940: Trends and Issues. *Sociology of Education,* 66 (3): 191-120 and Lindert, P. (2004). Growing Public: Social Spending and Economic

	Growth Since the Eighteenth Century, Vol. 2: Further Evidence. Cambridge: Cambridge University Press]
Natural resource exports around 1980 (percentage value of total exports)	World Bank Data 1999 (CD-Rom)
Physical capital (ln) about 1925	Enflo, K. and Baten, J. (2006). Estimates of Early Capital Stock Series. Working Paper Univ. Tuebingen/Lund.
Population 1500 (in millions)	Maddison, A. (2003). The World Economy Historical Statistics**.** Paris: OECD Publishing.
Southern Latitude	Comin, D., Easterly, W., Gong, E., 2010. Was wealth of nations determined in 1000 BC? American Economic Journal: Macroeconomics 2 (3), 65–97.

Appendix E: Educational performance in 1820 and in 1964-2003

Rank	country	Abcc 1820	Rank	country	cognitive
1	Finland	1	1	Taiwan (Chinese Taipei)	5.452
1	Japan	1	2	Korea, Rep.	5.338
1	Sweden	1	3	Singapore	5.330
1	Denmark	1	4	Japan	5.310
2	Switzerland	0.99	5	Macao-China	5.260
2	Slovenia	0.99	6	Hong Kong	5.195
2	Belgium	0.99	7	Estonia	5.192
2	Netherlands	0.99	8	Switzerland	5.142
3	Norway	0.98	9	Liechtenstein	5.128
3	Germany	0.98	10	Finland	5.126
4	France	0.97	11	Netherlands	5.115
5	Czech Rep.	0.96	12	Czech Rep.	5.108
5	Austria	0.96	13	Australia	5.094
5	Estonia	0.96	14	Austria	5.089
6	Canada	0.95	15	Slovak Rep.	5.052
6	United Kingdom	0.95	16	Hungary	5.045
7	Korea, Rep.	0.91	17	Belgium	5.041
8	Bulgaria	0.89	18	France	5.040
8	Hungary	0.89	19	Canada	5.038
9	Australia	0.88	20	Sweden	5.013
10	New Zealand	0.87	21	Ireland	4.995
11	Romania	0.86	22	Slovenia	4.993
11	China	0.86	23	New Zealand	4.978
12	United States	0.85	24	Denmark	4.962
12	Spain	0.85	25	Germany	4.956
12	Hong Kong	0.85	26	United Kingdom	4.950
12	Italy	0.85	27	China	4.939
12	Portugal	0.85	28	Iceland	4.936
13	Ireland	0.84	29	Russian Fed.	4.922
14	Latvia	0.83	30	United States	4.903
15	Slovak Rep.	0.81	31	Poland	4.846
16	Uruguay	0.79	32	Malaysia	4.838
17	Poland	0.77	33	Norway	4.830
18	Brazil	0.72	34	Spain	4.829
19	Argentina	0.66	35	Latvia	4.803
20	Russian Fed.	0.64	36	Bulgaria	4.789
21	Israel	0.62	37	Lithuania	4.779
22	Colombia	0.6	38	Italy	4.758
22	Mexico	0.6	39	Israel	4.686
23	Lithuania	0.59	40	Luxembourg	4.641
24	Chile	0.57	41	Greece	4.608
24	Serbia	0.57	42	Thailand	4.565
25	Greece	0.55	43	Portugal	4.564
25	Moldova	0.55	44	Romania	4.562
26	Peru	0.51	45	Cyprus	4.542
27	India	0.47	46	Moldova	4.530

