Essays on the economic history of numeracy in Spain

Doctoral Thesis
in order to obtain the title of Doctor
from the Faculty of Economics and Social Sciences
at the University of Tübingen

presented by

María del Carmen Pérez Artés

Tübingen

2020

Date of oral defense:	27.07.2020
Dean:	Professor Dr. rer. soc. Josef Schmid
1st supervisor:	Professor Dr. Jörg Baten
2nd supervisor:	Professor Dr. Carmen Sarasúa

Acknowledgments

First of all, I would like to thank my thesis supervisor, Jörg Baten, for giving me the opportunity to obtain my PhD and to be part of the wonderful research team at the chair of Economic History in Tübingen. Besides providing me insightful comments and advice, he also gave me the chance to attend to numerous international conferences and present my research. I owe him special thanks for assisting me in acquiring a working knowledge in economics and econometrics. Furthermore, I would like to thank my second supervisor, Carmen Sarasúa for her enormous support, for her comments on my thesis and for sharing her database, which was used in chapter 4. She also allowed me to participate in her research project financed by the Spanish Government (HAR2017-85601-C2-1-P).

The research group in Tübingen also provided me with ideas and important feedback for my research. I am particularly grateful to Thomas Keywood, Laura Maravall, Jessica Baier and Elisabeth Kempter. I would also like to acknowledge the support that I received from the chair of Economic History at the University of Almería; in particular from Andrés Sánchez Picón, José Joaquín García Gómez and Víctor Luque de Haro, for dedicating their time to help me to improve my research and obtain Spanish books during my stay in Germany. I also appreciate the financial support I received to present an early version of my thesis at the European Graduate School for Training in Economic and Social History Research at the University of Krakow (2017).

Finally, I would like to thank my parents and my brother for their encouragement and support, and Lukas for making my path easier to follow.

Contents

List of Figures

List of Tables

1. Introduction

1.1 Human capital in Spain prior to the 19th century

The theory of human capital has been a topic broadly studied over the past century (Schultz 1961; Becker 1962; Mincer 1974). In the 1960s, Schultz (1961) claimed that:

> *Although it is obvious that people acquire useful skills and knowledge, it is not obvious that these skills and knowledge are a form of capital, that this capital is in substantial part a product of deliberate investment, that it has grown in Western societies at a much faster rate than conventional (nonhuman) capital, and that its growth may well be the most distinctive feature of the economic system* (Schultz 1961:1).

Becker (1964), another founding father of the theory of human capital added:

> *It is clear that all countries which have managed persistent growth in income have also had large increases in the education and training of their labor forces* (Becker 1964 [1994]: 24).

Since then, many researchers have dedicated their work to estimating human capital in past societies and assessing the impact of human capital on economic growth. As education and training are considered the most important investments in human capital (Becker 1964), traditional studies have used the following indicators to estimate it. The most common measure of human capital is literacy (Cipolla 1969; Romer 1989). Barro (1991) used school-enrolment rates as a proxy for human capital while Barro and Lee (1993) constructed estimates of educational attainment by sex for persons aged 25 and over. In a follow-up study, Barro and Lee (1996) then used years of schooling by sex at various levels of education as their measure. In a more recent paper, Baten and Van

Zanden (2008) utilised book production as a proxy for advanced literacy skills. Using this indicator as well as years of secondary schooling, De Pleijt and Van Zanden (2016) found that human capital formation was the primary driver of the growth that occurred during the Little Divergence.

For Spain, the topic of this dissertation, the ability to sign has allowed researchers to estimate levels of literacy for societies where direct evidence of literacy rates does not exist. This was the case between the sixteenth century and the first half of the nineteenth century. However, neither the sources used, nor the samples are uniform and representative in all cases, showing significant bias. Normally, the individuals recorded in these documents were the most educated and those who belonged to wealthier social strata (Viñao Frago 1999; Rodríguez and Bennassar 1978; Vincent 1987; de la Pascua Sánchez 1989).

The situation changed in 1860, with what is considered the first modern census in Spain. For the first time, a source that systematically recorded the ability of the entire Spanish population to read and write exists. Núñez (1992) analysed this source in detail, paying attention to differences by province and gender. Despite some heterogeneity, she concluded that from 1860 to 1930, Spaniards progressed from a position of very constrained literacy where no more than 30 percent of the population was literate, to almost universal literacy.

Internationally, Spain was among the countries in the second wave of transition to literacy. While the countries of northern Europe and the United States had become practically fully literate by the mid-nineteenth century, those of southern Europe, Japan and Australia did so in the first decades of the twentieth century. By the end of the Second World War, Latin America and certain Asian countries had achieved near-universal

literacy, while the African continent, despite significant progress, lagged behind (Núñez 1992).

In 1860, the north of Spain, with the exception of Galicia, was the most literate Spanish population. The Basque Country, Navarra, Old Castile, Asturias, León and Santander had populations with literacy rates above the national average (44%), whereas those from Valencia, Murcia, Eastern Andalusia and the islands were lagged behind. Finally, the least literate population was in the Mediterranean southeast (Núñez 1992: 93). Recently, Beltrán-Tapia et al (2019) delve into the analysis of regional differences in Spain during the same period at the municipal level. The authors conclude that the greatest reduction in geographic inequality in literacy occurred with the creation of the Ministry of Public Instruction and the state beginning to finance primary education in 1900.

However, how could we obtain representative evidence of human capital formation for societies and periods where traditional sources of education indicators were incomplete? Numeracy, or the ability to deal with numbers, allows us to obtain a more comprehensive sample through age statements, as these can be found in a greater number of sources than alternative measures of human capital (A'Hearn et al 2009). As I study Spanish human capital formation during the sixteenth century and the first half of the eighteenth century, this is the primary human capital indicator used in this thesis. In order to assess numeracy, I employ "age heaping" methodology using the ABCC index. As explained in the following chapters, this method considers the share of individuals able to state their precise age in years, in contrast to those who report an age rounded to a multiple of five (Crayen and Baten 2010a).

During the last decade, several scholars have used this indicator to estimate levels of human capital when traditional sources were scarce[1]. However, very little is currently known about numeracy levels in a Spanish context. Álvarez and Ramos Palencia (2018), using both literacy and numeracy, suggested that in Castile circa 1750, human capital could have contributed to income inequality. The authors prove a positive relationship between human capital and male labour earnings in Spain for the provinces of Palencia, Guadalajara and Madrid. Gómez-i-Aznar (2019) studied numeracy rates in eighteenth century Catalonia, finding that the level of numeracy (73% of the inhabitants were able to state their ages correctly) was relatively high before Industrial Revolution. Juif et al. (2019) found that Jews and New Christians in Spain and Portugal had a substantial advantage in numeracy (around 20% higher numeracy level) over the Catholic majority during the inquisition era. In a later period, Beltrán et al. (2018) analysed age-heaping and literacy in Spain between 1877 and 1930 showing that age heaping remained unchanged during the second half of the nineteenth century, improving significantly from 1920.

Some research has also been done on Spaniards in the Latin American colonies. For example, Juif and Baten (2013) argued that the Spanish settlers were twice as likely to be numerate as the Peruvian Inca Indios. Calderon et al. (2020) have recently found that in late pre-independent Mexico, numeracy was similar to that of peripheral Europe and there were significant ethnic inequalities (*españoles* represented the group with higher rates while *indios* and *mulatos* had the lower rates). Juif (2015) established that

[1] Just to mention some of them see (Crayen and Baten 2010b; Manzel, Baten and Stolz 2012; Tollnek and Baten, 2017; Baten and Fourie 2015).

the poor and least educated population from the Canary Islands moved to Cuba in nineteenth century to work in agriculture.

The aim of this investigation is to increase our knowledge of human capital in Spain during periods (the sixteenth and eighteenth centuries) for which empirical evidence is scarce. The findings of this research will contribute to filling the gap in this topic. Apart from estimating the numeracy levels of Spaniards, I address important research questions posed by economic historians. Was there already a relationship between inequality and human capital in the early Modern Era? What was the self-selection of migrants to Latin America like during the sixteenth century? Did the level of parental human capital, among other factors, have any influence on the schooling and child labour decisions of their children in eighteenth century Castile?

1.2 Outline of the dissertation

This dissertation consists of three chapters that approach different aspects of the Spanish human capital between the sixteenth and eighteenth centuries. The second chapter "Land inequality and numeracy in Spain during the 17th and 18th century", written with Jörg Baten, has been accepted for publication in *Historia Agraria. Revista de agricultura e historia rural*[2].

Chapter two addresses the debate about whether the elites who owned most of the land, and therefore had the strongest political influence, aided or hampered human capital formation. While some authors found that landed elites promoted investments in mass

[2] Jörg Baten co-authored this chapter, contributing approximately 20% of the work to this paper.

schooling (Andersson and Berger 2019), others claimed that large landownership restricted human capital and investment (Galor et al. 2009; Baten and Hippe 2018; Beltrán Tapia and Martínez-Gallarraga 2018). However, all research carried out so far focuses on the nineteenth and twentieth centuries. In this article, we analyse the relationship between human capital formation and land inequality in early modern Spain using an individual-level analysis. We employ a new dataset from the *padrones* (local nominative population censuses) and the Cadaster of Ensenada (1750). Following Clark and Grey (2014), we use "farmer share" (the proportion of farmers of the total of agricultural population) as our land equality indicator. We found that farmer share was always positively correlated to regional numeracy (as opposed to regions with *latifundistas* and many day labourers). In accordance with the literature, we concluded that numeracy among farmers was higher than among agricultural workers (Tollnek and Baten 2017).

The selectivity of migrants and the level of human capital that they transferred is a prominent factor studied by human capital researchers. The majority of these studies found that migrants were positively self-selected during the nineteenth and twentieth centuries (Humphries and Leunig 2009; Quiroga 2003; Beltrán and Salanova 2017; Juif and Quiroga 2019). Additionally, those who migrated from Europe to Latin America during the nineteenth century were, on average, more literate than those who stayed (Sánchez Alonso 2007). However, as stated above, the numeracy level of migrants from the Canary Islands to Cuba in the nineteenth century was lower than the level of those who stayed (Juif 2015). Using new micro data compiled from published passenger lists, in chapter three (Numeracy selectivity of Spanish migrants in Hispanic America (16^{th} - 18^{th} centuries) I analyse the human capital compositions of Spanish migrants who emigrated to colonial Spanish America during the sixteenth and eighteenth centuries. I find that Spanish migrants were already positively self-selected in terms of numeracy in

the sixteenth century. On the other hand, and in accordance with the literature, colonial Hispanic American societies were not especially unequal in the eighteenth century (Baten and Fourie 2015; Calderon et al. 2020; Dobado González and García Montero 2010; Arroyo Abad and van Zanden 2016).

Chapter four (Human Capital, schooling and child labour in New Castile (Spain) in the 18th century) focuses on eighteenth century Castile. Using a database from the Cadaster of Ensenada (circa 1750), I show that family socioeconomic characteristics affected the parental decisions about child labour and the schooling of their children - such as the occupation of the head of the family, the occupation of the mother, the human capital level of the parents, the size of the family, the birth order of the children and the ratio of school-children to teachers or the cost of school (at a municipal level). Although some research has focused on the age that girls and boys started to work, the kind of tasks that they undertook or the schooling among children in Spain during eighteenth and nineteenth century, there has not been much research on family backgrounds (Borderías 2013; Borrás Llop 2002a; Borrás Llop 2002b; Borrás Llop 2002c; Campos Luque 2014; Camps 2002; Hernández 2013; Humphries 2013; Sarasúa 2002a; Sarasúa 2002b; Sarasúa 2013). Moreover and related to chapter two, I find that farmers were more interested in the investment of human capital for their children than day laborers. Finally, Chapter five concludes.

1.3 References

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

Álvarez, B. & Ramos Palencia, F. 2018. Human capital and earnings in eighteenth-century Castile. *Explorations in Economic History*, 67: 105–133.

Andersson, J. & Berger, T. 2019. Elites and the expansion of education in nineteenth-century Sweden. *The Economic History Review*, 72(3): 897-924.

Arroyo Abad, L., van Zanden, J. L. 2016. Growth under Extractive Institutions? Latin American Per Capita GDP in Colonial Times. *The Journal of Economic History*, 76(04): 1182–1215

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

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

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

Baten, J. & Hippe, R. 2018. Geography, land inequality and regional numeracy in Europe in historical perspective. *Journal of Economic Growth*, 23(1): 79–109.

Baten, J. & Van Zanden, J L. 2008. Book production and the onset of modern economic growth. *Journal of Economic Growth*, 13: 217-235.

Baten, J., & Fourie, J. 2015. Numeracy of Africans, Asians, and Europeans during the early modern period: new evidence from Cape Colony court registers. *The Economic History Review*, 68(2): 632–656.

Becker, G. 1962. Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5, Part 2): 9 – 49.

Becker, G. 1964 [1994]. *Human capital revisited. In Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. The University of Chicago Press (Third edition).

Beltrán Tapia, F. J. & de Miguel Salanova, S. 2017. Migrants' self-selection in the early stages of modern economic growth, Spain (1880-1930). *The Economic History Review*, *70*(1): 101–121.

Beltrán Tapia, F. J. & Martinez-Galarraga, J. 2018. Inequality and education in pre-industrial economies: Evidence from Spain. *Explorations in Economic History*, 69: 81–101.

Beltrán Tapia, F. J., Díez-Minguela, A., Martinez-Galarraga, J., & Tirado-Fabregat, D. 2018. Two stories, one fate: Age-heaping and literacy in Spain, 1877-1930. *EHES Working Papers in Economic History*, 139.

Beltrán-Tapia F.J.; Díez-Minguela, A.; Martínez-Galarraga J. & Tirado-Fabregat, D. 2019. Capital humano y desigualdad territorial. El proceso de alfabetización en los municipios españoles desde la ley Moyano hasta la guerra civil. *Estudios de Historia Económica*, 74, Banco de España.

Borderías, C. 2013. Salarios infantiles y presupuestos familiares en la Cataluña Obrera, 1856-1920. In J. M. Borrás Llop (Ed.), *El trabajo infantil en España, 1700-1950*: 371–408. Barcelona: Icaria.

Borrás Llop, J. M. 2002a. Aprender trabajando. La actividad de niñas y niños en tierras de regadío (la Vega del Tajuña a comienzos del siglo XX). In C. Sarasúa & L. Gálvez (Eds.), *¿Privilegios o eficiencia? Mujeres y hombres en los mercados de trabajo*: 157–183. Alicante: Universidad de Alicante.

Borrás Llop, J. M. 2002b. El trabajo infantil en el mundo rural español, 1849-1936. Género, edades y ocupaciones. In J. M. Martínez Carrión (Ed.), *El nivel de vida en la España rural, siglos XVIII-XX*: 497–547. Alicante: Universidad de Alicante.

Borrás Llop, J. M. 2002c. Mercado laboral, escolarización y empleo infantil en una comarca agrícola e industrial (el Valles Occidental, 1881-1910). *Cuadernos de Historia Contemporánea / Departamento de Historia Contemporánea*.

Borrás Llop, J. M. 2005. Schooling and child farm labour in Spain, circa 1880-1930. *Continuity and Change*.

Calderón-Fernández, A., Dobado-González, R., & García-Hiernaux, A. 2020. Numeracy in Central New Spain during the Enlightenment. *Revista De Historia Económica / Journal of Iberian and Latin American Economic History*: 1–35. doi:10.1017/S0212610919000387

Campos Luque, C. 2014. La tasa de actividad femenina a mediados del siglo XIX en Andalucía: el caso de Antequera. *Investigaciones de Historia Económica - Economic History Research*, 10(3): 191–201.

Camps, E. 2002. Trabajo infantil y estrategias familiares durante los primeros estadios de la industrialización catalana (1850-1925). Esbozos a partir del estudio de un caso. *Cuadernos de Historia Contemporánea*: 263–279.

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

Clark, G. & Gray, R. 2014. Geography is not destiny: geography, institutions and literacy in England, 1837–63. *Oxford Economic Papers,* 66 (4): 1042–1069.

Crayen, D. & Baten, J. 2010a. Global trends in numeracy 1820-1949 and its implications for long-term growth, *Explorations in Economic History*, 47: 82-99.

Crayen, D. & Baten, J. 2010b. New evidence and new methods to measure human capital inequality before and during the industrial revolution: France and the US in the seventeenth to nineteenth centuries. *The Economic History Review*, 63(2): 452–478.

De la Pascua Sánchez, M.J. 1989. Aproximación a los niveles de alfabetización en la provincia de Cádiz: las poblaciones de Cádiz, El Puerto de Santa María, Medina Sidonia y Alcalá de los Gazules entre 1675-1800. *Trocadero, Revista de Historia Moderna y Contemporánea*, 1: 51-65.

De Pleijt, A.M. & Van Zanden, J.L. 2016. Accounting for the "Little Divergence": What drove economic growth in pre-industrial Europe, 1300-1800?. *European Review of Economic History*, 20(4): 387-409

Dobado González, R., & García 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.

Galor, O.; Moav, O., & Vollrath, D. 2009. Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. *The Review of economic studies,* 76 (1): 143–179.

Gómez-i-Aznar, È. 2019. Human capital at the beginnings of the 18th century Catalonia: age-heaping and numeracy in a changing economy, *Documentos de Trabajo (DT-AEHE)*, No 1904.

Hernández, R. 2013. La mano de obra infantil en la Castilla rural del siglo XVIII. El trabajo del niño es poco, pero el que lo desprecia un loco. In J. M. Borrás Llop (Ed.), *El trabajo infantil en España, 1700-1950*: 91-115. Barcelona: Icaria.

Humphries, J. (2013). Childhood and child labour in the British industrial revolution. *The Economic History Review*, 66(2): 395–418

Humphries, J. and Leunig, T. 2009. Was Dick Whittington taller than those he left behind? Anthropometric measures, migration and the quality of life in early nineteenth century London?. *Explorations in Economic History*, 46: 120–131.

Juif, D. & 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 (2): 227-241.

Juif, D. 2015. Skill selectivity in transatlantic migration: The case of canary islanders in Cuba. *Revista De Historia Económica / Journal of Iberian and Latin American Economic History*, 33(02): 189–222.

Juif, D., & Quiroga, G. 2019. Do you have to be tall and educated to be a migrant? Evidence from Spanish recruitment records, 1890–1950. *Economics & Human Biology*, 34: 115–124.

Juif, D., Baten, J., & Pérez-Artés, M.C. 2019. Numeracy of religious minorities in Spain and Portugal during the inquisition era. *Revista De Historia Económica / Journal of Iberian and Latin American Economic History*, 38(1): 147-184.

Manzel, K., Baten, J., & 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.

Mincer, J. 1974. *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research.

Núñez, C. E. 1992. *La fuente de la riqueza: educación y desarrollo económico en la España contemporánea*. Alianza.

Quiroga, G. 2003. Literacy, education and welfare in Spain (1893–1954). *Paedagogica Historica*, *39*(5): 599–619.

Rodríguez, M. C. & Bennassar, B. 1978. Signatures et niveau culturel des témoins et accusés dans les procès d'inquisition du ressort du Tribunal de Tolède (1525-1817) et du ressort du Tribunal de Cordoue (1595-1632). *Cahiers du monde hispanique et luso-brésilien*, 31: 17–46.

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

Sánchez Alonso, B. 2007. The Other Europeans: Immigration into Latin America and the International Labour Market (1870–1930). *Revista De Historia Económica / Journal of Iberian and Latin American Economic History*, 25(03): 395–426.

Sarasúa, C. 2002a. Aprendiendo a ser mujeres: las escuelas de niñas en la España del siglo XIX. *Cuadernos de Historia Contemporánea / Departamento de Historia Contemporánea*.

Sarasúa, C. 2002b. El acceso de niñas y niños a los recursos educativos en la España rural del siglo XIX. In J. M. Martínez Carrión (Ed.), *El nivel de vida en la España rural, siglos XVIII-XX*: 549–609. Alicante: Universidad de Alicante.