28	Cyprus	0.39	47	Serbia	4.447
29	Thailand	0.37	48	Armenia	4.429
29	Indonesia	0.37	49	Swaziland	4.398
29	Philippines	0.37	50	Uruguay	4.300
30	Armenia	0.25	51	India	4.281
31	Turkey	0.21	52	Jordan	4.264
32	Albania	0.18	53	Iran	4.219
33	Egypt	0.08	54	Nigeria	4.154
			55	Colombia	4.152
			56	Macedonia	4.151
			57	Turkey	4.128
			58	Bahrain	4.114
			59	Zimbabwe	4.107
			60	Palestine	4.062
			61	Chile	4.049
			62	Kuwait	4.046
			63	Egypt	4.030
			64	Mexico	3.998
			65	Lebanon	3.950
			66	Argentina	3.920
			67	Indonesia	3.880
			68	Tunisia	3.795
			69	Albania	3.785
			70	Saudi Arabia	3.663
			71	Philippines	3.647
			72	Brazil	3.638
			73	Ghana	3.603
			74	Botswana	3.575
			75	Morocco	3.327
			76	Peru	3.125
			77	South Africa	3.089

Appendix F: The exclusion restriction for instrumental variables

One of the biggest challenges in any instrumental variable approach is the requirement of the exclusion restriction, which implies that the instrumental variables do not have a direct influence on the ultimate dependent variable except via the instrumented variable. In his seminal paper, Easterly (2007) studied the applicability of the exclusion restriction of relative soil and climatic suitability by using both theoretical reasoning and econometric tests. One possibility for such a direct causal channel is the possibility that wheat/rice and sugar have different effects on the wealth of the local population. This wealth difference could be a potential direct causal influence on cognitive abilities because those might depend on different investment possibilities. On the other hand, Easterly argues convincingly that the difference in the wealth effects of those agricultural goods are quite limited compared to all of the other goods that countries are producing.

Another potential violation of the exclusion restriction could stem from the widely discussed concept of the "natural resource curse" (Easterly 2007). Exceptionally high incomes from raw material exports might generate rents that in turn could lead to political economy problems (Sachs und Warner 1995, Auty 1993, see also the recent review by Frankel 2010). Sugar cane is a primary product that might produce such high windfall profits, for example. Isham et al. (2005) have developed a theory of "point-source" agricultural exports. Typical cases are exports such as sugar cane. The idea is that the "point-source" export revenues can more easily be captured by ruling elites than "diffuse" exports such as wheat and rice. Easterly (2007) argues that if these "resource curse" effects operate via inequality, the exclusion restriction is of course not violated. Most of the studies discussing these issues emphasize that the behavior of rich elites and their interactions with the institutional environment is the main issue, which is consistent with the inequality story (Engerman and Sokoloff 1997, Isham et al. 2005).

Nevertheless, one can still imagine that the resource curse works through other channels. One strategy to address these issues is to directly include additional controls for a resource-oriented export structure and determine whether inequality, measured with the sugar/wheat-rice suitability variable, turns insignificant. We again follow the literature in applying this strategy (Easterly 2007). We construct a variable of the share of raw material and mining exports relative to the country's total exports. The "resource curse" variable is insignificant for those cases that could be included and does not affect the significance of land inequality measured by the suitability variable (Table 6.5).

What about the theoretical properties of the instrumental variable "low population density in 1500 interacted with southern location"? Low population density could, for example, have a negative impact on schooling, because commuting costs are much higher in sparsely populated countries. However, the population that experienced this situation around 1500 does not have much in common with the population in the 20th century. When large-scale immigration to the southern cone countries and Australia started in the 19th century, the composition of the population changed completely. Immigration countries such as Argentina, South Africa and Australia have a reputation for much higher schooling independent of high or low population density.

We also consider a potentially similar concern that might arise regarding our second measurement (even if it turns out to be less crucial to the overall analysis above). For example, we could imagine that the "low population density in 1500- southern latitude" countries such as Argentina and Australia were not only unequal in their land distribution but also invested differently in public schooling. However, we expect those countries to have actually invested more in schooling, which should have affected modern math/science skills positively (whereas their land inequality was high, which should have had a negative effect). Nevertheless, using the same strategy and explicitly including school enrollment in the 1930s as an additional control, we again find that measured land inequality is still significant (Table 6.5, Model 2 and 3).

Appendix G: Sources of the ABCC 1820 estimates in detail (for reference abbreviations, see below)

Western Europe

Austria 1880 (Rothenbacher); Belgium 1856, 1866, 1880, 1890 (Rothenbacher); Switzerland 1860, 1870, 1880, 1888 (Rothenbacher); Denmark 1870, 1880, 1890 (Rothenbacher); Finland 1880, 1900 (Rothenbacher); France 1851, 1856, 1861, 1866 (Rothenbacher); Germany 1880 (Hippe and Baten); Ireland 1851, 1861 (Hippe and Baten); Italy 1871, 1931, 1936 (Rothenbacher); Netherlands 1849, 1859, 1869, 1879, 1889, 1899 (Rothenbacher); Norway (Crayen and Baten), using: Norway. Census of Norway 1865 and 1900. Statistics Norway, Oslo; Portugal: Stolz, Yvonne, Baten, J. and Jaime Reis, "Portuguese Living Standards 1720-1980 in European Comparison – Heights, Income and Human Capital", *Economic History Review 66-2* (2013): 545-578; Spain 1900 (Hippe and Baten); Sweden 1880, 1900 (Rothenbacher); Switzerland 1860, 1870, 1888, 1900 (Rothenbacher); United Kingdom 1851 and 1881 (Crayen and Baten), using: Anderson, M. et al., 1979. National sample from the 1851 census of Great Britain [computer file]. Supplied by History Data Service, UK Data Archive (SN: 1316). Colchester, Essex; Schuerer, K., Woollard, M., 2002. National sample from the 1881 census of Great Britain [computer file]. Supplied by History Data Service, UK Data Archive (SN: 4375). Colchester, Essex;

Eastern and Southeastern Europe

Albania 1918 (Hippe and Baten), using: Eberhart, Helmut et al. (2010), Preliminary dataset "Albanische Volkszaehlung von 1918", entstanden an der Karl-Franzens-Universita¨t Graz unter Mitarbeit von Helmut Eberhart, Karl Kaser, Siegfried Gruber, Gentiana Kera, Enriketa Papa-Pandelejmoni und finanziert durch Mittel des Oesterreichischen Fonds zur Foerderung der wissenschaftlichen Forschung; (FWF). On early 19th C numeracy trends in the Balkans, see (Habsburg 1880); ; Armenia (Russian Empire 1897); Azerbeidshan (Russian Empire 1897); Belarus (Russian Empire 1897); Bulgaria 1893 (Hippe and Baten); Croatia 1880 (Rothenbacher); Czech lands 1880 (Rothenbacher); Cyprus, first benchmark decade is the 1870s (Crayen and Baten), on the change of numeracy in the region 1820-70, see Turkey; Estonia (Russian Empire 1897); Georgia (Russian Empire 1897); Greece 1903 (Hippe and Baten). Values are for the 1830s; Hungary 1880 (Rothenbacher); Kazakhstan (Russian Empire 1897); Kyrgystan (Russian Empire 1897); Lithuania (Russian Empire 1897); Latvia (Russian Empire 1897); Moldova (Russian Empire 1897); Poland 1880 (Habsburg part: Rothenbacher), 1880 (Prussian part: Hippe and Baten) and 1897 (Russian part: Russian Empire 1897)); Romania 1880 (Habsburg part: Rothenbacher; Romanian part assumed equal); Russia (Russian Empire 1897); Serbia 1867 (Crayen and Baten), based on friendly communication by Siegfried Gruber, who collected visitation data on a number of Serbian villages. Siegfried Gruber, Karl-Franzens-Universität Graz, Centre for Southeast European History, Project ‘‘Kinship and Social; Security"; Slovenia 1880 (Habsburg 1880); Slovakia 1880 (Habsburg 1880); Tajikistan (Russian Empire 1897); Turkmenistan (Russian Empire 1897); Ukraine (Russian Empire 1897); Uzbekistan (Russian Empire 1897)