Sarasúa, C. 2013. ¿Activos desde cuándo? La edad de acceso al mercado de trabajo en la España del siglo XVIII. In J. M. Borrás Llop (Ed.), *El trabajo infantil en España, 1700-1950*: 61–80. Barcelona: Icaria.

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

Tollnek, F. & Baten, J. 2017. Farmers at the heart of the 'human capital revolution'? Decomposing the numeracy increase in early modern Europe. *The Economic History Review,* 70 (3): 779–809.

Viñao Frago, A. 1999: Alfabetización y primeras letras (siglos XVI-XVII). In A. Castillo (Ed.): *Escribir y leer en el siglo de Cervantes* : 39-84. Barcelona: Gedisa Editorial.

Vincent, B. 1987. Lisants et non-lisants des royaumes de Grenade et de Valence a la fin du XVI siècle. In CNRS (Ed.), *De l'alphabetisation aux circuits du livre en Espagne, XVI-XIX siècle*, París, CNRS: 85-104.

2. Land inequality and numeracy in Spain during the 17th and 18th century[3]

Abstract:

We assess the relationship between land inequality and human capital at the end of the early modern period, focusing on individual-level evidence from Spain. Our main finding is that land inequality had already had a significant negative effect on the formation of human capital there in the late-seventeenth and eighteenth centuries. We argue that this reflects the important role of a social structure based on farming families (as opposed to latifundia and day laborers) in the development of numeracy. This is consistent with earlier studies, which argued that farming households could (1) maintain a relatively favourable nutritional standard as a precondition for cognitive skills, (2) limit child labour and (3) encourage numeracy due to its demand by farming activities. Our results are robust, as they include several control variables and potential confounding variables.

[3] Co-authored by Jörg Baten. He contributed approximately 20% of the work to this paper. This chapter is based on a paper published in *Historia Agraria revista de agricultura e historia rural.*

2.1 Introduction[4]

Recently, an agricultural dimension was added to the debate about the determinants of growth and obstacles to development from a long-term perspective. Galor, Moav, and Vollrath (2009) developed a model in which a stronger position for large landowners relative to industrial entrepreneurs prevents human capital formation and, consequently, economic development. In other words, the size distribution of agricultural holdings would have played a central role because the political incentives of large landowners made substantial investments in human capital less likely. While entrepreneurs benefited from the accumulation of human capital by the masses and thus, had an incentive to support public education, large landowners were not willing to pay taxes for primary schooling, for example. The result of this impasse had an effect on the pace of the transition from an agricultural to an industrial economy, contributing to unequal economic growth across countries. Baten and Hippe (2018) confirmed this theory and came to the conclusion that it was mostly the agricultural south and east of Europe where large landownership restricted human capital and investment around 1900. In England, France, as well as in the most industrial parts of the Habsburg Empire, however, this effect was not visible. For the nineteenth century United Kingdom, Clark and Gray

[4] The authors would like to thank the anonymous reviewers of *Historia Agraria* for their comments to improve this article. An early version of this paper was presented at the European Historical Economics Society Conference held at the University of Tübingen in September 2017. This research has benefited from comments made by Daniel Oto-Peralías, Rowena Gray, Carmen Sarasúa and Andrés Sánchez Picón. The authors also thank Pilar Erdozáin and Luis Garrido for sharing the data of Olite and Laujar de Andarax respectively. Mari Carmen Pérez-Artés acknowledges the funding from the Ministerio de Economía y Competitividad of Spain (HAR2017-85601-C2-1-P).

(2014) found no correlation between land inequality and literacy at a local level, as this was a country in which the industrial revolution started early. Consistently, for nineteenth century agricultural Spain, Beltrán and Martinez Galarraga (2018) used the census of 1860 and found that land inequality was negatively correlated with male education.

However, all this refers to nineteenth and twentieth century evidence, when the industrial revolution was well under way. Until now, no study has addressed this relationship for the early modern period, which is the main focus of this article. We assess the relationship between land inequality and human capital for the seventeenth and eighteenth centuries (including some limited evidence on the sixteenth century), focusing on individual-level evidence from Spain. Our main finding is that land inequality also had a significant negative effect on human capital formation for the early modern period. In early modern Spain, industrial development was negligible and educational investment was not very relevant for the majority of the population, hence Galor et al.'s (2009) theory for the nineteenth century does not apply here, as the authors mentioned[5]. What was the causal mechanism instead? Building on earlier studies, we argue that farming families provided a relatively favourable nutritional standard, so that their descendants could acquire human capital (Tollnek and Baten 2017; Baten et al. 2014). Moreover, farming parents were able to provide some basic skills at home. This was very relevant for the

[5] During the seventeenth and eighteenth centuries, the industry of Andalusia and the crown of Castile were typical of an agricultural economy with a low level of mechanization. The only two industrial Andalusian cities of the nineteenth century comparable to the Catalan or Basque provinces were Antequera and Linares (Parejo 2009). The Segovia textile industry or the royal textile factories in Castile are other examples of the Spanish industry during Old Regime (García Sanz 1996; Clayburn la Force 1964).

early modern period since families were the main agents providing education during modern times (Peña Díaz 2012).

To be more specific, farmers had advantages along four causal channels: Firstly, during crisis situations (the crisis of the seventeenth century, but also during short-term crises), direct access to nutrients was very important for the development of numerical skills. Malnutrition was more prevalent for agricultural sector workers who could not provide high quality food to their children, given that prices rose substantially during periods of bad harvests. The farmers, in contrast, could decide to consume more of their produce in their own households, even when high prices provided incentives to sell. This implies that farming households could access better nutrition in crisis periods, avoiding the numeracy deprivation that results from severe malnutrition, from a protein deficiency in particular (Baten et al. 2014). (2) In addition to relatively good nutritional access, many farmers' children were not burdened by child labour, whereas day labourer households depended on it, disincentivising schooling (Tollnek and Baten 2017). (3) Farmers were also more willing to invest in the skills of their children, as they would need them to run the farms, whereas the demand for skills by agricultural labourer parents might often have been lower (Beltrán Tapia and Martinez-Galarraga 2018)[6]. We will also study below whether other social groups imitated the farmers in regions with a high farmers' share, although the evidence on this will be indirect. (4) Towards the end of the period in particular, the elites who owned land were concerned that "excessive" education of the poor would make them abandon manual labour. In the regions dominated by large-scale agriculture, the wealthy actively hindered school attendance of the lower income groups

[6] Furthermore, although the quality of formal instruction was poor, the children of the farmers had more stable schooling over more years, even if we take the months of absenteeism due to the cycles of agricultural tasks into account (Borrás Llop 2002b).

(Kagan 1981). Only a few villages received school donations from pious or charitable *Señores* (lords) which then benefited lower income groups. However, this was the exception rather than the rule. Kagan (1974: 25) claimed that: "Consequently, Spain's peasantry, too poor to support a schoolmaster, too hard working to take time out for classes, remained overwhelmingly illiterate until the opening years of the twentieth century." As such, it seems reasonable that in areas with a lower number of landless peasants, the farmers and *Señores* decided to invest more in education.

In this article, we focus on the determinants of numeracy in early modern Spain. Due to more detailed sources (we have more evidence on Andalusia and no evidence on the Northwestern coast and Catalonia) we pay particular attention to Andalusia (Figure 2.1). Evidence on the sixteenth century covers two Andalusian provinces (Cordoba and Seville), while for the seventeenth and eighteenth centuries we have a broad mix of Spanish regions. Andalusia was an economic and urban centre during this early period; indeed, it was one of the most urbanized regions across Europe[7]. This region is the southernmost point of mainland Europe and, with more than 87,000 square kilometers, its area is larger than several European countries (Parejo 2009: 11)[8]. Another important fact is that Andalusia benefited from the accumulation of colonial traffic with America; it was the starting point of the trade with the New World. Seville particularly enjoyed its monopoly in trade with America from the sixteenth century, until it was overtaken by Cadiz in 1717 (Marcos Martín 2000).

[7] The current term Andalusia comes from the territorial reform of 1833, when this domain included the Kingdom of Granada. Previously, it referred to the Kingdoms of Cordoba, Seville and Jaen, incorporated into the Crown of Castile in the thirteenth century (Parejo 2009).

[8] Andalusia is larger than Ireland, Luxembourg, Denmark and Belgium.

A new dataset from *padrones* (local nominative population censuses) and Cadaster of Ensenada is analysed here[9]. Table A1 offers a description of the sources. Although more than half of our sample is from Andalusia, we were able to include other regions in Spain to obtain more representative results. We weight our analysis, below, to give the Central and Northern regions their appropriate influence on our results. The sample covers the period from 1580 to the middle of the eighteenth century. Our sample is composed of 26,851 individual observations mentioning age, of which 17,145 also contain occupational data. This evidence allows us to provide a long-term perspective of land inequality and numeracy.

As far as we are aware, until now, no individual-level analysis on this topic exists for early-modern Europe. Only in Spain, and in Andalusia in particular, were local censuses containing both ages and occupations taken from as far back as the sixteenth century. We use the inequality proxy suggested by Clark and Gray (2014) as our main explanatory variable. This proxy is based on the idea that in regions where large estates were prevalent, the agricultural workforce mainly consisted of agricultural laborers who did not own farms and were not called "farmers." In contrast, in regions of small and medium sized farms, farmers represented a high share of the agricultural workforce[10].

We use age-heaping-based estimates of numeracy for the dependent variable. The underlying methods were developed in the last decade, especially for societies and periods where sources of other education indicators were incomplete. Numeracy, or the

[9] The *Catastro of Ensenada* (1750-1756) is the name given to the investigation carried out in the territories of the Crown of Castile on the property and income of the householders, as well as on their family and servants (Camarero Bullón 2002).

[10] This proxy has also been used in the nineteenth century study of Beltrán Tapia and Martinez-Galarraga (2018).

ability to deal with numbers, allows us to obtain a more comprehensive sample from early modern Spain. Age statements can be found in a greater number of sources than alternative measures of human capital (A'Hearn et al. 2009). This proxy has also been used by Álvarez and Ramos Palencia (2018) to assess the relationship between human capital and male labour earnings in Spain for the provinces of Palencia, Guadalajara and Madrid. They found that numeracy had an influence on earnings, supporting the relevance of numeracy among economies in early modern Spain. The relationship between numeracy and economic growth is even stronger than that for school enrolment or literacy, as the recent economic growth literature has shown: Hanushek and Woessmann (2012), for example, argued that math and science skills were crucial for economic success in the twentieth century. They concluded that numerical skills matter the most for economic growth by considering cross-country evidence as well as the success of migrants from various countries to the U.S., for example.

The remainder of the paper is structured as follows: section 2.2 introduces the historical context of land inequality and human capital in modern Spain. Section 2.3 follows with the explanation of the methodology and the data used in this study. In section 2.4, our empirical results and descriptive analysis are presented. Section 2.5 presents the conclusions.

2.2 Land inequality and human capital in modern Spain

2.2.1 The origin of land inequality

The agrarian reform law of September 1932 blamed "the *latifundium* for the backwardness and the pitiful conditions of the workers in the countryside" (Gónzalez de

Molina 2014: 28)[11]. However, Carmona and Simpson (2003: 19) argued that these institutions were not the cause of the low levels of production and productivity, rather that *latifundia* "reflected" the low level of development in agriculture. Our study combines these views by studying whether regions dominated with farming households displayed higher levels of numeracy.

We first provide some detailed background on Andalusia, the region for which we have ample evidence. We later discuss the (often middle-sized) farm characteristics of central and northern Spain. The structure of landownership in Andalusia was characterised by, on the one hand, a large share of large landholdings in the kingdoms that had been incorporated into Castile in the thirteenth century and, on the other hand, a majority of small properties in the Kingdom of Granada (Parejo 2009). In the Guadalquivir valley, large landownerships were predominant. As early as the fourteenth century, the nobility was interested in these lands and accumulated them in a regimen of large properties, being fully consolidated by the middle of the eighteenth century. Both the high nobility and the lower regional nobility owned very large estates in municipalities of the Guadalquivir riverside (Mata Olmo 1984). On the other hand, in Granada and Almería, the formation of *latifundia* began later and was restrained by the mountainous terrain of the area. This was favourable for small and medium farmers and for the repopulation after the *Morisco* uprising of 1568-70[12]. After this event, the state

[11] *Latifundia* refer to the large private farms in the south of Extremadura, Castile and the Guadalquivir Valley. Apart from the predominance of large rustic patrimonies and *latifundia*, the irrigated agriculture of the interior of Andalusia and Murcia and the production of wine regions of the south required a large workforce and therefore, of a large number of day labourers (González de Molina 2014).

[12] This uprising had its precedent in January of 1567, when a royal law obliged all "*Moriscos*" (Muslims forcibly converted to Christianity) to become "real" Christians within a year. This

distributed additional plots to Christian settlers and also prevented the accumulation of *latifundia* [13].

Warlords, nobles, clergy, religious orders and the church were the main beneficiaries of the Castilian conquest of Al Andalus. From the thirteenth century onwards, the concentration of landownership increased due to the purchase and sale of land by privileged groups, such as titular nobility and urban merchant classes. This tendency also justifies the origin of the day labourer (*jornaleros)* in Andalusia. Although, in the east, large properties were less represented in rural areas, day labourers made up the majority of the population on the Mediterranean coast (Arenas Posadas 2016). During the modern era, the power of rural elites increased. These elites originated in the lordships that were granted during the reign of the Catholic Monarchs, especially related to the conquest of Granada. These oligarchs were enriched through the accumulation of land, leases and cereal specialisation. Whether through economic, family or political ties, wealthy farmers had access to the privileges of the nobility. During the reigns of Charles V (1516-1556) and Philip II (1556-1598), the local lords and oligarchies usurped communal lands in southern Spain that had been fundamental for the subsistence of the peasant economies. Day labourers suffered from long working days and low wages (Peña

episode, also known as the war of the Alpujarras, is the last episode of the Islamic and Christian conflict that lasted almost eight centuries. This rebellion ended with a massive deportation of all previously Muslim families of the Kingdom of Granada. In 1609, the expulsion of the last *Moriscos* from Spain took place (Andújar Castillo 2004).

[13] Calculating the share of agricultural area relative to the total area, the lowest proportion was only 35% in the Kingdom of Jaen while Seville and Cordoba had 59% and 57% respectively. 61.8% of the Kingdom of Granada was agricultural due to the better utilisation of land caused by a more rational division of land than in the Guadalquivir Valley (see more on Artola et al. 1978).

Díaz 2012) and, by the end of the eighteenth century, the nobility, the church and municipalities owned most of the land (Carmona and Simpson 2003)[14]. Later on, during the nineteenth century, landless peasants still had to endure the poor conditions of income and labour, while rich landowners found enough workers for their estates (González de Molina 2014). However, in the nineteenth century, land accumulation decreased among the privileged classes of the old regime, and during the next century, the predominant landowner class of the southern *latifundia* began to lose political prominence in state government as well as at the regional and local levels (Mata Olmo and Naranjo-Ramírez 1997).

As stated above, two and a half centuries separated the conquest of Lower Andalusia and the Kingdom of Granada, which led to some institutional differences between both territories. Furthermore, after the Conquista, the repopulation of the *Bético* valley mainly consisted of people coming from northern Spain, whereas the one of the Granada region was administered by the western Andalusians[15]. Another peculiarity was the presence of a substantial Muslim community in Eastern Andalusia. Even after the expulsion of the Muslims, the socioeconomic and institutional reality in the Kingdom of Granada was different from the prevailing situation in Jaen, Cordoba and Seville. In the eighteenth century, the dissimilarities within Andalusia were also visible in economic indicators such as the ratio between the number of day-labourers and farm owners. The share of farmers (*labradores* and *hortelanos*, taking only males) relative to the total

[14] For central Spain, Santiago-Caballero (2011) demonstrated that the income inequality among grain producers decreased in Guadalajara at the end of the eighteenth century. It was due to the possibility that small peasants had to increase the size of their lands as a result of the redistribution of common lands privatized by the central government.

[15] The term *Bético* refers to the provinces of Cordoba, Seville, Huelva, Jaen, and Cadiz.

number of male occupations (males, age 25+) in 1785-87 according to the census of *Floridablanca* is, in this sense, quite heterogenous between Andalusian regions: in the provinces of Almería and Granada, this farmers' share was as high as 42% and 31% respectively; it was 24% in Jaen, and 20% in Malaga, 14% in Seville and 18% in Cordoba and a negligible 6% in Cadiz[16]. In central and northern Spain, the farmers' share was much higher: Navarra and Guadalajara had the highest shares – 84% and 64%. However, the farmers' share is not just a correlate of the north-south differences. For example, La Rioja had a relatively low farmers' share (29%), a rate that was below that of provinces such as Almería, Caceres (42%) and Badajoz (40%). Here and in the following we distinguish between "provinces" and "regions", the latter comprising several provinces (the regions are visible in Figure 2.1, provinces are compatible to today's provinces).

2.2.2 Human capital in Spain since the 16th century

A widely used indicator for studying human capital in pre-census periods has been literacy. Several studies used the ability to sign as a proxy for literacy (Delgado Criado 1993; Viñao Frago 1999). The presence or absence of signatures on documents was considered the only direct evidence for measuring education levels. Only in 1797, with the census of Godoy, direct data about the schooling process became available. Finally, in 1860, a Spanish census included information about the ability of inhabitants to read

[16] Ponsot (1986: 28) studied the distribution of the property for 17 municipalities in western Andalusia by the middle of the eighteenth century. Only in two cases were found that small and medium-sized owners had some relevance (Espartinas and Montilla located in Seville and Cordoba) while the major owners were the majority (for example, in Carmona and Medina Sidonia in Seville and Cadiz).

and write for the first time. Núñez (1992) studied and analysed the relationship between human capital and economic growth in contemporary Spain, exploiting this source.

The ability to sign has allowed researchers to estimate levels of literacy for different regions of Spain. However, one of the problems with this method is the representativeness of the available samples; the types of sources (fiscal sources, testimonies, marriage records, death record etc.) are usually not uniform for all regions or even within the same location for different years. In addition, the same sources often have different levels of representativeness; for example, sometimes the samples represent only the wealthier and presumably more educated social strata which makes it difficult to reach conclusions (Viñao Frago 1999). These studies reveal that the capacity to sign increased during the sixteenth century, but decreased again in the seventeenth century (Viñao Frago 1999). Rodríguez and Bennassar (1978) studied the interior Andalusian regions of Andújar, Iznatoraf, Úbeda and Cordoba using the testimonies of the accused by the inquisition. Vincent (1987) used fiscal sources and assessed the literacy of the *Moriscos* in Granada in 1570. Literacy in Cadiz has also been explored by de la Pascua Sánchez (1989) using wills during the late seventeenth century.

Throughout our period, the family was the main agent providing education either because they could afford to pay a teacher or if within the family one member knew how to write and read (and probably basic numeracy skills) was in charge of teaching the rest (Peña Díaz 2012). There was substantial numeracy in Spanish farm households before the widespread introduction of schooling, hence the acquisition of numerical skills could only have happened in the family and the household (Tollnek and Baten 2017; Borrás Llop 2002a; Álvarez and Ramos Palencia 2018). Only very few families could afford a teacher during the early modern period. For the children of the poorest neighbourhoods,

the local communities and parishes sometimes paid an annual amount to a teacher, but schools were few. Moreover, the control over the training of teachers would not begin in Andalusia until the beginning of the eighteenth century (Peña Díaz 2012).

During the eighteenth century, when local communities in some parts of Europe paid for teachers and schools, the large Andalusian landowners were not interested in paying taxes to promote education for their day labourers. As Arenas Posadas (2016: 375) has argued:

> "illiteracy and the absence of training contribute to immobilizing the labour force in the territory, thus promoting the excess of labour and, consequently, low wages".