Asia/ Oceania and Africa:

Australia: Meinzer, Nicholas (2013) "The selectivity of migrants to Australia: a new methodological approach". Unpubl. Master thesis Univ. Tuebingen; China: Baten, J., Debin Ma, Stephen Morgan and Qing Wang (2010) "Evolution of Living Standards and Human Capital in China in the 18-20th Centuries: Evidences from Real Wages, Age-heaping, and Anthropometrics", *Explorations in Economic History* 47-3: *347-359*; Egypt: 1848 (Census of Cairo), 1907 (Census of Egypt: The Statistical Department of the Ministry of Finance Egypt, 1907. Statistical yearbook of Egypt. 3rd census of Egypt 1905. Cairo, The Government Press; Hong Kong Baten, J., Debin Ma, Stephen Morgan and Qing Wang "Evolution of Living Standards and Human Capital in China in the 18-20th Centuries: Evidences from Real Wages, Age-heaping, and Anthropometrics", *Explorations in Economic History* 47-3 (2010): *347-359*; India: 1891 (Census of India, 1891 (Bombay, Madras, North-Western Provinces) Indian Empire Census of 1891, 1901, 1911 and 1921. The Superintendent of Government Printing India, Calcutta; Indonesia: Southeast Asia estimate of Baten, J, and Johan Fourie "Numeracy in the 18th Century Indian Ocean Region"): ERSA Working Paper No. 270, complemented with evidence used also in the study Baten, J., Mojgan Stegl and Pierre van der Eng (2013). The Biological Standard of Living and Body Height in Colonial and Post-colonial Indonesia, 1770–2000", *Journal of Bioeconomics 15*: 103-122; Japan: Ministry of Internal Affairs and Communications, 1882. First Statistical Yearbook of the Japan Empire. Population statistics of the Province of Kai 1879 (today's Yamamashu Prefecture). Government Publications, Tokyo; Korea: Baten, J and Kitae Sohn: "Back to the 'Normal' Level of Human-Capital Driven Growth? A Note on Early Numeracy in Korea, China and Japan, 1550–1800", University of Tübingen Working Papers in economics and finance, No. 52; New Zealand: Meinzer, Nicholas (2013) "The selectivity of migrants to Australia: a new methodological approach". Unpubl. Master thesis Univ. Tuebingen; Philippines: Southeast Asia estimates of (Crayen and Baten) and Baten, J, and Johan Fourie "Numeracy in the 18th Century Indian Ocean Region," ERSA Working Paper No. 270, complemented with evidence by Kathrin Grether (2012), Langfristige Humankapitalentwicklung auf den Philippinen im international Vergleich. Unpubl. BA Thesis Univ. Tuebingen; Thailand: Southeast Asia estimates of (Crayen and Baten), and Baten, J, and Johan Fourie "Numeracy in the 18th Century Indian Ocean Region," ERSA Working Paper No. 270; Turkey (Russian Empire 1897): evidence on the province of Kars; see the discussion in Crayen and Baten (2010) about the representativeness of the province. See also on the Ottoman census of 1831 Starbatty, Peter (2011). Humankapitalentwicklung im Osmanischen Reich 1760-1810. Regionale und ethnische Unterschiede. Unpubl. BA Thesis Univ. Tuebingen.

The Americas:

Argentina (Manzel, Baten, Stolz 2012), based on Argentina: National census data of 1869 and 1895, published in Somoza, J., 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; Brazil (Manzel, Baten, Stolz). ; Canada: (Crayen and Baten), using the 1852 and 1881 Historical Censuses of Canada (Canada East, Canada West, New Brunswick and Nova Scotia). Université de Montréal, Montréal; Chile : Robert Pertschy (2012), Regionale Unterschiede der langfristigen Humankapitalentwicklung in Chile im 19. Jahrhundert. Unpubl. BA Thesis Univ. Tuebingen; Colombia (Manzel, Baten, Stolz); Ecuador (Manzel, Baten, Stolz); complemented with new evidence by Christian Schneider (2011), Das Humankapital in den Regionen Ecuadors, Unpubl. Diploma Thesis Univ. Tuebingen; Mexico (Manzel, Baten, Stolz); Peru (Manzel, Baten, Stolz); Complemented with evidence by Sabin Guettler (2011), Verbreitung der Bildungsinnovationen in Peru und Ecuador im 18. und 19. Jahrhundert, Unpubl. Diploma Thesis Univ. Tuebingen; United States: 1850, 1860, 1870, 1880, 1900: Ruggles, S., Alexander, J.T., Genadek, K., Goeken, R., Schroeder, M.B., and Sobek, M. (2010). Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota; Uruguay (Manzel, Baten, Stolz).

Reference abbreviations above:

(Crayen and Baten): Crayen, D., and Baten, J. (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. *Explorations in Economic History*, 47(1): 82-99.

(Habsburg 1880): Austro-Hungarian census of 1880, published as Österreichische Statistik, Band 1, Heft 1–3, Band 2, Heft 1–2 and Band 5, Heft 3, 1882–1884. The evidence covers Austria, Bosnia and Herzegovina, Croatia, Czech Republic, Hungary, Slovakia and Slovenia. We merged Austrian, Russian, and German regional statistics to obtain weighed averages for the modern territories of Ukraine and Poland.

(Hippe and Baten): Hippe, R., and Baten, J. (2012) The Early Regional Development of Human Capital in Europe, 1790 – 1880, *Scandinavian Economic History Review*, 60(3): 254-289.

(Manzel, Baten and Stolz) Manzel, K., Baten, J. and Stolz, Y. (2012) "Convergence and Divergence of Numeracy: The Development of Age Heaping in Latin America, 17th to 20th Century", *Economic History Review* 65(3): 932–960. Detailed sources are listed in their online appendix p.4/5

(Rothenbacher): Rothenbacher, F. (2002). *The European Population 1850-1945*. Basingstoke: Palgrave Macmillan.

(Russian Empire 1897): First General Russian Empire Census of 1897. издание центрального статистического комитета министерства внутренних (1899).

Appendix H: How path-dependent is numeracy over time? (Pairwise correlations for dataset used in the Figures above. All correlation coefficients are statistically significant at the 10% level)

	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940
1820	1.00												
1830	0.98	1.00											
1840	0.83	0.91	1.00										
1850	0.83	0.90	0.96	1.00									
1860	0.79	0.87	0.93	0.98	1.00								
1870	0.68	0.78	0.87	0.95	0.98	1.00							
1880	0.65	0.75	0.84	0.94	0.98	1.00	1.00						
1890	0.76	0.83	0.85	0.95	0.97	0.96	0.96	1.00					
1900	0.70	0.76	0.82	0.93	0.96	0.96	0.97	0.98	1.00				
1910	0.76	0.82	0.87	0.95	0.95	0.93	0.94	0.99	0.97	1.00			
1920	0.74	0.79	0.83	0.94	0.95	0.94	0.95	0.99	0.98	0.99	1.00		
1930	0.61	0.61	0.65	0.76	0.75	0.73	0.74	0.80	0.86	0.84	0.86	1.00	
1940	0.73	0.75	0.77	0.88	0.90	0.87	0.89	0.91	0.95	0.89	0.91	0.91	1.00