Apart from low wages, day labourers had to face times of unemployment due to bad weather or times when there was no work in agriculture (Bernal 1987; Carmona and Simpson 2003). This is consistent with the findings of Álvarez and Ramos Palencia (2018) for Guadalajara, Madrid and Palencia where human capital (literacy and numeracy) influenced male labour earnings during the eighteenth century. In contrast to Denmark, which developed a human capital-intensive form of agriculture, the proximity of owners to agricultural production was not given, in addition to a number of other differences[17].

Andalusia did not reach levels above 30% of literacy until the twentieth century. The western provinces, rural areas and the female population had the lowest literacy rates (Arenas Posadas 2016: 351; Sarasúa 2002). At the national level, in 1900, Andalusia held

[17] In the case of Andalusia, large absentee landowners owned extensive properties in the South (Carmona and Simpson 2007). Although absenteeism could be an obstacle to promoting human capital, it does not seem to affect agricultural production (Simpson and Carmona 2017).

an average position in terms of literacy; but by 1950, it had dropped to the lowest level in all of Spain (Arenas Posadas 2016: 352).

2.3 Methodology and data

The regions considered in this research are illustrated in Figure 2.1. Table 2.1 specifies the number of observations by province and period. Table 2.6 in the appendix contains a description of the sources[18].

To measure land equality, we use the ratio between the number of farmers and the overall agricultural population, an indicator suggested by Clark and Gray (2014). Our definition of farmers depends on the contemporaneous naming of occupations. "Farmers" (*labradores)* were not only those who owned land, but also those who rented land and ran a farm of a substantial area. Hence, a day labourer (*jornalero*) who was usually not possessing or controlling land, would not be identified as farmer by contemporary census takers (Tollnek and Baten 2017). Although quantitatively almost irrelevant, we also include "*hortelano*" in the same category as farmers, since they usually also had some control over plots of land that were intensively farmed and they could provide better nutrition to their children in crisis situations[19]. Although hortelanos were obviously not farmers, we included them for simplicity in the variable "farmers' share" (justified by their small number). In order to assess the plausibility of the farmers' shares based on our

[18] Within these sources, we analysed a convenient sample and we took care not to select only special groups.

[19] The difference between "*labrador*" and "*hortelano*" lies in the type of land they own. For the former it was rain-fed for the latter it was irrigated (Bermúdez Méndez and Martín Chicano 2007).

sample, we can calculate a similar farmers' share for the Floridablanca census (even if the Floridablanca census was recorded somewhat later, in 1785-87). The correlation is very strong (Figure 2.2, aggregated on province level). A large share of both our-sample-based farmers' shares and the Floridablanca-based farmers' shares are in the 20 to 40 percent range. Our sample is slightly more urban (hence a lower farmers' share for Sevilla, for example) and more Andalusian. This difference is mostly compensated for by our weighting procedure.

In order to assess numeracy, we employ the "age heaping" methodology using the ABCC index[20]. This method considers the share of individuals who are able to state their precise age in years, in contrast to those who report an age rounded to a multiple of five. For instance, an individual could state "I am 45" when he or she is 44 in reality, but did not know it exactly. Numeracy and literacy are robustly correlated, though basic mathematical skills diffused earlier than literacy. In addition, the potential biases caused by counting cultures and the institutional settings of censuses have been thoroughly discussed throughout the numeracy literature, but the results did not invalidate the age heaping method (Tollnek and Baten 2017). Accordingly, we can argue that, just as signature rates in official documents, despite their limitations, can serve as proxy for basic literacy (Reis 2005; Rodríguez and Bennassar 1978), age heaping can serve as a proxy for basic numeracy.

The ABCC index is a simple linear transformation of the Whipple index (1), derived by A'Hearn et al. (2009). The ABCC index (2) allows for an easier interpretation and yields an estimate of the share of individuals who state their age precisely:

[20] The term "ABCC" results from the initials of the authors' last names plus that of Gregory Clark, who commented on their paper.

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

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

This index ranges from 0 to 100, where 100 indicates no heaping patterns on multiples of five; meaning that the entire society has skills in basic numeracy. The age groups we use are in increments of ten years; 23 to 32, 33 to 42 etc. We omitted the age range 63 to 72, as this group offers relatively few observations, especially for the seventeenth and eighteenth centuries when mortality was relatively high (Schofield and Reher 1994). Crayen and Baten (2010) analysed age effects carefully and found that they do not have a strong influence once the birth cohort effect is controlled for: older individuals may round more strongly, but mostly because they were born earlier. The only exception is the youngest group, age 23-32, which needs an adjustment of 25% that we calculated in our sample (Crayen and Baten 2010)[21].

While the ABCC index refers to averages of groups (by region and birth decade, for example), it is also possible to analyse the likelihood of individuals to report a rounded

[21] Moreover, a potential bias could result from counter-checking by the officials who collected the local censuses. We looked at each source by itself to assess whether numeracy was close to 100 percent in local communities and times in which this could not be expected. This phenomenon of counter-checking occurred in some Russian and Korean sources, for example, as described by Baten, Szołtysek and Campestrini (2017) as well as Baten and Sohn (2017). They therefore decided to discard a part of their sources. In Spain, government officials were not counter-checking sources to the same extent, as we do not observe this phenomenon of numeracy being very close to 100 percent.

age. This can be done by assigning the binary variable "numerate" which is coded as 1 for those who report an unrounded age and 0 otherwise (Juif and Baten 2013; Tollnek and Baten 2017). The binary variable can be analysed with Logit or Probit regression models or by using a linear probability model (LPM) with heteroskedasticity-robust standard errors. For the result to be interpreted in ABCC-values under the LPM, it needs to be multiplied by 125 (by 100 to move from a fraction between 0 and 1 to a percentage, and by an additional 25 to account for the fact that 20% of the population actually do have ages ending in 0 or 5).

How representative is the sample? Fortunately, the availability of evidence in Spain resulted in a quite widespread geographic distribution (Figure 2.1). Most regions can be covered in the seventeenth and eighteenth centuries, except the northwestern coast and Catalonia. We have more observations on Andalusia, but we can adjust this overrepresentation by assigning smaller weights to Andalusian observations and larger weights to the other provinces (see the notes in Table 2.3 for details). Socially, our local censuses are quite representative, because they include all social strata, as can be seen from the occupational information. We also took care that we did not only record a special effect in the Cadaster that might have reflected a special sub-population (such as the nuns in a monastery or the merchant quarter of a city, for example). We have rather drawn samples that cover various parts of cities and villages, if the archival situation allowed us to do so. As a definition, we will call cities and villages "local communities" in the following. In general, we distinguish between local communities, provinces and regions (as in Figure 2.1).

Finally, is the population of each local community sufficiently covered by at least some observations? We calculated the approximate share of our sample, relative to the

total population in the earliest reliable census, the Floridablanca census (1785-87)[22]. As a result, in only 10 local communities, our sample represented less than 10% of the total population older than 25 years of age, while for 48 local communities we could obtain more than one tenth of the overall population (see Table 2.7 in the appendix)[23]. As there were differences in the archival survival rates in various local communities, we needed to weigh the samples in order to obtain regional representativeness anyways.

Finally, we analysed whether the observations for which we have occupations and those for which we do not have occupations are comparable. The numeracy index of those with occupations was 64.3 and the one without occupations was 66. Hence the numeracy index difference is only 1.7 points, which is a very small difference that can easily be caused by composition effects.

[22] Using this census, we calculated the inhabitants who were more than 25 years old (given the way in which the Floridablanca census aggregates the information, it is not possible to take it from 23 years of age) by local community. We divide the number of persons in our sample by the census total, even if our sample refers to an earlier period. Due to the lack of reliable census sources for occupations in the sixteenth, seventeenth and early eighteenth century, it is not possible to obtain reliable census totals per local community for earlier periods.

[23] The ten cases of less than 10% refer mostly to Andalusia, for which we have overall a very high number of observations anyways. In other words, if we would have a 10 percent share for these Andalusian local communities, our regional representativeness would actually be smaller. The same is the case for the urban share – our sample has slightly more urban cases than the general Spanish population, hence we would have a less representative sample, if Écija, Córdoba etc. would be presented by a 10% sample.

2.4 Descriptive analysis and regression results

Table 2.2 shows the descriptive statistics. The mean of the variable "numerate" in our sample is 0.57, which indicates that slightly less than a half of our sample reported an age ending in 0 or 5. The mean farmers' share, which is our main explanatory variable of interest in this study, is 0.33, with a standard deviation of 0.27, defined as fraction of occupations between 0 and 1.

In order to assess the influence of the farmers' share on numeracy, we performed logit and linear probability model (LPM) regressions. The LPM is described in the following equation, which applies similarly to the logit model.

$$Numerate_{itr} = \alpha + \beta_1 Farmershare_{tr} + \beta_2 Farmer_i + \beta_3 age23\text{-}32_i + \beta_4 age43\text{-}52_i + \beta_5 age53\text{-}62_i + \beta_6 City_r + \beta_7 Female_i + \mu_r + \gamma_t + \varepsilon_{itr}$$

$_i$ indicates each respective individual, $_t$ indicates the decade of birth and $_r$ denotes the region in which the individual was born at the local community level. The variable to be explained is *numerate*, coded as 0 when age is stated as a multiple of five, and 1 otherwise. *Farmershare* is the proportion of farmers in the agricultural sector of our sample and *Farmer* is a dummy for farmers. *Age23-32$_i$* corresponds to the group of individuals aged between 23 and 32, following the same idea for *Age43-52$_i$* and *Age53-62$_i$*. *City* is a dummy for cities with more than 20,000 inhabitants according to the *Floridablanca* census carried out in 1787 and *Female* is a dummy for females. The model includes region fixed effects (μ_r) that reflect the historical regions in Spain from Figure 2.1. We also control for time fixed effects (γ_t), using half-century periods from 1580 to 1760. Finally, the equation allows for a constant term (α) and an error term (ε_{itr}). The

model is also weighted by the proportion of inhabitants by historical regions in the *Aranda* census (1768)[24].

To measure the effect of farmers' shares on numeracy, based on the occupational information for 17,145 cases, we calculated the farmers' share of each local community and period. Our inequality data provide 117 observations combining local communities and birth centuries. We then assigned this farmers' share in a given local community and century to all 26,851 individuals: We include all cases where age is reported, even if occupational information is not contained for each individual, but for a sufficient number of occupations in a specific local community and birth century.

Table 2.3 shows the results of the effect of farmers' shares on numeracy. We cluster the observations at the local community and birth decade level. Weights establish representativeness for the regions included in columns 2 and 3, but there is not a substantial difference to the unweighted regression in Column 1. Columns 1 and 2 include both males and females. In the last column, we only analyse the males of our sample. We control for the characteristic of being a farmer and different groups of age[25]. Interestingly, if we include the inequality proxy "farmers' share" the farmer coefficient by itself does not show a significant difference, relative to persons who are not farmers[26]. Consequently,

[24] See note Table 2.3.

[25] Following Reher (1994) for seventeenth and eighteenth century, we categorise the region as rural for local communities with less than 5,000 inhabitants, urban with more than 5,000 and city with more than 20,000 inhabitants. Unfortunately, we cannot control for local community fixed effects, as this would move the focus to the modest variation over time, which would seem less reliable – considering potential measurement error – compared to the substantial cross-sectionals variation in our sample.

[26] Some of the coefficients for higher ages are statistically significant and negative, which might be either caused by the fact that people tend to forget their ages when they reach their 50s and

we conclude that the social structure in regions with high farmers' shares also affected numeracy beyond the farmer group itself. The only logical explanation for this are external effects: people with other occupations (for example, craftsmen and skill-intensive services) who lived in regions dominated by farmers behaved more similarly to (and perhaps imitated) farmers, compared to craftsmen and others in regions not dominated by farmers, but by agricultural day-laborers and *latifundia*: The ones in the farmer-dominated regions also invested more time in their offspring, sent their child less often to work, and provided slightly higher quality of nutrition than in the latifundia regions. We do not have direct qualitative evidence on this, but presenting this indirect quantitative evidence on these external effects is already interesting.

As a caveat, we note that the number of cases in our individual-level regression should not be taken as proof of high reliability, as the explanatory variable "farmers' share" varies by local community and century. Nevertheless, in all specifications, our equality measure farmers' share had a large positive impact on numeracy. The variable "city" never appears significantly correlated. In this analysis, women do not have a significant disadvantage once we control for farmers' share. It should be taken into account that mothers had a very important role in farming households (Tollnek and Baten 2017). Table 2.4 performs the same analysis in a logit model. The results are nearly the same as those obtained in the LPM. R-Squares are generally low, suggesting that at the individual level a substantial random variation accounts for large part of the overall variation. However, the p-value of significance suggests that the farmers' share has a substantial influence.

60s years of age, or by the fact that they were born in earlier birth decades. The research by Crayen and Baten (2010, Appendix) suggests the latter.

To test whether the results are potentially driven by a small number of outliers, we construct a residual plot by regressing numeracy on the most important explanatory variables (city, female, and century fixed effects figure 2.3). In a second step, we regress the main explanatory variable of interest, the farmers' share, on all of these variables except numeracy. In both steps we saved the residuals, of numeracy and farmers' share, respectively. These can be interpreted as the residual value of both variables, after removing the influence of the other explanatory variables. In order to make it easier to interpret, we aggregate all locations at the provincial level and century. For example, our evidence on Cuenca, Soria and Avila had a high land equality (indicated by the high residual farmers' share) in the eighteenth century, and at the same time a high residual numeracy. In contrast, eighteenth century Cadiz, Jaen and Cordoba had both low residual land equality and numeracy[27]. Outlying observations to the upper left were Seville, Madrid and Navarra: residual numeracy was higher than expected based on land inequality. For Seville and Madrid, the urban effect might be particularly important and not be fully captured by the large-city-dummy variable (which was also assigned to smaller urban centres)[28]. Murcia had a relatively low level of numeracy in spite of its comparatively high land equality (but it should be noted that Murcia is only represented by Lorca). This might be caused by the difficulties in maintaining Murcia´s irrigation agriculture in the eighteenth century due to the lack of water and due to privatization during the seventeenth and eighteenth centuries. Concentration and privatisation affected not only day labourers, but also farmers in Murcia. Only the landlords from the capital,

[27] There is a high intertemporal persistence, as Beltrán Tapia et al. (2018) found for the nineteenth and twentieth centuries that the lowest numeracy indices were also in the Andalusian provinces.

[28] For Navarra, we cannot exclude the possibility that the sample is too small to yield a reliable estimate.

who received regular payments from their tenants, benefited from it (Pérez Picazo and Lemeunier 1985). However, in sum, we observe that residual numeracy strongly corresponds with the residual farmers' share.

We also considered endogeneity and a potentially confounding role of skill-selective migration (Appendix 2.8.2). Both these potentially confounding factors appear to have only a very modest influence on the results.

How large are the numeracy differences between farmers and agricultural labourers individually? While we already included a farmer variable in the previous regression comparing farmers with non-farmers, here we are interested in the differences between farmers and day-labourers, as well as the differences between other occupational groups and day-labourers. Hence, in the first column of Table 2.5 we test the difference between being a day-labourer and having a non-agricultural occupation or being a farmer. In the first column we include region fixed effects. In the second, we use fixed effects for each local community. In both models, time fixed effects are also considered. In both cases, the coefficients of numeracy for the farmers are significantly positive. In other words, we observe that the difference in numeracy between farmers and day labourers was 7.1 percentage points in the first specification and 4.8 in the second, which controls for local community fixed effects. This result is smaller, but with the same sign as in Catalonia in the eighteenth century, where the farmers had a 14 percentage point advantage (Gómez-i-Aznar 2019). *In sum, the agricultural day-labourers had a much lower numeracy level than the non-agricultural occupations* (i.e. services and crafts).

How did these numerical differences develop over time? Figure 2.4 portrays the numeracy trends by occupation groups for the sixteenth to the eighteenth century. The sixteenth century evidence cannot be directly compared in level terms, because we have

only three Andalusian regions for the sixteenth century. But the relative numeracy ranking of occupations might still be interesting: farmers, day labourers and other occupations had much lower numeracy in sixteenth century compared to the seventeenth century across Spain. Moreover, for the seventeenth and eighteenth centuries, we have evidence on all regions. We observe that the farmers started at the same level as the day labourers in Andalusia in the sixteenth century. For the regionally broader data of the seventeenth century, numeracy was much higher for all occupation~~'s~~ groups. Farmers and day labourers both still had quite low numeracy. By the eighteenth century, farmers almost reached the level of tradesmen, craftsmen and workers in administration. The gap in numeracy between farmers and the rest of the agricultural sector confirms earlier research about inequality in Spain by Alvarez-Nogal and Prados de la Escosura (2013), who found an increase in Spanish inequality (and land rent to wage ratios) from the early sixteenth century, after the Spanish medieval economy, with its strong urban and pastoral elements, disappeared (see Santiago-Caballero 2011 on Guadalajara).

The final question is whether the farmers' share remained stable over time, increased or declined. We can only trace this trend for all three centuries for Cordoba and Écija, located in Andalusia, where occupation was reported systematically for all the three periods. We observe that the farmers' share fell from around 18 percent to 2 percent between the sixteenth and eighteenth century (Figure 2.5). Clark and Gray (2014) argued, that this indicator proxies equality, hence we observe a strong increase in inequality, but with some caveats in this case: in two cities, the outskirts had a substantial share of farmers in the early period, but this phenomenon vanished over time as farmers disappeared in the larger towns, according to our evidence. Whether a similar decline from a higher starting point occurred, as in Cordoba and Écija, cannot be assessed for lack of evidence. To the extent that Cordoba and Écija are representative, this might

reflect a tendency of declining farmers' shares in Andalusia overall. Bernal (1987: 3) has shown that the number of day laborers for a sample of 20 local communities in Seville represented 54% of the workforce in 1620, increasing to 70% in 1754. By the end of the eighteenth century, this group would be 78%, on average, for the four Andalusian kingdoms, reaching their maximum in Seville and Cordoba. It would be one element implying slower numeracy progress in this region, relative to other European regions[29].

2.5 Conclusions

We conclude that *the land equality indicator "farmers' share" always had a significant positive effect on regional numeracy*. We also observe higher numeracy among farmers in the eighteenth century than among agricultural workers.

We argue that this relationship can be explained by the behaviour of (often middle-sized) farm households and the social structure in the regions dominated by these. Earlier studies emphasised advantages of farm households via four causal channels. Firstly, during crisis situations, farmers could benefit from their control over nutrients. This was very important for the development of numerical skills among their children. Agricultural sector workers could not provide high quality food to their children, especially not in crisis years, hence the children suffered from severe protein malnutrition (Baten et al. 2014). Apart from relatively good nutrition, some farmer children were not burdened with child labour, whereas day labourer households depended on child labour,

[29] In a much later period, the share of landless workers declined again. According to Carmona et al. (2019), the relative number of landless workers declined between 1860 and 1930. This was partly due to the falling ratio between land prices and rural wages and partly because of the exodus of the rural population to the cities.

inhibiting schooling (Tollnek and Baten 2017). Farmers were also more willing to invest in the skills of their children, as they would need them to run the farm, whereas the demand for skills by agricultural labourer parents might often have been lower. Finally, especially towards the end of the period, the elites who owned land prevented investment in the education of the poor. These hypotheses about farmer behaviour are consistent with the results of our study, as we find a consistently positive impact of the farmers' share. In contrast, comparing the farmers with all other occupational groups in the same regression, we do not find a significant farmer coefficient (only relative to day-laborers, farmers were more numerate). Consequently, the social structure in regions with a high farmers' share apparently also affected numeracy beyond the fact that some people were farmers. The only logical explanation for this are external effects: people with other occupations (for example, craftsmen and skill-intensive services) who lived in regions with a high farmers' share imitated (or behaved similar to) farmers, investing more time in their offspring's numeracy, requiring less child labour of them, providing slightly higher quality of nutrition than in other regions. We do not have direct qualitative evidence on this, but presenting our indirect quantitative evidence on these external effects can be considered a first step to gain insights on this externality.

This also has wider implications for understanding the history of world inequality. Scheidel (2017) describes the process of growing inequality in world economic history as follows: Farm size distribution played an important role. On one hand, kings and other rulers were interested in having a large share of farmers with medium sized plots, because their second and third sons were often recruited into the military. On the other hand, the nobility and others among wealthy social strata were keen on increasing their landownership and often forced small and medium sized farmers into servitude or agricultural labour and took over the land. A similar struggle can be observed for Spain

during the *Reconquista*. In the western and north western Andalusian territories, the nobility and similarly interested religious orders succeeded in allocating a large share of the land to their own *latifundia*. In contrast, in central Spain and the south-eastern Kingdom of Granada, which was only conquered after a long period of peace, the Spanish Crown succeeded in distributing most of the land to medium and small farmers and later protecting them against the nobility which might have otherwise expropriated the land (Oto-Peralías & Romero-Ávila 2016).

We add an economic process to this mechanism: the reduction of the share of small and medium farms retards human capital formation and hence impedes economic development. Therefore, the struggle between the ruler and medium sized farm owners on the one hand and the nobility on the other not only had a military consequence but an economic one as well.

For the example of Spain, as late as the first half of the twentieth century, less than 1% of holdings accounted for 57% of the area in Western Andalusia (Carmona and Simpson 2007: 348). Although after the Spanish Civil War (1936-1939) the active agrarian population began to decline in Spain, the provinces with *latifundia* in Andalusia continued being the ones with the greatest number of day laborers (Bernal 1987: 4); at the same time this was the region with the lowest literacy share (Arenas Posadas 2016: 352).

Our findings might also add an important notion to the investment issue in the late nineteenth century, as human capital differences tend to be persistent over time (Baten and Juif 2014): Physical and human capital tend to be complementary (Galor et al. 2009). The lack of numeracy in unequal regions might have reduced the profitability of physical capital investment due to this complementarity.

In sum, Spain can provide the most solid insights into the farmers' share and numeracy relationship, because it is the only country of the world for which occupations and ages are reported in local censuses for repeated years of the early modern period. We have evidence for the sixteenth, seventeenth and eighteenth centuries that allowed for the analysis of the effect of farmers' shares on numerical characteristics of the population. This certainly provides intriguing insights for Spain, but also more general conclusions about the role of farmers' shares in human capital formation throughout world economic history.

2.6 References

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

Àlvarez-Nogal, C. & Prados de la Escosura, L. 2013. "The Rise and Fall of Spain (1270-1850)", *Economic History Review,* 66(1): 1-37.

Álvarez, B. & Ramos Palencia, F. 2018. Human capital and earnings in eighteenth-century Castile. *Explorations in Economic History*, (67): 105–133.

Andújar Castillo, F. 2004. La cuestión morisca: de la General Conversión a la guerra y el destierro, *Andalucía en la Historia,* Nº4: 16-21.

Arenas Posadas, C. (2016). *Poder, economía y sociedad en el sur. Historia e instituciones del capitalismo andaluz.* Fundación Pública Andaluza Centro de Estudios Andaluces.

Artola, M; Bernal, A.M., & Contreras, J. 1978. *El latifundismo, propiedad y explotación, SS. XVIII-XIX*, Madrid, Mº de Agricultura.

Baten, J. & Hippe, R. 2018. Geography, land inequality and regional numeracy in Europe in historical perspective. *Journal of Economic Growth,* 23 (1): 79–109.

Baten, J.; Crayen, D., & Voth, HJ. 2014. Numeracy and the Impact of High Food Prices in Industrializing Britain, 1780–1850. *Review of Economics and Statistics,* 96 (3): 418–430.

Baten, J.; Juif, D. 2014. "A Story of Large Land-Owners and Math Skills: Inequality and Human Capital Formation in Long Run Development, 1820-2000", *Journal of Comparative Economics,* 42(2): 375- 401.

Baten, J.; Sohn, K. (2017) "Numeracy in Early Modern Korea, Japan, and China: the Age-Heaping Approach", *Japan and the World Economy 43:* 14-22.

Baten, J.; Szoltysek, M. & Campestrini, M. 2016. "Girl Power" in Eastern Europe? The human capital development of Central-Eastern and Eastern Europe in the seventeenth to nineteenth centuries and its determinants. *European Review of Economic History*, 21(1): 29-63.

Beltrán Tapia, F. J. & Martinez-Galarraga, J. 2018. Inequality and education in pre-industrial economies: Evidence from Spain. *Explorations in Economic History,* 69: 81–101.

Beltrán Tapia, F. J., Díez-Minguela, A., Martinez-Galarraga, J., & Tirado-Fabregat, D. 2018. Two stories, one fate: Age-heaping and literacy in Spain, 1877-1930. *EHES Working Papers in Economic History*, 139.

Bermúdez Méndez, M. & Martín Chicano, P. 2007. *Coín 1752 según el Catastro de Ensenada. Transcripciones y análisis crítico*. Diputación Provincial de Málaga.

Bernal, A.M. 1987. Latifundios, jornaleros y paro agrícola, *Revista de Estudios Andaluces*, 8: 67-86.

Borrás Llop, J. M. 2002a. Aprender trabajando. La actividad de niñas y niños en tierras de regadío (la Vega del Tajuña a comienzos del siglo XX). In C. Sarasúa & L. Gálvez (Eds.), *¿Privilegios o eficiencia? Mujeres y hombres en los mercados de trabajo*: 157–183. Alicante: Universidad de Alicante.

Borrás Llop, J. M. 2002b. Mercado laboral, escolarizacion y empleo infantil en una comarca agricola e industrial (el Valles Occidental, 1881-1910). *Cuadernos de Historia Contemporánea*, 24: 147-149.

Camarero Bullón, C. 2002. El Catastro de Ensenada, 1749-1759: diez años de intenso trabajo y 80.000 volúmenes manuscritos, *CT: Catastro*, 46: 61-88.

Carmona, J. & Simpson, J. 2003. *El Laberinto de la agricultura española. Instituciones, contratos y organización entre 1850 y 1936*. Prensas Universitarias de Zaragoza.

Carmona, J. & Simpson, J. 2007. Economías de escala, organización de patrimonios y obstáculos a una reforma agraria. Andalucía, 1880-1936. In R. Robledo and S.M. López (Coord.), *¿Interés particular, bienestar público?: grandes patrimonios y reformas agrarias*, Prensas Universitarias de Zaragoza.

Carmona, J., Rosés, J., & Simpson, J. 2019, The question of land access and the Spanish land reform of 1932, *Economic History Review,* 72 (2): 669–690

Clark, G. & Gray, R. 2014. Geography is not destiny: geography, institutions and literacy in England, 1837–63. *Oxford Economic Papers*, 66 (4): 1042–1069

Clayburn la Force, J. 1964. Royal textile factories in Spain, 1700-1800. *The Journal of Economic History*, 24 (3): 337-363.

Crayen, D. & Baten, J., 2010. Global trends in numeracy 1820-1949 and its implications for long-term growth, *Explorations in Economic History*, 47: 82-99.

De la Pascua Sánchez, M.J. 1989. Aproximación a los niveles de alfabetización en la provincia de Cádiz: las poblaciones de Cádiz, El Puerto de Santa María, Medina

Sidonia y Alcalá de los Gazules entre 1675-1800. *Trocadero, Revista de Historia Moderna y Contemporánea*, 1: 51-65.

Delgado Criado, B. 1993. *Historia de la educación en España y América. La educación en la España moderna (siglos XVI-XVIII)*. Madrid: Fundación Santa María.

Galor, O.; Moav, O., & Vollrath, D. 2009. Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. *The Review of economic studies*, 76 (1): 143–179.

García Sanz, A. 1996. Verlagssystem y concentración productiva en la industria pañera de Segovia durante el siglo XVIII, *Revista de historia industrial*, 10: 11-36.

Gómez-i-Aznar, È. 2019. Human capital at the beginnings of the 18th century Catalonia: age-heaping and numeracy in a changing economy, *Documentos de Trabajo (DT-AEHE)*, No 1904.

González de Molina, M. 2014. La tierra y la cuestión agraria entre 1812 y 1931: latifundismo versus campesinización. In M. González de Molina (Ed.), *La cuestión agraria en la historia de Andalucía. Nuevas perspectivas.* Fundación Pública Andaluza Centro de Estudios Andaluces.

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

Juif, D. & 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 (2): 227-241.

Kagan, Richard L. 1974. *Students and society in early modern Spain*. Baltimore: Hopkins

Kagan, Richard L. 1981. *Universidad y sociedad en la España moderna*. Madrid: Tecnos.

Marcos Martín, A. 2000. *España en los siglos XVI, XVII y XVIII. Economía y sociedad*. Barcelona: Crítica.

Mata Olmo, R. 1984. Transformación en regadío y evolución de la gran explotación agraria: el ejemplo de la Depresión del Guadalquivir, *Agricultura y sociedad*, 32: 193-228.

Mata Olmo, R. & Naranjo-Ramírez, J. 1997. La Geografía rural y el estudio de la tenencia de la tierra en España. *Asociación de Geógrafos Españoles*.

Núñez, C. E. 1992. *La fuente de la riqueza: educación y desarrollo económico en la España contemporánea*. Alianza.

Oto-Peralías, D. & Romero-Ávila, D. 2016. The economic consequences of the Spanish Reconquest: the long-term effects of Medieval conquest and colonization. *Journal of Economic Growth,* 21 (4): 409–464.

Parejo, A. (2009). *Historia económica de Andalucía contemporánea*. Síntesis.

Peña Díaz, M. (Ed.). 2012. *Breve historia de Andalucía*. Fundación Pública Andaluza Centro de Estudios Andaluces.

Pérez Picazo, M.T. & Lemeunier, G. 1985. *Agua y coyuntura económica. Las transformaciones de los regadíos murcianos (1450-1926)*, Geo Crítica, 58.

Ponsot, P. 1986. *Atlas de Historia Económica de la Baja Andalucía*. Granada: Editoriales Andaluzas Unidas.

Reher, D.S. 1994. Ciudades, procesos de urbanización y sistemas urbanos en la península ibérica, 1550-1991, in M. Guàrdia, F.J. Monclús & J.L. Oyón (Eds.), *Atlas Histórico de las ciudades europeas. Península Ibérica*: 1-30. Salvat-Centre de Cultura Contemporàni de Barcelona: Barcelona.

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

Rodríguez, M. C. & Bennassar, B. 1978. Signatures et niveau culturel des témoins et accusés dans les procès d'inquisition du ressort du Tribunal de Tolède (1525-1817) et du ressort du Tribunal de Cordoue (1595-1632). *Cahiers du monde hispanique et luso-brésilien,* (31): 17–46.

Santiago-Caballero, C. 2011. Income inequality in central Spain, 1690–1800. *Explorations in Economic History,* 48 (1): 83–96.

Sarasúa, C. 2002. El acceso de niños y niñas a los recursos educativos en la España rural del siglo XIX. In J. Martínez Carrión (ed.), *El nivel de vida en la España rural, siglos XVIII-XX*: 549-609. Universidad de Alicante.

Scheidel, W. (2017). *The Great Leveler Violence and the History of Inequality from the Stone Age to the Twenty-First Century*. Princenton University Press.

Schofield, R. & Reher, D. 1994. El descenso de la mortalidad en Europa. *Boletín de la Asociación de demografía histórica*, XII (1): 9-32.

Simpson, J. & Carmona, J. 2017. Too many workers or not enough land? The experience of land reform in Spain during the 1930s, *Historia Agraria*, 72: 37-68.

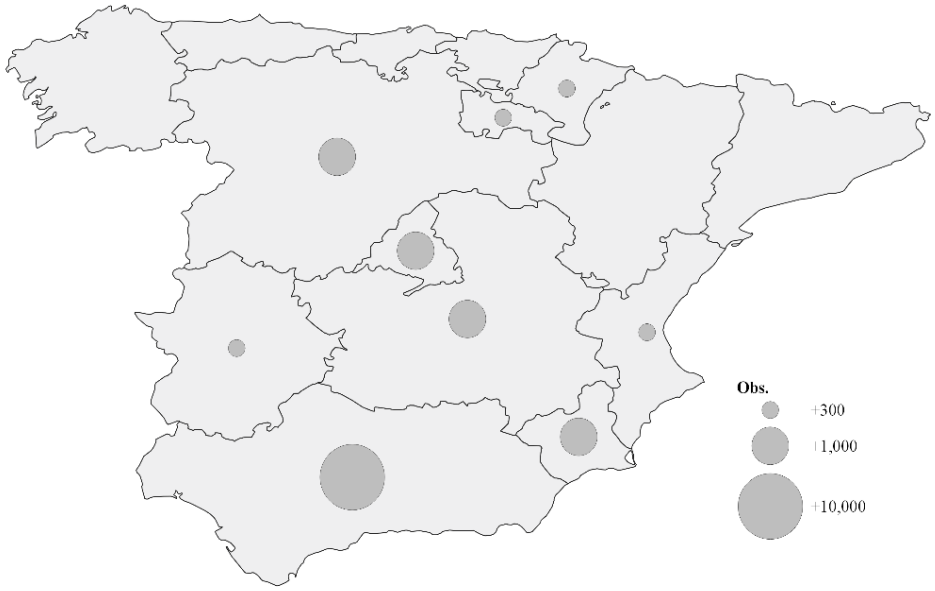
Tollnek, F. & Baten, J. 2017. Farmers at the heart of the 'human capital revolution'? Decomposing the numeracy increase in early modern Europe. *The Economic History Review,* 70 (3): 779–809.

Viñao Frago, A. 1999. Alfabetización y primeras letras (siglos XVI-XVII). In A. Castillo (Ed.): *Escribir y leer en el siglo de Cervantes*: 39-84. Barcelona: Gedisa Editorial.

Vincent, B. 1987. Lisants et non-lisants des royaumes de Grenade et de Valence a la fin du XVI siècle. In CNRS (Ed.), *De l'alphabetisation aux circuits du livre en Espagne, XVI-XIX siècle*, París, CNRS: 85-104.

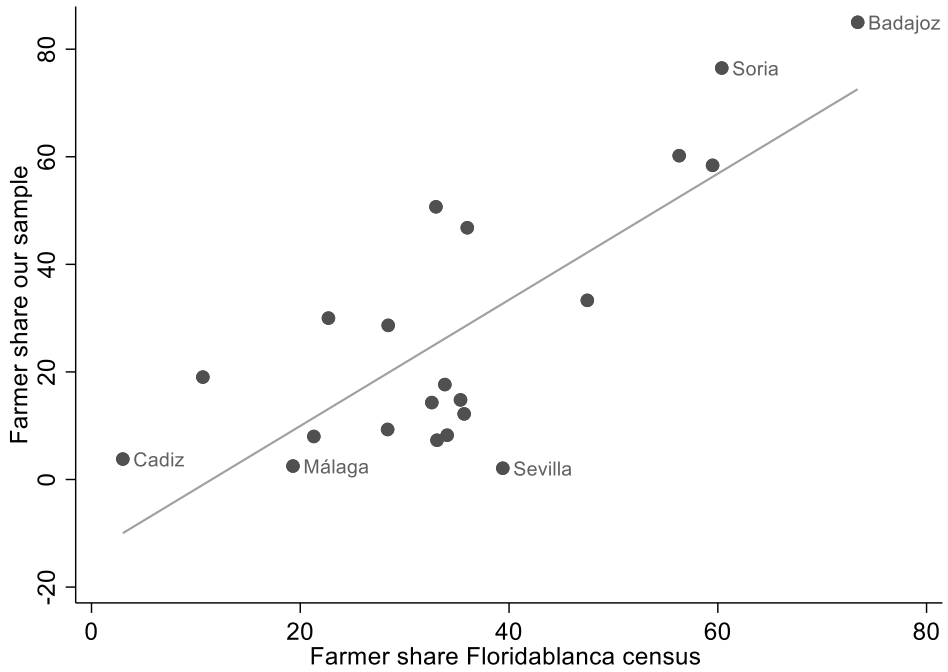
2.7 Figures and Tables

Figure 2.1 Location and sample (birth decade 1580-1760)



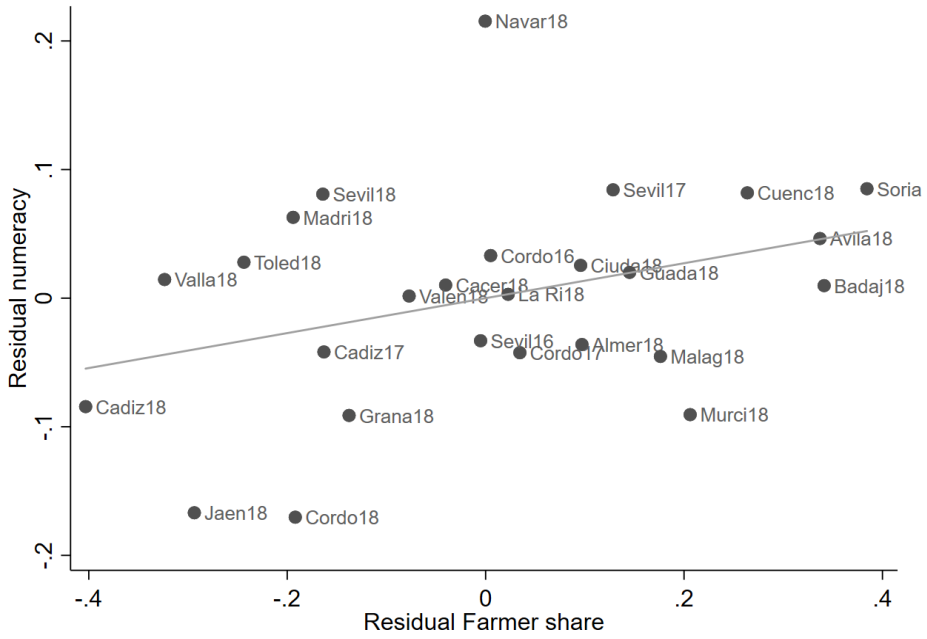
Source: see section 2.3 of this text

Figure 2.2 Comparison of the farmers' share in the Floridablanca census and in our sample



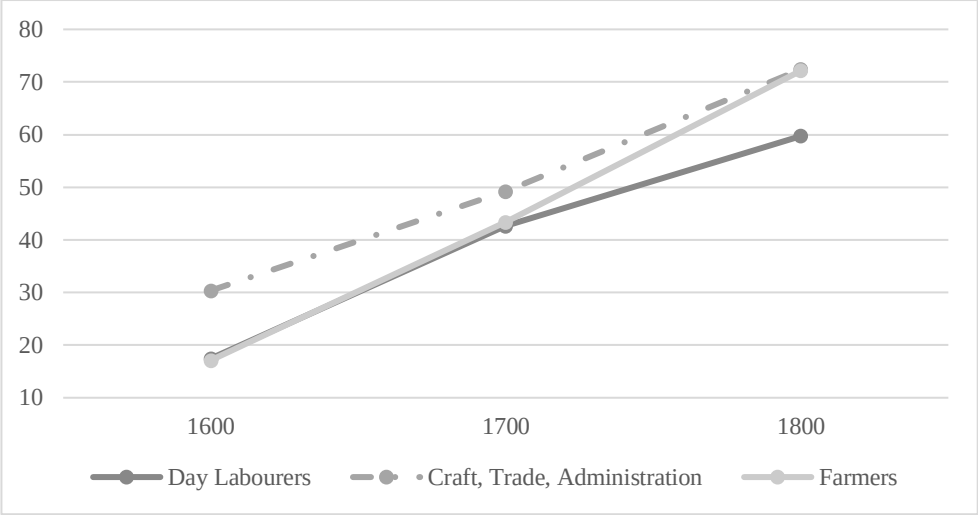
Note: we aggregate the farmers' share here for only the local communities for which we have numeracy data. For example, Murcia is only represented by Lorca, Valencia only by Sueca. Consequently, this comparison does not aim at representativeness for the provinces.