Appendix I: Inequality and Human Capital in OLS Regressions with Constant Observation Numbers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Robust regression	No	No	No	No	No	Yes	Yes	No	No	No	No	No
WLS	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Early numeracy	1.56***	1.51***	1.21***	0.87***	0.84***	0.87***	0.96***	0.78***	0.88***	1.26***	1.15***	0.81***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Early land inequality		-1.52***	-1.70***	-1.52***		-1.03***		-1.31***		-1.48*	-1.28**	-1.10***
(around 1890)		(0.00)	(0.00)	(0.00)		(0.00)		(0.00)		(0.05)	(0.05)	(0.00)
Income inequality (1890)					-2.07**		-1.00*		-1.93**	-0.87	-0.26	-0.83
					(0.03)		(0.08)		(0.03)	(0.18)	(0.69)	(0.41)
GDP/c 1910			0.19	0.13	0.07	-0.07	-0.14	-0.02	-0.09			-0.02
			(0.14)	(0.39)	(0.66)	(0.38)	(0.20)	(0.84)	(0.36)			(0.68)
Tropic share				-0.30	0.05	-0.81***	-0.67***	-0.51**	-0.17	-0.47***	-0.53***	-0.44**
				(0.36)	(0.91)	(0.00)	(0.00)	(0.03)	(0.56)	(0.00)	(0.00)	(0.02)
Population density				0.04	0.05	0.08***	0.08***	0.06*	0.05			0.05
				(0.39)	(0.20)	(0.00)	(0.00)	(0.09)	(0.15)			(0.24)
Institutional quality				-0.01	-0.00	0.00	0.01	-0.01	-0.00			-0.00
				(0.39)	(0.94)	(0.83)	(0.14)	(0.36)	(0.84)			(0.43)
Fertility				-0.05	-0.10	-0.05	-0.05	-0.08*	-0.12**			-0.09**
				(0.41)	(0.11)	(0.18)	(0.29)	(0.09)	(0.03)			(0.04)
Ethnic Fractionalization				-0.02	-0.33	0.24	0.13	0.17	-0.19			0.13
				(0.95)	(0.46)	(0.19)	(0.56)	(0.56)	(0.52)			(0.65)
Constant	3.39***	4.41***	3.32***	4.08***	4.58***	5.12***	5.32***	5.17***	5.81***	4.99***	4.74***	5.43***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	30	30	30	30	30	30	30	30	30	30	30	30
Adj. R-Squared	0.54	0.69	0.70	0.73	0.73	0.93	0.86	0.85	0.81	0.73	0.81	0.85

Notes: Dependent variable: Cognitive Skills, average test scores 1964 to 2003. Early numeracy: ABCC 1820; Early land inequality: Gini coefficient; Institutional quality: Polity 2; Fertility: refers to 1950; Fractionalization: Index from Alesina et al. (2003); Population density and GDP/c are in logs; Specifications (6) to (12) are weighted by the square root of population numbers of each country. All specifications include heteroskedasticity robust standard errors and clustered standard errors by world region. Models (6) and (7) were estimated with R-command mlrobust. Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10

Appendix J: Inequality and Human Capital in OLS Regressions with Prisoner and Volunteer Soldier-based Estimates of Income Inequality set to Missing Value