Figure 2.3 Relation of residual farmers' share and residual numeracy, on a provincial aggregate leve



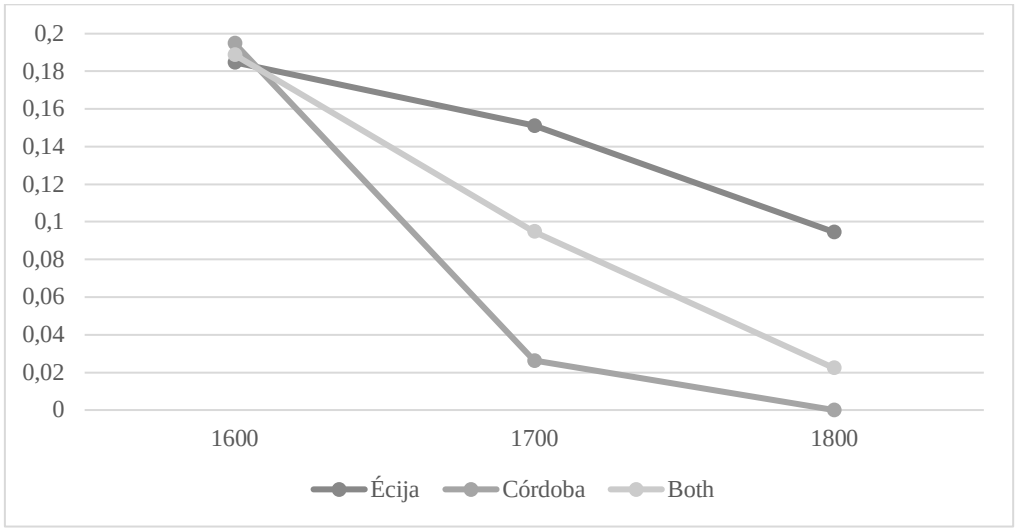
Note: in the regression analysis, we used 117 local community-birth century units. Here we aggregated by province and birth century, in order to make the figure more easily understandable.

Figure 2.4 Numeracy of farmers, agricultural laborers and other occupations.



Note: 1600 refers to Andalusia only (Cordoba and Écija), 1700 and 1800 to all of Spain. "1600" is the sixteenth century etc.

Figure 2.5 Share of farmers in Cordoba and Écija (the two local communities with continuously reported occupations), relative to other day labourers



Note: "1600" is the sixteenth century etc.

Table 2.1 Nº Observations by province and birth century

Province	Nº Total Observations 16th	17th	18th	Nº Observations with occupations province	16th	17th	18th
Almeria		224	1196	Almeria		130	733
Avila		22	130	Avila		22	130
Badajoz		22	98	Badajoz		22	98
Caceres		41	324	Caceres		41	324
Cadiz		549	196	Cadiz		379	180
Ciudad Real		9	109	Ciudad Real		9	109
Cordoba	253	1283	1300	Cordoba	202	630	905
Cuenca		35	208	Cuenca		35	182
Granada		718	4613	Granada		373	2167
Guadalajara		191	1442	Guadalajara		85	735
Jaen		36	909	Jaen		35	857
La Rioja		69	285	La Rioja		61	258
Madrid		44	219	Madrid		44	219
Málaga		110	1206	Málaga		50	308
Murcia		191	939	Murcia		191	939
Navarra			337	Navarra			140
Seville	303	549	337	Seville	222	424	303
Soria		306	1787	Soria		292	1747
Toledo		740	5162	Toledo		445	2780
Valencia			324	Valencia			304
Valladolid		7	28	Valladolid		7	28
	556	5,146	21,704		424	3,275	13,674
Total	26,851			Total	17,145		

Source: see section 2.3 of this text

Table 2.2 Descriptive statistics

Variable	Obs.	Mean.	Std. Dev.
Numerate	26,851	0.57	0.50
Farmers' share	26,851	0.33	0.27
Farmer	26,851	0.14	0.35
Day Labourer	26,851	0.17	0.38
Age 23-32	26,851	0.33	0.47
Age 43-52	26,851	0.22	0.42
Age 53-62	26,851	0.15	0.35
City*	26,851	0.21	0.41
Female	26,851	0.34	0.47

*More than 20,000 inhabitants.
Note: at the individual level, all these variables are coded as 0 or 1.
Source: see section 2.3 of this text

Table 2.3 The effect of land equality indicator "farmers' share" on individual numeracy (the likelihood of individuals not to report a rounded age) using a linear probability model (LPM)

	(1)	(2)	(3)
Farmers' share	12.14**	9.65**	9.59**
	(0.034)	(0.024)	(0.032)
Farmer	0.38	0.36	-0.02
	(0.849)	(0.888)	(0.994)
Age 23-32	2.75**	0.13	-1.25
	(0.043)	(0.960)	(0.686)
Age 43-52	-4.38*	-5.33*	-3.92
	(0.067)	(0.071)	(0.232)
Age 53-62	-2.15	-10.04	-3.93
	(0.727)	(0.161)	(0.555)
City	1.19	-0.44	2.16
	(0.847)	(0.948)	(0.784)
Female	2.17	0.21	
	(0.235)	(0.908)	
Constant	24.61***	31.49***	24.92**
	(0.004)	(0.001)	(0.011)
Observations (individuals)	26,851	26,851	17,777
Adjusted R-squared	0.04	0.04	0.04
Time FE	YES	YES	YES
Region FE	YES	YES	YES

The dependent variable is 1 if the individual reported an unrounded age, 0 otherwise. The constant refers to male non-farmers living in local communities of fewer than 20,000 inhabitants aged 33-42. Time fixed effects are half centuries and region fixed effects are historical regions. We

clustered by local community of birth and birth decade. We use the weights with the analytic weight function for the population of census (columns 2 and 3). We weighted by the population share of Aranda census by historical regions. This implies that local communities are stronger weighted, for which we have less observations relative to the total observations in the censuses. We use stata´s analytic weights, including "[aw=pop]". Our local communities are classified as follows according to the classification of the Aranda census by historical regions: Andalusia: Almería, Almuñécar, Bérchules, Bubión/Capileira, Colomera, Cordoba, Écija, Estepona, Granada, Iznalloz, Jaen, Laujar de Andarax, Loja, Málaga, Montilla, Navas de San Juan, Puerto de Santa María, Villanueva del Rey; Castilla La Nueva: Abenójar, Alovera, Arganda, Cavanillas, El Casar, Marchamalo, Móstoles, Pinto, Saelices, Toledo, Villanueva de la Torre, Yunquera de Henares; Castilla La Vieja: Adanero, Adradas, Aguaviva de la Vega, Aguilar y Montuenga, Alcubilla del Marqués, Aldea de San Esteban, Aldeasenor, Alentisque, Almaluez, Almarza, Almazán, Andaluz, Arcos de Jalón, Arévalo, Atauta, Fuente El Sol, Inestrillas, Logroño, Ontalvilla de Almazán, Torreandaluz, Ziria; Extremadura: Alía, Valdecaballeros; Murcia: Abanilla, Abrán, Albudeite, Lorca; Navarra y País Vasco: Olite; País Valenciano: Sueca. Robust p-Values are given in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.4 The effect of the land equality indicator "farmers' share" on individual numeracy (the likelihood of individuals not to report a rounded age) using a Logit model (Marginal effects reported)

	(1)	(2)	(3)
Farmers' share	12.57**	9.88**	9.84**
	(0.031)	(0.021)	(0.029)
Farmer	0.45	0.44	0.06
	(0.829)	(0.867)	(0.983)
Age 23-32	2.88**	0.21	-1.19
	(0.035)	(0.936)	(0.711)
Age 43-52	-4.53*	-5.43*	-4.00
	(0.063)	(0.077)	(0.243)
Age 53-62	-2.28	-10.40	-4.01
	(0.720)	(0.157)	(0.554)
City	1.28	-0.38	2.18
	(0.833)	(0.955)	(0.772)
Female	2.23	0.25	
	(0.230)	(0.894)	
Observations (individuals)	26,851	26,851	17,777
Time FE	YES	YES	YES
Region FE	YES	YES	YES
Pseudo R2	0.0296	0.0296	0.0296

The dependent variable is 1 if the individual reported an unrounded age, 0 otherwise. The constant refers to male non-farmers living in local communities of fewer than 20,000 inhabitants aged 33-42. Time fixed effects are half centuries and region fixed effects are historical regions. We clustered by local community of birth and birth decade. Weights establish representativeness for the regions included in columns 2 and 3 (see note on Table 3). Robust p-Values are given in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 2.5 How large was the numeracy difference between farmers and agricultural labour (and non-agricultural occupations)?

	(1)	(2)
Farmer	7.11***	4.76*
	(0.004)	(0.087)
All non-agric. Occupations	10.18***	9.73***
	(0.000)	(0.000)
Age 23-32	-1.44	-1.38
	(0.489)	(0.509)
Age 43-52	-5.54***	-6.01***
	(0.004)	(0.002)
Age 53-62	-6.86	-10.71**
	(0.124)	(0.034)
Constant	23.70***	46.88***
	(0.000)	(0.000)
Observations	15,901	15,901
Adjusted R-squared	0.04	0.05
Time FE	YES	YES
Region FE	YES	NO
Local community FE	NO	YES

Note: The dependent variable is 1 if the individual reported an unrounded age, 0 otherwise. The constant refers to agricultural laborers aged 33-42. Time fixed effects are half centuries, region fixed effects are historical regions and LC fixed effects are for each local community. Weights establish representativeness for the regions included (see note on Table 3). Robust p-Values are given in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

2.8 Appendices

2.8.1 Description of the sources

Table 2.6 Description of the sources

Local community	Year of Source	Source
Abanilla	1756	Cadaster of Ensenada, Family Search
Abarán	1756	Cadaster of Ensenada, Family Search
Abenójar	1752	Cadaster of Ensenada, Family Search
Adanero	1752	Cadaster of Ensenada, Family Search
Adradas	1752	Cadaster of Ensenada, Provincial Historical Archive Soria
Aguaviva de la Vega	1752	Cadaster of Ensenada, Provincial Historical Archive Soria
Aguilar de Montuenga	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Albudeite	1756	Cadaster of Ensenada, Family Search
Alcubilla del Marques	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Aldea de San Esteban	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Aldeasenor	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Alentisque	1752	Cadaster of Ensenada, Provincial Historical Archive Soria
Alía	1752	Cadaster of Ensenada, Family Search
Almaluez	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Almarza	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Almazán	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Almería	1753	Cadaster of Ensenada, Family Search
Almuñecar	1752	Cadaster of Ensenada, Provincial Historical Archive Granada
Alovera	1751	Cadaster of Ensenada, National Historical Archive Madrid
Andaluz	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Arcos de Jalon	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Arevalo	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Arganda	1753	Cadaster of Ensenada, Regional Archive Madrid
Atauta	1753	Cadaster of Ensenada, Provincial Historical Archive Soria

Berchules	1752	Cadaster of Ensenada, Provincial Historical Archive Granada
Bubion/Capileira	1750	Cadaster of Ensenada, Provincial Historical Archive Granada
Cavanillas	1751	Cadaster of Ensenada, National Historical Archive Madrid
Colomera	1752	Cadaster of Ensenada, Provincial Historical Archive Granada
Cordoba	1643	Padrón, Family Search
Cordoba	1693	Padrón, Family Search
Cordoba	1718	Padrón, Family Search
Cordoba	1761	Padrón, Family Search
Ecija	1645	Padrón, Family Search
Ecija	1704	Padrón, Family Search
Ecija	1775	Padrón, Family Search
El Casar	1751	Cadaster of Ensenada, National Historical Archive Madrid
Estepona	1752	Cadaster of Ensenada, Provincial Historical Archive Granada
Fuente El Sol	1752	Cadaster of Ensenada, Family Search
Granada	1752	Cadaster of Ensenada, Provincial Historical Archive Granada
Inestrillas	1752	Cadaster of Ensenada, Family Search
Iznalloz	1753	Cadaster of Ensenada, Provincial Historical Archive Granada
Jaen	1771	Padrón, Family Search
Laujar de Andarax	1751	Cadaster of Ensenada, Provincial Historical Archive Almería
Logroño	1751	Cadaster of Ensenada, Family Search
Loja	1750	Padrón, Family Search
Loja	1753	Cadaster of Ensenada, Provincial Historical Archive Granada
Lorca	1756	Cadaster of Ensenada, Family Search
Málaga	1751	Cadaster of Ensenada, Municipal Archive Málaga
Málaga	1776	Padrón, Municipal Archive Málaga
Marchamalo	1752	Cadaster of Ensenada, National Historical Archive Madrid
Montilla	1752	Cadaster of Ensenada, Family Search
Móstoles	1753	Cadaster of Ensenada, Regional Archive Madrid

Navas de San Juan	1752	Cadaster of Ensenada, Family Search
Olite	1786	Floridablanca, Municipal Archive Olite
Ontalvilla de Almazán	1753	Cadaster of Ensenada, Provincial Historical Archive Soria
Pinto	1753	Cadaster of Ensenada, Family Search
Puerto de Santa María	1719	Padrón, Family Search
Puerto de Santa María	1734	Padrón, Family Search
Puerto de Santa María	1762	Padrón, Family Search
Saelices	1752	Cadaster of Ensenada, Family Search
Sueca	1794	Padrón, Family Search
Toledo	1752	Cadaster of Ensenada, Regional Archive Madrid
Torreandaluz	1752	Cadaster of Ensenada, Provincial Historical Archive Soria
V. de la Torre	1751	Cadaster of Ensenada, National Historical Archive Madrid
Valdecaballeros	1753	Cadaster of Ensenada, Family Search
Villanueva del Rey	1750	Cadaster of Ensenada, Family Search
Yunquera de Henares	1751	Cadaster of Ensenada, National Historical Archive Madrid
Ziria	1753	Cadaster of Ensenada, Provincial Historical Archive Soria

Table 2.7 Share of individuals in our sample in the 18th century

CCAA	province	pl	N° sample (23-62)	N° inhab. Floridablanca (age 25-50)	% sample
Andalucía	Sevilla	Écija	337	17,599	1,9
Andalucía	Cádiz	Puerto de Santa María	196	8845	2,2
Andalucía	Córdoba	Córdoba	576	19665	2,9
Andalucía	Málaga	Málaga	815	26423	3,1
Andalucía	Almería	Almería	346	7404	4,7
Andalucía	Granada	Granada	1410	28696	4,9
Murcia	Murcia	Lorca	523	9238	5,7
La Rioja	La Rioja	Logroño	182	3172	5,7
Andalucía	Córdoba	Montilla	539	6641	8,1
Andalucía	Jaén	Jaén	753	8322	9,0
Murcia	Murcia	Abarán	79	751	10,5
Andalucía	Granada	Loja	753	5648	13,3
Murcia	Murcia	Albudeite	147	1058	13,9
Murcia	Murcia	Abanilla	190	1305	14,6
Comunidad Valenciana	Valencia	Sueca	324	2223	14,6
Andalucía	Málaga	Estepona	391	2257	17,3
Castilla León	Valladolid	Fuente El Sol	28	130	21,5
Castilla La Mancha	Cuenca	Saelices	208	774	26,9
Castilla La Mancha	Ciudad Real	Abenójar	109	353	30,9
Madrid	Madrid	Pinto	219	704	31,1
Extremadura	Badajoz	Valdecaballeros	98	314	31,2
Castilla La Mancha	Toledo	Arganda	352	1116	31,5
Castilla León	Ávila	Adanero	130	391	33,2
Andalucía	Jaén	Navas de San Juan	156	449	34,7
La Rioja	La Rioja	Inestrillas	103	296	34,8
Andalucía	Granada	Almuñécar	497	1395	35,6
Extremadura	Cáceres	Alía	324	824	39,3
Navarra	Navarra	Olite	337	708	47,6
Andalucía	Córdoba	Villanueva del Rey	185	372	49,7

Andalucía	Granada	Bubion/Capileira	528	1017	51,9
Castilla León	Soria	Aldeaseñor	52	100	52,0
Castilla León	Soria	Almazán	567	1055	53,7
Castilla La Mancha	Toledo	Toledo	4454	8216	54,2
Castilla León	Soria	Aguaviva de la Vega	93	169	55,0
Andalucía	Granada	Iznalloz	433	767	56,5
Castilla León	Soria	Alentisque	66	113	58,4
Castilla León	Soria	Ontalvilla de Almazán	50	85	58,8
Castilla León	Soria	Torreandaluz	33	56	58,9
Castilla León	Soria	Ziria	156	262	59,5
Castilla León	Soria	Aldea de San Esteban	31	52	59,6
Castilla León	Soria	Almarza	123	205	60,0
Castilla León	Soria	Adradas	56	93	60,2
Castilla León	Soria	Arcos de Jalón	125	207	60,4
Castilla León	Soria	Arévalo	60	99	60,6
Andalucía	Granada	Bérchules	487	801	60,8
Castilla León	Soria	Alcubilla del Marques	51	83	61,4
Andalucía	Granada	Colomera	505	811	62,3
Castilla León	Soria	Atauta	52	82	63,4
Madrid	Madrid	Móstoles	356	548	65,0
Castilla La Mancha	Guadalajara	Marchámalo	329	505	65,1
Castilla León	Soria	Andaluz	35	53	66,0
Castilla La Mancha	Guadalajara	V. de la Torre	107	146	73,3
Castilla La Mancha	Guadalajara	El Casar	384	518	74,1
Andalucía	Almería	Láujar Andarax	850	1124	75,6
Castilla León	Soria	Almaluez	117	140	83,6
Castilla La Mancha	Guadalajara	Alovera	163	183	89,1
Castilla La Mancha	Guadalajara	Yunquera	277	303	91,4
Castilla La Mancha	Guadalajara	Cabanillas	182	195	93,3
Castilla León	Soria	Aguilar y Montuenga	120	no data in Floridablanca census	

Source: see section 2.3 of this text

2.8.2 Potential Caveats

First, 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 to numeracy, 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 would be able to buy more land. These peasants might also influence political activity in favor of land reforms, as Cinnirella and Hornung (2016) 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 nearby cities, for example.

Oto-Peralías and Romero-Ávila (2016) and Beltrán Tapia and Martinez-Galarraga (2018) recently advocated the Reconquista events as an instrument of land inequality (a similar instrument was used by Baten and Hippe 2018). The advantage of the speed of Reconquista is intrinsically exogenous in nature, as it depended more on military status during the medieval period than on any economic characteristic of the territories that were reconquered. Hence, Reconquista speed is most likely very exogenous. Moreover, Oto-Peralías and Romero-Ávila (2016) show that the inequality of land in Spain had its origins in the *Reconquista* during the Middle Ages. It was the rapid phase of the *Reconquista* during the thirteenth century, which caused the large land inequality, that is, three to five centuries before our period. Therefore, it is not likely that numeracy determined the farmers' share. This was mostly reinforced during the following centuries. The share of lords and military orders slightly increased their landholdings during the fifteenth,

sixteenth and seventeenth centuries (De Albornoz de la Escosura 1963)[30]. As a result, during the repopulation process, a small number of aristocratic families and ecclesiastical entities emerged as owners of large properties, especially in the southwest of Spain (Tortella 2000). Consequently, Oto-Peralías and Romero-Ávila (2016) have argued that the rate of *Reconquista* determined the distribution of regional income. A slow expansion contributed to set better political institutions and equitable distribution of land such as in the north of the Duero valley, for example. Beltrán Tapia and Martinez-Galarraga (2018) used the *Reconquista* as an instrument in the nineteenth century; their results show that the timing of *Reconquista* was positively correlated to the landownership structure.

Another potential issue could be migration. For example, we could imagine that more numerate people moved to regions where land inequality was less prevalent. Migratory intra-rural movements related to agricultural labour in the south were studied by Florencio Puntas and López Martínez (2000). They found that since the middle of the fifteenth century, there has been evidence of seasonal migrations related to agricultural work in the region of Seville. Seasonal emigration in Andalusia was widespread and typical of the whole period, whereas there was not as much permanent migration within the regions of Andalusia. The same results were shown by Bernal (1987) who studied the mobility of day laborers in the Guadalquivir Valley during the sixteenth and eighteenth centuries. Although this mobility was of medium or long distance (Eastern Andalusians in the western zone or Spaniards from the north who went down to the south to harvest) all were not permanent. Furthermore, Sánchez Picón (1988) has studied migratory movements for the province of Almeria in Eastern Andalusia. The migrations were

[30] Through the mayorazgos (family holdings that were inherited by the firstborn) the nobility contributed to this increase not allowing the dispersion of lands.

mainly seasonal, carried out by harvesters who, since the eighteenth century, had gone to the Andalusian countryside as a subsistence strategy. Additionally, for the north of Spain, there is evidence of temporary migrations during the eighteenth and nineteenth centuries (Sarasúa 1994). Ultimately, during the period studied, day laborers were unable or showed no interest in emigrating permanently (Carmona and Simpson 2003).

In general, poor, but numerate individuals did not typically earn enough to buy or develop sufficient skills to rent farms in this early period (Baten and Hippe 2018). It would not matter whether one farmer moved to another region; biases from migration only occur if labourers from *latifundia* regions could buy or rent farms in other districts and hence migrate to these regions permanently. However, this is a very unlikely scenario for early modern societies[31].

[31] Although some day labourers rented land from landowners, this practice was not the usual means to earn capital (Carmona and Simpson, 2003: 115).

2.8.3 References

Baten, J. & Hippe, R. 2018. Geography, land inequality and regional numeracy in Europe in historical perspective. Journal of Economic Growth, 23 (1): 79–109.

Beltrán Tapia, F. J. & Martinez-Galarraga, J. 2018. Inequality and education in pre-industrial economies: Evidence from Spain. Explorations in Economic History, 69: 81–101.

Bernal, A.M. 1987. Latifundios, jornaleros y paro agrícola, Revista de Estudios Andaluces, 8: 67-86.

Carmona, J. & Simpson, J. 2003. El Laberinto de la agricultura española. Instituciones, contratos y organización entre 1850 y 1936. Prensas Universitarias de Zaragoza.

Cinnirella, F. & Hornung, E. 2016. Landownership concentration and the expansion of education. Journal of Development Economics, 121: 135-152.

De Albornoz de la Escosura, A. 1963. El "milagro" de la agricultura española. La distribución de la propiedad. Comercio Exterior: 764–767

Florencio Puntas, A. & López Martínez, A. 2000. Las migraciones estacionales agrarias en Andalucía anteriores al siglo XX, Boletín de la Asociación de Demografía Histórica, XVIII (I): 71-100.

Oto-Peralías, D. & Romero-Ávila, D. 2016. The economic consequences of the Spanish Reconquest: the long-term effects of Medieval conquest and colonization. Journal of Economic Growth, 21 (4): 409–464.

Sánchez Picón, A. 1988. Marchar a las Andalucías: un episodio migratorio en la Almería del siglo XIX. I Encuentro de Cultura Mediterránea, Confederación Española de Cajas de Ahorro, Cajalmería

Sarasúa, C. 1994. Las emigraciones temporales en una economía de minifundio: los montes de Pas, 1758-1888, Boletín de la Asociación de Demografía Histórica, XII (2/3): 165-179

Tortella, G. 2000. The Development of modern Spain. An economic history of the nineteenth and twentieth centuries. Harvard University Press

3. Numeracy selectivity of Spanish migrants in Hispanic America (16th - 18th centuries) [32]

Abstract

This paper assesses the human capital composition of Spanish migrants who went to colonial Hispanic America during the sixteenth to eighteenth centuries. To estimate the numeracy levels of the Spaniards who left Spain to settle in the colony, I use the age-heaping based method to measure the human capital. The main finding is that the Spanish migrants were positively selected. Differences are observed in the human capital of those who chose to settle in Mexico, with a higher level of numeracy, than those who chose Peru. These differences could be due to the viceroyalty structure and the presence of religious orders that encouraged the emigration of people with greater human capital to Mexico. Finally, it seems that inequality between Spaniards and natives, in terms of human capital, was larger in Mexico at the end of the sixteenth century reducing the gap circa 1710.

[32] The author would like to thank Antonio García-Abásolo and María del Carmen Martínez Martínez for providing information about the works that contain the data used in this research. The author is also grateful for comments on this paper to the organizers and participants in the Summer School in Economic and Social History (Vila Viçosa, Portugal, 26th – 29th June 2019) and to Carmen Sarasúa, Joerg Baten and Laura Maravall.

3.1 Introduction

The selectivity of migrants is a key factor to understand the impact of international migrations. It has been argued that during the nineteenth and 20th century internal migrants were, on average, positively self-selected in terms of health and human capital in Great Britain and Spain (Humphries and Leunig 2009; Quiroga 2003; Beltrán and Salanova 2017; Juif and Quiroga 2019). Sánchez Alonso (2007) found that during the nineteenth and 20th centuries migrants from Southern Europe to Latin America were more literate than those who stayed. In contrast, using the numeracy as a proxy for human capital, Juif (2015) observed that European immigrants in Cuba in nineteenth century were negatively self-selected. The crisis in the agricultural sector of the Canary Islands in the nineteenth century encouraged the poor and less educated population to move to Cuba to work in its plantation system. Despite this interest there are no studies analysing the characteristics and self-selection of migrants since the sixteenth century from European regions to Latin America.

On the other hand, the attempt of explaining the high inequality of Latin America has become one of the central topics in economic history (Coatsworth and Summerhill 2010). One branch of the literature focuses on colonial institutions and their negative impact on contemporary economic development (Acemoglu and Robinson 2013; Sokoloff and Engerman 2000). However, Arroyo Abad and van Zanden (2016) have shown a substantial increase in real wages between the 1550s and 1780s in colonial Mexico and Peru. Some researchers have also argued that high inequality was a phenomenon of the nineteenth century as a result of export-oriented economic growth (Coatsworth 2008; Williamson 2009). Focusing on living standards, especially heights,

Dobado González and García Montero (2010) have shown that, from an international comparative perspective, Bourbon America was not a particularly unequal society.

One way to contribute to both debates (migrants self-selection and level of inequality) is to assess human capital, which is a widely recognised determinant of long-term economic growth (Becker 2008; Barro 2001; Schultz 1961). *The Oxford Handbook of Latin America History* edited by Jose C. Moya (2010) has two chapters that provide a literature review of the Historiography of *Nueva España* and Colonial Spanish South America. In none of them the topic of human capital is mentioned. In the same book, Coatsworth and Summerhill (2010: 419) claim: "*the most glaring omission involves the lack of systematic studies of education […], skill acquisition, and human capital formation…*". Instead, we know literacy (or illiteracy: per cent of the population which can neither read or write) rates for the countries of Central and Latin America since the nineteenth century after independence. In 1870 Argentina had between 75-80% population above 15 years old illiterate, Chile 70-80% and Cuba 70-75%. In 1890 the percentages were 55-60% in Argentina, 60-65% in Chile and Cuba and 80-85% in Mexico (Sánchez Alonso 2007: 416). In 1877 in Spain, the illiterate rates were 38% in Madrid, the less illiterate region, and 80% in eastern Andalusia, the most illiterate region in the peninsula. In 1887, these rates were 34% and 77% respectively (Núñez 1992: 132). The national averages were 67% in 1877 and 62% in 1887 (Núñez 1992: 94).

In the last years the number of studies that assess the human capital in colonial Hispanic America has started to flourish. Using the age-heaping technique it is possible to estimate the numeracy level, as a proxy for human capital, for those societies where the lack of data hampers to know the literacy rates. Although various authors have estimated numeracy levels in Latin America, only a few of them have studied the time

period before the independence. Manzel et al. (2012) performed the first attempt to estimate the development of human capital for seven Latin American countries from the seventeenth to the twentieth century. They concluded that Argentina, Mexico and Peru experienced a rise until the late eighteenth century reducing the gap from 50 to 30 per cent separating them from Western Europe by the 1780s. However, between the late eighteenth century and the early nineteenth century, coinciding with the wars of independence, the numeracy levels stagnated for both countries. At the same time those of Western Europe soared, triggering the divergence between both regions. During the onset of the globalization this inequality was even higher (Baten and Mumme 2010). Calderón-Fernández et al. (2020) found that the population of central Mexico during the eighteenth century had levels of numeracy similar to those of Italy and Portugal. Furthermore, around 1820 Ireland and most of Eastern Europe had lower numeracy skills than the Mexican population[33].

However, within all these studies that quantify human capital and aim to identify its determinants, very few analyse the impact of migration. Juif and Baten (2013) found that Spanish settlers had twice the numeracy level than those of Peruvian Inca *Indios* in the Andean region. Focusing on Brasil, Stolz et al. (2013) established that in those states where most emigrants arrived, the human capital grew strongest even after controlling for educational expenditures. Sánchez Alonso (2007: 418) concluded that Southern European migrants in Latin America during the nineteenth century "carried higher literacy rates than native populations. On the whole, Latin America benefited clearly from European immigration". The same is claimed by Droller (2018) who demonstrated that

[33] Numeracy levels (ABCC index) of Mexican population during eighteenth centruy: 67.9 in Oaxaca, 64.1 in Mexico City and 63.7 in other 24 localities (Calderon-Fernández et al. (2020: 14).

the most qualified Europeans contributed to increasing literacy rates in the Pampas (Argentina) and played a decisive part at the beginning of the industrialisation.

Using the age-heaping technique to estimate the numeracy skills as a proxy for human capital, this work contributes to the literature in two ways. The main result is that migrants moving to Hispanic America during the sixteenth and eighteenth century were positively self-selected. Then I am providing new evidence about the self-selection of migrants studying a period for which data is scarce. In a second step, I analyse the inequality in terms of human capital between Spanish migrants and the indigenous Mexican population during the late seventeenth century and at the beginning of the eighteenth century. The data reveals that Spaniards had a higher level of numeracy than the native Mexicans, but it tended to converge over time.

The database comprises 31,089 individual observations including migrants and the control sample (non-migrants). Most of the data involve Andalusians, the largest group of Spaniards who left the country to go to Hispanic America. To complement the analysis, a smaller sample of migrants from the rest of the peninsula is also assessed. The remainder of the paper is structured as follows: Section 3.2 briefly introduces the historical context of colonial Spanish America. Section 3.3 follows with the explanation of the methodology and the data used in this study. In section 3.4, the empirical results and descriptive analysis are provided. Section 3.5 concludes.

3.2 Historical background of Hispanic America: 15th-18th centuries

3.2.1 Spanish conquest of the American continent

Columbus arrived with fifteen hundred Spanish settlers on the island Hispaniola in 1493. La Isabella, in honour of the queen of Castile, was the first European town founded on the American continent. After a couple of years Santo Domingo, today capital of the Dominican Republic, became the capital of Hispaniola in 1496. Twenty-five years after the first voyage of Columbus in 1492, Spanish colonies were established on the four largest islands of the Caribbean: Hispaniola, Cuba, Puerto Rico, and Jamaica (Restall and Lane 2011).

Chronologically the conquest of the rest of the territory continued as follows: in 1508 from Hispaniola, Juan Ponce de León took Puerto Rico. In 1509, Juan de Esquivel settled Jamaica. The first Spanish town on the American mainland (Tierra Firme) was founded in 1511 called Santa María la Antigua de Darién. At the same time, Cuba was conquered. During 1519 to 1521 Hernán Cortés conquered the Aztec Empire (located in the current area that goes from southern Mexico to Guatemala). From 1532 to 1536 Francisco Pizarro did the same with the Inca empire placed in the Peruvian Andes, being the major conquest in South America. The foundations of Santa Fe de Bogotá, of Santiago de Chile and Guadalajara, took place in 1538, 1541 and 1542, respectively (Bakewell 1997).

The territorial organization was divided into two levels. There was a superior government in charge of the supervision of the general administrative activity. On the second level, there were different institutions and foundations for the issues of justice, war and finance. The districts of superior government, *Nueva España*, Peru, Santo

Domingo, Guatemala and New Granada, grouped several provinces in the seventeenth century. While in the first two there was a viceroy, in the others there was a president of the *audiencia*. They depended on the *Consejo de Indias* and the king of Spain[34]. However, in the provinces of Chile and Rio de la Plata, although *audiencias* established seemed subordinated to the viceroy of Peru, they exercised their government with relative independence. Since 1543 the administration of the viceroyalties of *Nueva España* and Peru was composed of two organs: the viceroy and the *Audiencia*. Although the viceroy was president of the *Audiencia*, which dealt with the administration of justice, the power of the viceroy was separated from it (Ramos Pérez and Lohmann Villena 1985).

At the same time, in the crown of Castile in the Iberian Peninsula, it became necessary to create an institution to control all matters related to the New World. To carry this objective out, in 1503 the *Casa de la Contratación de las Indias* was created in the city of Seville[35]. This organism was the administrator in the displacement of the emigrants to the New World (Martínez Martínez 1993). During the seventeenth century the departure of the fleets was in the cities of Seville, San Lúcar de Barrameda (Cádiz) and the city of Cádiz. In 1680 Cádiz became the head of the *Indias* fleets, therefore it was the only harbour from which the trip to the Hispanic America could be undertaken[36]. Epidemics and the problems of river navigation were among the reasons behind this

[34] The *Consejo de Indias* was created in 1524 grouping all the ministers appointed by the Crown for the functioning of administrative and judicial matters (Alvar Ezquerra 2003).

[35] The *Casa de Contratación* was an institution that regulated commercial traffic between the Indias and the Peninsula. It also dealt with the technical aspects of navigation (Alvar Ezquerra 2003).

[36] *Indias* was the Spanish term to refer to Hispanic America since Columbus believed incorrectly that he had achieved his objective of reaching Asia (known in Spanish as Indias) (Elliott 1984).

decision. Finally, with the Bourbons, in 1717 the *Casa de la Contratación* was moved to Cádiz (Peña Díaz, 2012).

After 1700, the Hispanic world changed. The "door to the New World" switched from Seville to Cadiz, and the monarchy of the Bourbons replaced those of the Habsburgs. The war of succession in Spain (1701-1714) led to the second era of significant migrations (Hugon 2019). One of the objectives of Phillip V was the restriction of the functions and capabilities of the Consejo de Indias. Phillip V also changed the social origin of the bureaucratic and representative positions from the nobility to the bourgeoisie and the professional military. The change of the Casa de Contratación from Seville to Cádiz was to formalise an influential bourgeoisie (Menéndez Pidal 1988). Other substantial reforms in the government of the Bourbons were the addition of a third viceroyalty, New Granada in 1739 and the one of Río de la Plata in 1776. The purpose of creating New Granada was to mitigate the inefficiency in the audiencia at Santa Fé de Bogotá and to enlarge the capacity of collecting taxes due to the rise of the gold in this region. The motivation for the creation of Río de la Plata was similar since Buenos Aires was rising economically and commercially (Bakewell 1997).

3.2.2 Education in colonial America during the Early Modern era

During the sixteenth century it became essential to transfer culture from the Old World to the New World, as evidenced in the *Leyes de Indias* (Indian laws)[37]. For example, in 1535 Charles V ordered the creation of schools to educate the children of the

[37] The *Leyes de Indias* were the laws enacted by the kings of the Hispanic monarchy to regulate all aspects of life in their territories in America.

caciques in the Catholic religion, "good habits" and the Spanish language[38]. In 1557, Philip II supported the continuation of the school for poor and *mestizo* children in Mexico City to prevent them from staying on the street. In this school, children also learned both Christian doctrine and good habits[39]. Since the beginning, the colonisers were concerned with the evangelisation and the instruction of the Americans, therefore, at the same time as the military conquest, schools and universities emerged. In all of Hispanic America thirty-three universities were founded since 1538, when n North America the first university-college was founded in Harvard in 1636. By 1769 only nine universities were created in the English colonies of North America (Stoeckel 1976: 45). In Spanish territories, the first was placed in Santo Domingo in 1538 (*Universidad de Santo Tomás de Aquino*). The school established in 1530 by Bishop Ramírez y Fuenleal was the precedent for the *Universidad de Santiago de La Paz* authorized by Charles V in 1540 and founded by the Dominicans. They aimed at protecting the rights of the Americans at the same time that they evangelized and taught them (Delgado Criado 1993; Catholic University of America 1967).

The first school in Mexico was founded by Pedro de Gante (Franciscan) in 1525. This institution was focused on the teaching of arts and crafts for American boys. Later, in the school of San José de los Naturales (1577), "the natives learned all kinds of trades" (tailors, carpenters, blacksmiths) and even learned to construct musical string instruments and to play them. During this century the chronicles wrote that this captures "the imitative

[38] Recopilación de las leyes de los reinos de Indias, Título veinte y tres, de los colegios y seminarios, Ley XI: 141.

[39] Recopilación de las leyes de los reinos de Indias, Título veinte y tres, de los colegios y seminarios, Ley XV: 142.

ability of the natives concerning any art of craft". The daughters of Caciques in Nueva España could also attend schools for girls since 1534 in Texcoco, Huejotzingo, Cholula and Mexico City (Catholic University of America: 149). Later, the San Juan de Letrán School accepted all children, without taking into account their ethnicity. When the young students had the right age, they started to work in an artisan workshop.

Regarding the universities, Lima and Mexico had the two most important universities. In 1551 both were official and royal obtaining the pontifical confirmation in 1571 and 1595, respectively. In addition, university academies were also frequently understood as small universities or faculties. Although theoretically these institutions were open to Spaniards and Americans, the incorporation of the natives was slow and scarce (Delgado Criado 1993).

Jesuits also played an important role as educators but mostly in rural places. Under the influence of the Crown, the Church was the main promoter of educational institutions. There were elementary schools where catechism, reading, writing, number and music were taught. From 1503 the missionaries were required to establish a house next to the church where the chaplain taught and evangelized the Americans in each new town. In some cases, children of conquerors and natives were instructed in the same centres. An example of this is the school of del Cercado, founded in Lima by Jesuits. In 1582, the seminary of indios of Tepotzotlan admitted all kinds of children, those who belonged to the nobility and those who did not. However, other centres were reserved for the sons of the caciques (Saavedra Inaraja 2008).

3.3 Sources and method

The database of this research consists mainly of two samples: Spanish migrants to Hispanic America and Spaniards that were born and stayed in Spain (non-migrants sample). The first sample is composed of 7,220 and the second of 26,685 individual observations of population aged 23-62.

The data of migrants were collected from the passenger books (*Libros de Asiento de Pasajeros*) created from the licenses granted to passengers to the *Indias*. The licenses were indispensable requirements to travel to the new world and were usually granted by the king. After arriving in Seville passengers had to verify their documents in the presence of the president and the judges of the *Casa de la Contratación*. These documents recorded the personal data of the passenger and the place of destination for which it was granted. However, fraud was common among those who did not obtain licences by legal means. Buying licenses, bribes, and even dressing up as sailors were the most common practices. This kind of emigration is impossible to account for (Martínez Martínez 1993).

The creation of the database utilised in this research has been possible using several published sources. The dataset of migrants to Spanish America contains individual data of 7,220 as stated above. The birth decades considered are between 1540 and 1750. The sources used are summarised in Table 3.1. In the original works cited in Table 3.1 the number of passengers is larger than 7,220 migrants. However, the age of the passengers was not always recorded. For this reason, the number of observations is significantly lower in my sample.