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Robust regression	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No
WLS	No	No	No	No	No	No	No	No	No	Yes	Yes	No	Yes	Yes
Early numeracy	1.48***	1.47***	1.30***	0.91**	0.85**	1.16**	0.95***	1.09***	1.13***	0.91***	0.92**	1.20***	1.13***	0.78***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.04)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.01)
Early land inequality (around 1890)		-1.37*	-1.50*	-1.39**		-1.44**		-0.91***		-1.19**		-1.60**	-1.24*	-1.24***
		(0.07)	(0.07)	(0.02)		(0.04)		(0.01)		(0.02)		(0.04)	(0.05)	(0.00)
Income inequality (1890)					-1.20*		-1.07		-0.84		-1.50**	-0.61	-0.24	-0.20
					(0.07)		(0.51)		(0.19)		(0.04)	(0.10)	(0.72)	(0.71)
GDP/c 1910			0.16	0.21	-0.01	0.46	0.26	-0.05	-0.12	0.01	-0.09			0.00
			(0.16)	(0.14)	(0.90)	(0.13)	(0.37)	(0.67)	(0.35)	(0.80)	(0.41)			(0.97)
Tropic share				-0.27	0.04	-0.42*	-0.12	-0.82***	-0.64***	-0.49***	-0.17			-0.52***
				(0.21)	(0.87)	(0.06)	(0.50)	(0.00)	(0.00)	(0.00)	(0.35)			(0.00)
Population density				0.03	0.09**	0.04	0.10*	0.07***	0.08***	0.05	0.08**			0.08***
				(0.38)	(0.03)	(0.48)	(0.10)	(0.00)	(0.01)	(0.21)	(0.04)			(0.01)
Institutional quality				-0.00	0.00	-0.00	-0.00	0.01	0.01	-0.00	-0.00			-0.01**
				(0.70)	(0.85)	(0.68)	(0.79)	(0.28)	(0.14)	(0.81)	(0.66)			(0.03)
Fertility				-0.02	-0.11**	0.03	-0.05	-0.01	-0.02	-0.04	-0.10			-0.07*
				(0.75)	(0.04)	(0.62)	(0.31)	(0.91)	(0.65)	(0.27)	(0.15)			(0.08)
Ethnic Fractionalization				-0.07	-0.22	-0.14	-0.58	0.35	0.16	0.09	-0.17			0.25
				(0.81)	(0.63)	(0.75)	(0.40)	(0.13)	(0.50)	(0.75)	(0.62)			(0.27)
Physical capital						-0.17	-0.22							
						(0.30)	(0.28)							
Constant	3.57***	4.39***	3.36***	3.24**	4.75***	2.36	4.02**	4.56***	4.83***	4.68***	5.43***	5.01***	4.74***	4.94***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.11)	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Observations	54	44	37	35	30	31	25	35	30	35	30	31	31	28
Adj. R-Squared	0.55	0.69	0.73	0.78	0.78	0.78	0.84	0.87	0.90	0.85	0.86	0.80	0.84	0.91

Notes: Dependent variable: Cognitive Skills, average test scores 1964 to 2003. Early numeracy: ABCC 1820; Early land inequality: Gini coefficient; Institutional quality: Polity 2; Fertility: refers to 1950; Fractionalization: Index from Alesina et al. (2003); Population density and GDP/c are in logs; Specifications (6) to (12) are weighted by the square root of population numbers of each country. All specifications include heteroskedasticity robust standard errors and clustered standard errors by world region. Models (6) and (7) were estimated with R-command mlrobust.
Robust p-values in parentheses: ***p<0.01, **p<0.05, *p<0.10.

We took into account the concerns by Bodenhorn, Guinnane and Mroz (2013) that samples of height into which individuals were selected (or into which they self-selected) by labor market forces could theoretically contain a bias. They demonstrate this potential mechanism using the British volunteer army of the 19th century and generalize that this might happen in all volunteer armies and also in prisons, which was another type of source. They mention, however, that this is not likely to happen if the armies were formed on the basis of general conscription, as the French army, for example. Another case of context in which sample selection bias is highly unlikely, were measurements performed by scholarly anthropologists. Fortunately, in our data, these types of samples accounted for more than 90 percent of the height-based inequality estimates, so we could easily run a robustness check excluding the potentially biased volunteer soldier and prisoner samples (please note that Bodenhorn et al. only point to the possibility of selection bias without claiming that this was always the case, and perhaps not even in the majority of volunteer army cases; they assess only the UK voluntary army).

Reference: Bodenhorn, H., Guinnane, T. W., and Mroz, T. (2013). Problems of Sample-selection Bias in the Historical Heights Literature: A Theoretical and Econometric Analysis. University of Yale Working Paper.

7. Summary, Policy Implications and Directions for Future Research

This thesis has assessed different determinants for long-run human capital formation, with a focus on the development of human capital in Latin America and the Iberian Peninsula from a historical perspective. A number of conclusions can be drawn from its results.