How many emigrants were there? First, we need to consider that as the opposite of the era of mass migration, for this time it is not possible to have a reliable long-run series of migration (Taylor and Williamson 1997). Considering the partial documentation

and the lack of knowledge of the volume of clandestine emigration, it has been estimated that in the sixteenth century 200,000 Spaniards went to the *Indias* and 305,000 in the seventeenth century (Mörner 1975: 64). During the eighteenth century, this number decreased to 55,000 emigrants (Hernández Sánchez-Barba 1954: 117-118). The total population of the Castilian crown was about 5 million inhabitants at the beginning of the sixteenth century, rising to 6,6 million at the end of the century. In 1712, the population was around 7,5 million. Thus, the proportion of the emigrated population was relatively significant (Hugon 2019).

Microstudies allow us to know the origin of these individuals and thus offer us a complete picture of the migratory phenomenon towards the New World. In the following, we compare several samples of migrants with our data in order to assess the representativeness of our sample relative to the migrants recorded in this other samples. Before 1520 it was possible to identify the place of birth of 5,481 individuals known to be in the *Indias*. The Andalusians were the largest group. Only the provinces of Sevilla and Huelva "*furnished over 30% of the total number of colonialists for the entire period*" (Boyd-Bowman 1956: 1156). It seems that this was the pattern during the next years, even centuries: Andalusia at the top followed by the regions of Extremadura, New Castile, Old Castile and Leon in the number of emigrants (Table 3.2)[40].

Based on a smaller sample of 1,263 emigrants for the years 1794-1796, Delgado Ribas (1982: 119) established the five most common regions of origin of emigrants as followed: Andalusia (24.6%), Castile and Leon (16.2%), the Basque Country (16.1%),

[40] Martínez Martínez (1993: 82) quantified 9,085 emigrants from Castile and Leon during the years 1517-1600 and a total of 11,345 during 1515-1700 (91).

Catalonia (15.1) and Galicia (11.2%). In our sample the migrants analysed were mostly from Andalusia, Castile and Leon which is in line with the literature (Figure 3.1).

How representative is our sample? Assuming that the numbers of migrants given by the literature for the sixteenth, seventeenth and eighteenth centuries are reliable, our sample represents 0,4%, 2% and 4%, respectively. For the two largest samples (Andalusia and Old Castile including Leon) we have assessed the share of rural and urban passengers. In Andalusia, 63.8% of the migrants were born in urban places and 36.2% from rural places. We need to consider that at that time Andalusia was the most urbanised region of Spain. From Old Castile and Leon, the shares were 36.7% and 63.3%, respectively[41].

The database of non-migrants has been collected from census and *padrones* (individual population counts)[42]. It includes different regions, rural and non-rural, for the following provinces in Spain: Almería, Ávila, Badajoz, Cáceres, Cádiz, Ciudad Real, Córdoba, Cuenca, Granada, Guadalajara, Jaén, La Rioja, Madrid, Málaga, Murcia, Navarra, Sevilla, Soria, Toledo, Valencia and Valladolid. The random sample contains 26,685 individuals of which 17,613 are men, and 9,072 are women, aged 23-62[43].

To estimate numeracy, the method used is the "age-heaping" technique (A'Hearn et al. 2009; Crayen and Baten 2010). This method is based on the share of individuals who are able to state their precise age in years instead of reporting an age rounded to a multiple of five. For instance, one person lacking the knowledge of their age could state

[41] Urban is defined as places with more than 5,000 inhabitants by the end of sixteenth century according to Reher (1994:3). The urban-rural share of Old Castile is based on Martínez Martínez (1993).

[42] See table 2.6 in appendix 2.8. This database has been previosly used by Pérez-Artés and Baten (forthcoming).

[43] 12,220 males and 5,674 females aged 33-62.

that she or he is 55 when she or he is actually 54. As it has been pointed out before that if the signature is used as a proxy for basic literacy, age heaping can serve as a proxy indicator for basic numeracy. The main findings of A'Hearn et al. (2009) are that the most basic mathematical skills diffused earlier than literacy and that there is a robust correlation of literacy and numeracy. They suggested an index that was later called the ABCC index[44]. It is a simple linear transformation of the Whipple index and enables the estimation of the share of individuals who state their age precisely:

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

The index ranges from 0 to 100 (100 means that everybody reports the correct age). The age groups used are 23-32, 33-42, 43-52, 53-62 and 63-72 years, due to the bias that younger and older groups could present (Crayen and Baten 2010).

As mentioned in the introduction, studies on the estimation of human capital in Spanish Latin America had never been addressed before 2010 due to the lack of data. Thus, numeracy allows us to estimate the basic mathematical skills of societies based on age data. Table 3.3 summarises all the samples in our database by birth decade, gender and migrants and non-migrants. The number of observations for the age groups between 23-32 and 33-62 years are shown to avoid possible bias as migrants usually consisted of young adult males between 20 and 30 years of age (Martínez Martínez 1993; Díaz Trechuelo 1990). In our sample, 61.3% of migrants are in the group of age 23-32. In total, we have 31,032 observations for the age group 23-62 and 19,046 for the age group 33-62. When comparing migrants and non-migrants birth decades are included from 1580

[44] "ABCC" are the initials of author's last names plus Gregory Clark, who commented on their paper.

since this is when the non-migrant sample starts. When only migrants are considered, we use the total sample: 7,220 observations.

One of the advantages of this research is to be able to estimate the level of human capital accurately. For example, we know the number of occupations of 3,879 emigrants and hence we could know their skills-levels. Only the abcc levels for the individuals including in the groups of "clergy", "highly qualified professionals" (administrators, governors, doctors) or "servants" can be calculated due to the number of observations (51, 93 and 3,605). The abcc index of these groups is as follows: highly qualified professionals (89), clergy (88) and servants (75). The abcc levels for professionals and clergy are in line with the literature of eighteenth century but the servant group seems to be slightly high (Tollnek and Baten 2017). However, more than 92% defined themselves as servants (71% of men and 21% of women) and this could lead to misunderstandings since "servant" in this context could have different meanings. For example, some individuals went to Hispanic America as servants of graduates and doctors for working as prosecutors or governors among others, namely, occupations that involve a high level of human capital. Others appeared as a servant of a master. It was probably because they did not get the license to travel. In this case they paid this lord in exchange for being able to travel with him on a servant's license (Martínez Martínez 1993). This practice continued even during the second half of the eighteenth century. Among the immigrants to Madrid looking for work, there were who offered themselves as servants of a master or family who went from Cadiz to America (Sarasúa 1994).

3.4 Analysis

Did the Spanish migrants present higher levels of numeracy than those who stayed? The positive self-selection of migrants in the internal migrations in Spain and Great Britain during the nineteenth and 20th century has been confirmed in terms of their literacy, heights and socio-economic status. In Great Britain, Humphries and Leunig (2009) studied a group of seamen born outside London in mid-nineteenth century. The authors found that the highest, literate and those who could remember the exact day, month and year when they were born, were more likely to leave London. Beltran and Salanova (2017) have showed that the literacy gap between Spanish migrants to Madrid and non-movers in Spain by province of origin was, on average in the period of time 1880-1887, 41% for men aged 16-30 and 37% for women in the same range of age. Quiroga (2003) has estimated that the literacy rate among the interprovincial migrants in Spain between 1893 and 1954 was 90%, while those who remained in their province of birth was 78%. If only the period of time between 1893 and 1899 is taken into account, this gap in literacy rates between movers and stayers (12%) was even higher: around 24% since the literacy rate of stayers was around 66%. Juif and Quiroga (2019) have proved that up to 1915, movers showed 15-20% higher literacy than stayers. After that and until mid-1920s, the gap shrank to 5-10%, converging both groups by 1950. According to the occupational groups, the share of white-collar workers was double among the movers and the share of students, professional and modern services was triple. According to heights, between 1893 and 1945 the height of movers within Spanish regions was on average one centimetre more than the height of stayers. This height gap reached almost three centimetres in the 1920s to early 1930s.

Our data confirms these previous findings: migrants were positively self-selected in terms of numeracy levels when comparing the two largest subsamples, Andalusia and Castile. As indicated before, I performed this analysis for the two age groups separately (23-62 and 33-62) since individuals at the age of 23 were more likely to emigrate than those aged 32: the assumption of similar representation of end digits (apart from heaping on multiples of 5) is violated. In any case, for the two age groups, we found a positive selection of emigrants. The difference is higher between the Andalusians (17.4% for individuals aged between 23-62 and 19.5% for individuals older than 33 years) than among the Castilians (16.5% and 18.9% respectively).

In order to estimate the selectivity of migrants more carefully, we carried out a linear probability model (LPM) and logit regressions. The LPM has the following specification and applies similarly to the logit model (Table 3.7 appendix).

$$Numerate_{itr} = \alpha + \beta_1 migrant_i + \beta_2 age2332_i + \beta_3 age4352_i + \beta_4 age5362_i + \mu_r + \gamma_t + \varepsilon_{itr}$$

$_I$ refers to each respective individual, $_t$ indicates the decade of birth, and $_r$ represents the place in which the individual was born. The main variable of interest is *numerate*, coded as 0 when the age stated was a multiple of five and 1 if otherwise. The variable *migrant* is equal to 1 if the individual is a migrant and 0 if not. *Age2332* is equal to 1 if the individual belongs to the age group between 23 to 32 years which applies the same way for *age4352* and *age5362*. The model includes region fixed effects (μ_r) that reflect the provinces of origin of Spaniards. We also control for time fixed effects (γ_t) for all half-century periods from 1600 to 1750. Finally, the equation allows for the constant term (α) and the standard errors (ε).

We tested different models, and in all of them, numeracy is high and correlates positively with being a male migrant (Table 3.5). Only females aged 23-62 appear slightly positively correlated. Columns 1 and 2 include all the observations (individuals aged between 23-62) and columns 3 and 4 include those that are older than 33. Columns 2 and 4 tested differences between migrants and non-migrant females and the rest of the columns include only males. In order to avoid potential biases because of the issue of the age of migrants, we control for all age groups.

Although it is difficult to make generalizations due to the period of time and geographic scope of my study, I would argue that the largest number of emigrants came from the most developed subpopulations with greater labour diversification (or were the most qualified within the less developed regions). One of the reasons why these individuals decided to go to the *Indias* was the "call of a relative". This is reflected in the documentation through the "call letters" that the emigrants already living in the *Indias* sent to their families. They usually illustrated a positive view of life in the colony and gave advise on how to do the crossing the best possible way (known in the classical literature on migration as chain migration). In addition, Hispanic America remained governed by the same monarch, with the same laws and the same behaviours as in Spain, but the hope of enrichment was added (Martínez Martínez 1993). This is supported by the literature: if the pull effect of the destination is stronger than the push effect from the provinces of origin, migrants are probably going to be positively self-selected due to their better life conditions (Juif and Quiroga 2019).

Figure 3.2 shows the evolution of numeracy of the Spaniards in the two countries where they went mostly, that is Mexico and Peru (44% and 22% respectively of our sample), by birth decades (1540-1710). The total number of observations is 4,355 (2,919

for Mexico and 1,436 for Peru). Emigrants who went to Mexico had a higher level of numeracy than those who chose to go to Peru. There was an increase in numeracy levels at the end of the sixteenth century. In 1580 we found the highest level of numeracy in both countries. Analysing book productions (considered as an alternative index of human capital) Peru, with 6 titles per million inhabitants, was also behind of Mexico with 8.5 during the eighteenth century. In terms of real GDP per capita, Mexico was ahead of Peru since 1650. Furthermore, the number of cities with population over 5,000 inhabitants was larger in Mexico in this year: 11 cities in Mexico and 6 in Peru (Arroyo Abad and Van Zanden 2016). Moreover, Mexico was the region of colonial Spain where markets, free labour and silver mining, were more developed. Also, wages were higher, and the working conditions were better there than in Peru (Salvucci 2014). During the next century, for the years that we have more than 100 observations, we see that these levels increase in both countries and stabilize in Mexico in the first two decades of the eighteenth century.

Although both countries had an important university (*La Real y Pontificia Universidad* of Mexico and the *Universidad de San Marcos* in Lima) in the sixteenth century, the school network was more widespread in Mexico than in Peru (Delgado Criado 1993). Around 1600 there were 36 schools, one university and one school-university in Mexico. In Peru there were also one university, and one school-university but in this case it have been accounted only 12 schools.

Furthermore, the number of cities with printing presses was larger in Mexico than in Peru (Figure 3.3)[45]. Mexico had five towns with presses by the end of the eighteenth century (Mexico, Puebla, Antequera, Guadalajara and Nueva Veracruz) while Peru had

[45] Convents and missions of religious orders are not shown on this map.

only one (Lima). Also, the number of print shops within the viceroyalty of Mexico was larger than in the Peruvian viceroyalty. Moreover, the imprints published in Mexico were more distinguished than those published in Lima, with better ink and greater industrial technique (Guibovich Pérez 2001; Del Palacio Montiel 2004)[46]. This could indicate a possible differentiated emigration between *Nueva España* and Peru where in the former the activities and works required people with greater human capital[47]. According to the occupations in our sample (1,040 for Mexico and 482 for Peru) the individuals in the "clergy" and "professional" groups were larger in Mexico than in Peru. In the first case these two occupational categories represented 3.6% of the total while in Peru it was 2.3%. In both cases "servants" were more than 90% of the individuals.

The religious orders had also more importance in Mexico than in Peru as literature has demonstrated. For example, in Mexico, regions where the Mendicant missions were active are positively correlated currently with a higher literacy rate of educational attainments until post-secondary levels (Waldinger 2017).

[46] The first printer in Mexico was Esteban Martín (in 1535) and the first printed work was *La escala espiritural para llegar al cielo* by San Juan Clímaco, translated by Fray Juan Estrada (Torre Revello 1940). In Puebla de los Ángeles it is not clear who was first printer. However, it is known that the press worked from 1642 to 1821 (although it was settled in 1640) producing 2,700 documents. In Antequera there was a printing workshop since 1687. Francisca Flores was the owner of it since 1720. In Guadalajara the first printer was Mariano Valdés (1792) and in Nueva Veracruz, Manuel López Bueno was the first official printer of the consulate since 1794. (Del Palacio Montiel 2004). In the Peruvian viceroyalty the first printing shop was founded in Lima in 1584 by the Italian printer Antonio Ricardo. Philip II through the Royal Decree on 22 August (1584) to the Viceroy and the Audience of Lima, ordered that a printing press be installed in the city (Torre Revello 1940; Guibovich Pérez 2001).

[47] We need to consider that the first years for which we have migrant departure data (not when they were born) are especially for the last two decades of the sixteenth century.

In *Nueva España* the Franciscans were stablished in 1523, Dominicans in 1526 and Augustinians in 1533. Finally, the Jesuits arrived in Mexico in 1572 continuing with the evangelization of the natives began by the previous religious orders. A year after their arrival, the Jesuits opened the first school in Mexico City (*Colegio de San Pedro y San Pablo*), founding up to 1751 other twenty schools outside the capital of the viceroyalty. The Franciscans had greater importance in *Nueva España* both in the number of schools and students, and in the results achieved. In 1531 there were almost twenty Franciscan convents to instruct the children of the most important people. In Peru there is not so much evidence. As in the rest of the colony, we could assume that, in general, there was a house next to each church as a school to teach the children of the *caciques*.

As for women's education, it is also in the *Nueva España* where we have the most direct and early references. In 1529 a Spanish midwife instructed the daughters of the lords of the region. Later the education of Indian girls was also established. In Peru, female education was mostly limited to instruct mestizos who were abandoned by their parents which was in this region a problem greater than in Mexico or Guatemala. It seems that there were also differences: in Mexico there is evidence that, at least the daughters of lords and principal American people, were taught to read and write while in Peru it was mostly house skills (Delgado Criado 1993)[48].

If we take into account the origins of emigrants in Mexico and Peru by region from the last decades of the sixteenth century to the first half of the eighteenth century,

[48] An example is the figure of *sor Juana Inés de la Cruz*, daughter of a *criolla* (born in America with European parents) and a Spanish captain, born in Mexico circa 1648. *Juana Inés de la Cruz, w*ho knew how to read and write at the age of three, decided to profess in the convent of San Jerónimo (only for criollas) in order to be an intellectual and have access to the culture, being an extraordinary poet. (Paz 1982).

the proportions are quite similar (Table 3.6). In Mexico, 65.8% of migrants were from Andalusia, 17.6% from Castile, 13.3% from the provinces of the north and 0,9% from the other Mediterranean area. In Peru these percentages were 74.2%, 17%, 7.4% and 0.5% respectively. The percentages that are missing to complete 100% correspond to the cases where we do not know the place of origin (2.4% in Mexico and 0.9% in Peru). Therefore, we can argue that apparently the origin regions of the migrants cannot account for these differences.

Were there high levels of inequality in Mexico to which the migrants have further contributed by their migration into the middle and upper classes? Figure 3.4 shows the numeracy levels of 731 people of Spanish origin and the numeracy levels of 992 indigenous people in Mexico during the first half of the eighteenth century. This is the most suitable comparison that can be done with the studies that we have about the native population's numeracy since in the native sample, including *Indios, mestizos, pardos* and other *castas*, we have in total 1,228 observations[49]. From 1680 to 1700, the levels of numeracy of both groups increased considerably although the numeracy level of natives was behind of those with Spanish origin. During the decade of 1680 the abcc level of Spaniards was 73% and during 1710 was 88%. For the same period of time, among the native group, these levels were 33% and 59%. Therefore, the gap between the two groups fell from 40% to 29% by 1710. During the second half of the eighteenth century, the gap between Spaniards from the peninsula and the rest of the population (including *indios* and the rest of *castas*) decreased to 13.5% in Mexico City (Calderón-Fernández et al. 2020: 19). In terms of heights also there were a reduction in the gap between European (*blancos*)

[49] The data comes from the 1740-4 censuses (Manzel et al. 2012).

and non-Europeans (*pardos)* in Southern Mexico since 1730s to 1780s: from about 4 centimetres to close zero (Dobado González and García Montero 2010: 266).

How large are the differences between migrants of European origin and indigenous people elsewhere in the world? As an international comparison in the Cape colony (current South Africa) the European settlers had higher levels of numeracy than non-Europeans during the late seventeenth century and the late eighteenth century, that is, more than 60% of abcc index (Baten and Fourie 2015). This difference indicates a more unequal society than Hispanic America were the higher gap in terms of abcc between Spaniards and natives was 40%.

Over time, these differences might have not disappeared but decreased as in the mining district of Pachuca, northeast of Mexico City. By 1520, labour shortage affected the forced recruitment of American workers. During the eighteenth century, the census of Real del Monte 1768 reveals that there were still social differences in the work done by the Americans. The occupations of the natives usually were the least qualified and most dangerous, but there were also qualified American workers as merchants, artisans and even a schoolteacher, musician and painter (Navarrete 2015). The same pattern of reduction of differences is found in the mines of the Real de Monte. Although there were laws prohibiting slave labour, there is evidence that it existed during the sixteenth century disappearing in the seventeenth century. Furthermore, in the work of the mines all the *castas* or ethnic groups were represented (Gaona Rivera 2019: 168)[50]. Moreover, religious orders played a role to reduce the differences among ethnic groups. In Mexico,

[50] 1798 the occupation of *barretero* (those who works with a bar, wedge or pick) was composed by: *criollos* (40%). *mestizos* (38%), *indios* (10%), *mulatos* (4%) and *castizos* (3%). Pawns were 40% of *mestizos*, 32% of *criollo*s and 23% of *indios*.

the Mendicant missions focused on the native population and the decrease of inequality (Waldinger 2017).