First, when assessing the genesis of the different development paths in American regions, not only the colonial legacy should be taken into account, but also a potential pre-colonial heritage. Institutions fostering inequality, for example by impeding schooling for the mass population, are claimed to have hampered the development of some regions. But most studies do not consider the role of "extracting" institutions that were established before the European conquest and mostly adopted later by the colonial administration, such as the "mita" system in the Inca Empire. Chapter number two assessed the "pre-colonial legacy" hypothesis for the Andean region by measuring the educational inequality of different social groups in the Inca Empire with the age-heaping technique.

Chapters three and four were dedicated to the cultural aspect of human capital formation. Chapter three discussed different hypotheses on why the Jewish religion incentivized investments into education. Moreover, the educational differences between this religious minority and the Catholic majority in the Iberian world in the era of the Spanish Inquisition were measured with the use of the age heaping methodology. The result was a significant advantage of the Jewish coreligionists, but also of the Protestant minority and other target groups of the Inquisition tribunal. This finding suggests that the Spanish Inquisition could have had a negative impact on the human capital formation in Spain and its colonies. Chapter four is a continuation of the foregone paper which analyzed the human capital advantage of the Jewish minority for the early 20th century in three Latin American countries.

Selective migration is a further determinant of human capital formation in a country. Highly skilled immigrants can improve the average level of human capital by their sole presence and through spill-over effects on the native population. Chapter number five assessed the skill level of migrants to three different Latin American countries. The educational level of immigrants relative to the native population in the destination countries was higher in Argentina than in Cuba, for example. This result was also determined by the main regions or countries of origin of the migrants. However, Argentina in the late 19th century could profit from a relatively highly skilled immigration that probably contributed to the economic development of the country. Argentina's policy had been to attract immigrants that contributed to its "culturalization". In Cuba, on the other hand, the majority of immigrants were attracted to work in large plantations and did not make an important contribution to the rise of the national educational level.

An important obstacle to human capital formation is high land inequality. Chapter number six addressed the relationship between land inequality and human capital formation in a global cross-country study and found a significant negative impact of land inequality. The main theory that explains this relationship argues that in regions with high land inequality large landowners had the financial backup to influence political decisions. They avoided the implementation of a tax-financed public education system, since education was not very productive in the agricultural sector and it could lead their labour force to seek work in another sector. Regions with very high land inequality such as Latin America turn out to perform relatively weakly in terms of human capital today.

Human capital is clearly a very important driving factor of wellbeing and economic growth. Education has a direct impact on economic development as well as other positive externalities, for example, for health and democracy. In the last decades, the huge relevance of human capital as the motor of technological advance and economic development for today's knowledge based economies has been recognized. However, human capital seems to have been quite persistent throughout history. The countries or regions with the highest educational levels two centuries ago turn out to be the leading countries in terms of education as well as economic development today. Since the effects of human capital are long lasting, it is important to assess what conditions favoured educational investments and what hampered them in such a persistent manner.

Policy actions directed at promoting economic development should aim to increase the level of human capital of the population. Policy makers ought to create an environment in which individuals have an incentive to invest into education and provide the necessary infrastructure. Creating equal opportunities in access to education for the whole population should be a very important goal. In countries in which agriculture still plays a crucial role, the implementation of land reforms to alleviate the concentration of this asset in the hands of a few could improve the access to schooling for the poorer section of the population and provide them with better social and economic opportunities. In developed countries, attracting highly skilled immigrants is a policy measure that can have positive effects on the development of a country.

Several areas may deserve attention for future research. This thesis focused mostly on very basic numerical skills measured by the age heaping technique. It would be fruitful to enrich worldwide human capital research by enlarging age heaping databases to other developing regions, especially the African continent, which is relatively understudied. A further contribution would be the construction of other measures of human capital, either of more advanced or of different skills. Yet another possibility is to include a health-related measure in human capital proxies. Health was already viewed as a component of human capital by the famous "human capitalist" Gary S. Becker in the 1960s and could be considered in empirical research. Empirical human capital research could profit from the combination of age heaping data with other complementary human capital measures. More and improved data are also an important tool for the validation of existing economic theories.