3.5 Conclusions

This paper contributes to the literature of migrations and human capital formation obtaining new empirical evidence for Hispanic America during the early modern period. The analysis was carried out at different levels. The main conclusion to be drawn is that Spanish migrants had higher human capital than those who stayed in the Iberian Peninsula from the sixteenth to the middle of the eighteenth century. In general, the numeracy of Spanish migrants was relatively high by the standards of the time (A'Hearn et al. 2009; Juif et al. 2019). This finding is in line with the results obtained by other researchers about positive migrant's self-selection in Great Britain and Spain during nineteenth and 20th century (Humphries and Leunig 2009; Quiroga 2003; Beltrán and Salanova 2017; Juif and Quiroga 2019).

Differences in numeracy levels of migrants are observed between those who went to Mexico and those who went to Peru. On average, migrants in Mexico had a higher level of human capital than those in Peru. These differences could be due to the religious orders that encouraged a higher network of schools in *Nueva España* and the viceroyalty characteristic: in Mexico the wages were higher, and the number of cities and the book production was larger than in Peru (Arroyo Abad and van Zanden 2016).

In a more detailed analysis, I assess the numeracy inequality among different ethnic groups. Levels of numeracy in Mexico between people of Spanish and native origins indicated a relative high inequality during the late seventeenth century, to which

the selectivity and relatively high level of numeracy of Spanish migrants might have further contributed. However, at the beginning of the next century these differences decreased. This result is consistent with the decrease differences in heights between whites and pardos in Mexico since 1730 (Dobado González and García Montero 2010).

3.6 References

Recopilación de las leyes de los reinos de las Indias. Mandadas imprimir y publicar por la majestad católica del rey Don Carlos II Nuestro Señor. Tomo Primero. Quinta Edición. Boix Editor, Impresor y Librero, calle de Carretas, Número 8. 1841.

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

Acemoglu, D. & Robinson, J. 2013. *Why nations fail. The origins of power, prosperity, and poverty*. New York, Crown Business.

Alvar Ezquerra, J. 2003. *Diccionario Espasa historia de España y América*. Madrid, Espasa Calpe.

Arroyo Abad, L., van Zanden, J. L. 2016. Growth under Extractive Institutions? Latin American Per Capita GDP in Colonial Times. *The Journal of Economic History*, 76 (04): 1182–1215.

Bakewell, P. J. 1997. *A history of Latin America. Empires and sequels,1450-1930*. Oxford, Blackwell.

Barro, R. J. 2001. Human Capital and Growth. *American Economic Review*, 91 (2): 12-17.

Baten, J. & 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(02): 279–305.

Baten, J., & Fourie, J. 2015. Numeracy of Africans, Asians, and Europeans during the early modern period: new evidence from Cape Colony court registers. *The Economic History Review*, *68*(2): 632–656.

Becker, G. S. 2008. *Human capital. A theoretical and empirical analysis, with special reference to education*. Chicago, The Univ. of Chicago Press, 3. ed.

Beltrán Tapia, F. J. & de Miguel Salanova, S. 2017. Migrants' self-selection in the early stages of modern economic growth, Spain (1880-1930). *The Economic History Review*, *70*(1): 101–121.

Boyd-Bowman, P. 1956. The Regional Origins of the Earliest Spanish Colonists of America. *Modern Language Association*, 71(5): 1152–1172.

Calderón-Fernández, A., Dobado-González, R., & García-Hiernaux, A. 2020. Numeracy in Central New Spain during the Enlightenment. *RHE/JILAEH*: 1–35.

Catholic University of America. 1697. *Educacion I.* New Catholic Encyclopedia. New York, McGraw-Hill: 148-157.

Coatsworth, J. H. & Summerhill, W.R. 2010. The New Economic History of Latin America: Evolution and Recent Contributions. In J. C. Moya (Ed.) *The Oxford handbook of Latin American history.* Oxford, Oxford University Press.

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

Crayen, D. & Baten, J. 2010. New evidence and new methods to measure human capital inequality before and during the industrial revolution: France and the US in the seventeenth to nineteenth centuries. *The Economic History Review*, 63 (2): 452–478.

Del Palacio Montiel, C. 2004. La imprenta y el periodismo en las regiones de México (1539-1820). *Comunicación y Sociedad*, 2:161-184.

Delgado Criado, B. 1993. La educación en la España moderna (siglos XVI-XVIII*)*. In Buenaventura Delgado (Coord.) *Historia de la educación en España y América*, Vol. 2, Madrid, Ediciones SM.

Delgado Ribas, J. M. 1982. La emigración española a América Latina durante la época del comercio libre (1765-1820) el ejemplo catalán. *Boletín Americanista*, (32): 115–137.

Díaz Trechuelo, M. L. 1990. *La emigración andaluza a América, siglos XVII y XVIII.* Sevilla, Junta de Andalucia.

Dobado González, R., & García 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.

Droller, F. 2018. Migration, Population Composition and Long Run Economic Development: Evidence from Settlements in the Pampas. *The Economic Journal*, (614)128: 2321–2352.

Elliott, J.H. 1984. The Spanish Conquest and settlement of America. In Leslie Bethell (Ed.) *The Cambridge History of Latin America*, 1: 149-206

Gaona Rivera, E. 2019. Trabajo, salarios y nivel de vida de los mineros de Real del Monte (México) en los siglos XVIII y XIX. *Dipòsit Digital de Documents UAB* (*Tesis Doctoral)*.

García-Abásolo, A. F., Quiles García, F. & Fernández, M. A. 2006. *Aportes humanos, culturales y artísticos de Andalucía en México, siglos XVI - XVIII*. Sevilla, Junta de Andalucía.

Guibovich Pérez, P. 2001. The Printing Press in Colonial Peru: Production Process and Literary Categories in Lima, 1584-1699, *Colonial Latin American Review*, 10 (2): 167-188.

Hernández Sánchez Barba, M. 1954. La población hispanoamericana y su distribución social en el siglo XVIII. *Revista de Estudios Políticos*, 78: 111–142.

Hugon, A. 2019. *La grande migration. De l'Espagne à l'Amérique: 1492-1700.* Paris, Vendémiaire.

Humphries, J. and Leunig, T. 2009. Was Dick Whittington taller than those he left behind? Anthropometric measures, migration and the quality of life in early nineteenth century London?. *Explorations in Economic History*, 46: 120–131.

Juif, D. & 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(2): 227–241.

Juif, D. 2015. Skill selectivity in transatlantic migration: The case of canary islanders in Cuba. *Revista De Historia Económica / Journal Of Iberian And Latin American Economic History*, 33(02): 189–222.

Juif, D., & Quiroga, G. 2019. Do you have to be tall and educated to be a migrant? Evidence from Spanish recruitment records, 1890–1950. *Economics & Human Biology*, 34: 115–124.

Juif, D., Baten, J., & Pérez-Artés, M.C. 2019. Numeracy of religious minorities in Spain and Portugal during the inquisition era. *Revista De Historia Económica / Journal of Iberian and Latin American Economic History*, 38(1): 147-184.

Macías Domínguez, I. 1999. *La llamada del Nuevo Mundo. La emigración española a América (1701-1750).* Sevilla, Universidad de Sevilla.

Manzel, K., Baten, J., & 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 Martínez, M. C. 1993. *La Emigración castellana y leonesa al Nuevo Mundo. 1517-1700*. Junta de Castilla y León, Consejería de Cultura y Turismo.

Menéndez Pidal, R. 1988. *Las Indias y la política exterior*. Madrid, Espasa-Calpe.

Mörner, M. 1975. La emigración española al Nuevo Mundo antes de 1810 un informe del Estado de la investigación. *Anuario de Estudios Americanos*, 32: 43–131.

Moya, J. C. (Ed.) 2010. *The Oxford handbook of Latin American history*. Oxford, Oxford University Press.

Navarrete, D. 2015. Trabajadores, artesanos y trajinantes. La participación indígena en la economía minera del México colonial: Real del Monte en la segunda mitad del siglo XVIII. *Nuevo Mundo Mundos Nuevos*. DOI : https://doi.org/10.4000/nuevomundo.67758

Núñez, C.E. 1992. *La fuente de la riqueza. Educación y desarrollo económico en la España Contemporánea*. Alianza Universidad.

Paz, O. 1982. *Sor Juana Inés de la Cruz o Las trampas de la fe*. Seix Barral.

Peña Díaz, M. 2012. Del confín de Europa al corazón del mundo: (ss. XVI-XVIII). In Manuel Peña Díaz (Ed.), *Breve historia de Andalucía*: 163-222, Fundación Centro de Estudios Andaluces.

Quiroga, G. 2003. Literacy, education and welfare in Spain (1893–1954). *Paedagogica Historica*, *39*(5): 599–619.

Ramos Pérez, D. & Lohmann Villena, G. 1985. *América en el siglo XVII. Los problemas generales*. Madrid, Rialp.

Reher, D.S. 1994. Ciudades, procesos de urbanización y sistemas urbanos en la península ibérica, 1550-1991. In M. Guàrdia, F.J. Monclús & J.L. Oyón (Eds.), *Atlas Histórico de las ciudades europeas. Península Ibérica*: 1-30, Salvat-Centre de Cultura Contemporàni de Barcelona: Barcelona.

Restall, M. & Lane, K. E. 2011. *Latin America in colonial times*. Cambridge, Cambridge University Press.

Saavedra Inaraja, M. 2008. *La forja del Nuevo Mundo. Huellas de la Iglesia en la América española*. Madrid, Sekotia.

Salvucci, R. 2014. Capitalism and Dependency in Latin America. In Larry Neal and Jeffrey G. Williamson (Eds.) *The Cambridge History of Capitalism*: 403-430. Cambridge, Cambridge University Press.

Sánchez Alonso, B. 2007. The Other Europeans: Immigration into Latin America and the International Labour Market (1870–1930). *Revista De Historia Económica / Journal Of Iberian And Latin American Economic History*, 25(03): 395–426.

Sarasúa, C. 1994. *Criados, nodrizas y amos. El servicio doméstico en la formación del mercado de trabajo madrileño, 1758-1868.* Siglo XXI de España editores.

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

Sokoloff, K. L. & 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.

Stoeckel, A. 1976. *Presidents, professors, and politics: the colonial colleges and the American revolution*. Ball State University. Department of History.

Stolz, Y., Baten, J. & Botelho, T. 2013. Growth effects of nineteenth-century mass migrations: "Fome Zero" for Brazil?. *European Review of Economic History*, 17(1): 95–121.

Taylor, A. M., & Williamson, J. G. 1997. Convergence in the age of mass migration. *European Review of Economic History*, 1(1): 27-63

Tollnek, F & Baten, J. 2017. Farmers at the heart of the "human capital revolution"? Descomposing the numeracy increase in early modern Europe. *The Economic History Review*, 70(3): 779-809

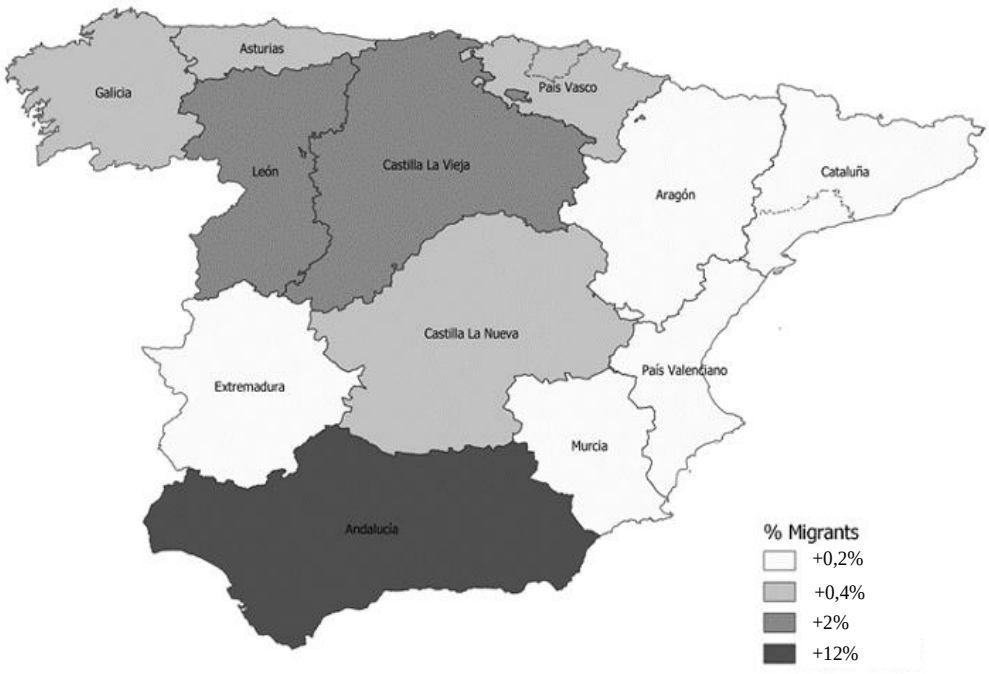
Torre Revello, J. 1940. *El libro, la imprenta y el periodismo en América, durante la dominación española.* Publicaciones del Instituto de Investigaciones históricas, Buenos Aires: Casa Jacobo Peuser.

Waldinger, M. 2017. The long-run effects of missionary orders in Mexico. *Journal of Development Economics*, 127: 355–378.

Williamson, J. 2009. History without Evidence: Latin American Inequality since 1491. *NBER Working Paper* Series, Nº14766.

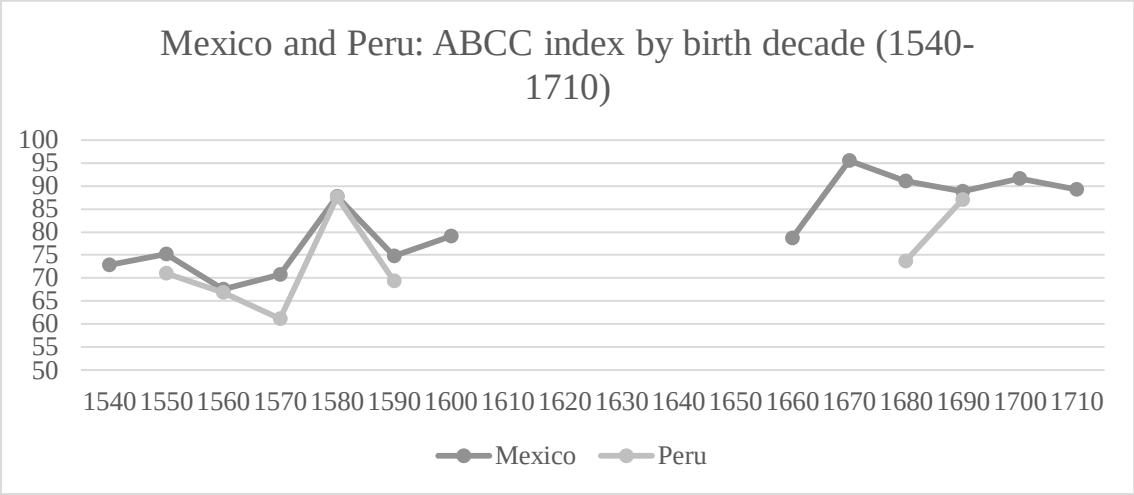
3.7 Figures and tables

Figure 3.1 Origin of emigrants to Hispanic America in our sample (1540-1750 birth decades)



Source: See table 3.1

Figure 3.2 Spaniards in Mexico and Peru: ABCC index by birth decade (1540-1710)

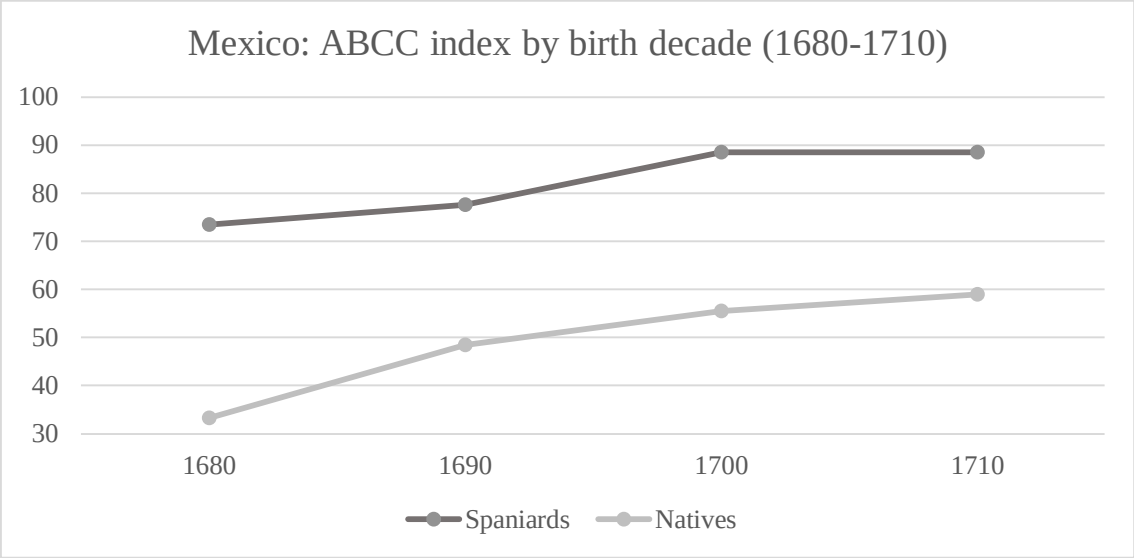


Source: see Table 3.1

Figure 3.3 Schools, printing presses and main universities in Mexico and Peru (16th-18th century)

Source: HGIS de las Indias and Catholic University of America (1967: 148-157)

Figure 3.4 Spaniards and indigenous Mexico: ABCC index by birth decade (1680-1710)



Sources: Mexico see table 3.1 and Hidalgo/Guanajuato/Oaxaca Manzel et al. (2012)

Table 3.1 Sources and Number of observations in my sample

Source	*Number of Observations*
García-Abásolo et al. (2006)	446
Macías Domínguez (1999)	1,329
Martínez Martínez (1993)	1,024
Díaz Trechuelo (1990)	4,421
Total	7,220

Table 3.2 Passengers to Hispanic America 1493-1600, by origin

	1493-1519		1520-1539		1540-1559		1560-1579		1580-1600		1493-1600	
	T	%	T	%	T	%	T	%	T	%	T	%
Andalusia	2,172	39.6	4,247	32.0	3,269	36.1	6,547	37.2	3,994	42.0	20,229	36.9
Extremadura	769	14.0	2,204	16.6	1,416	15.7	3,295	18.7	1,351	14.2	9,035	16.5
New Castile	483	8.8	1,587	12.0	1,303	14.4	3,343	19.0	1,825	19.2	8,541	15.6
Old Castile	987	18.0	2,337	17.6	1,390	15.4	1,984	11.3	970	10.2	7,668	14.0
Leon	406	7.4	1,004	7.6	559	6.2	875	5.0	384	4.0	3,228	5.9
Basque Country	257	4.7	600	4.5	396	4.4	515	2.9	312	3.3	2,080	3.8
Foreigners	141	2.6	557	4.2	332	3.7	263	1.5	229	2.4	1,522	2.8
Galicia	111	2.0	193	1.5	73	0.8	179	1.0	111	1.2	667	1.2
Val., Cat.+ Bal.	40	0.7	131	1.0	62	0.7	113	0.6	55	0.6	401	0.7
Aragon	32	0.6	101	0.8	40	0.4	99	0.6	83	0.9	355	0.6
Murcia	29	0.5	122	0.9	50	0.6	96	0.5	47	0.5	344	0.6
Navarra	10	0.2	71	0.5	81	0.9	112	0.6	52	0.5	326	0.6
Asturias	36	0.7	77	0.6	49	0.5	90	0.5	71	0.7	323	0.6
Canarias	8	0.1	31	0.2	24	0.3	75	0.4	24	0.3	162	0.3
Total	**5,481**	**100**	**13,262**	**100**	**9,044**	**100**	**17,586**	**100**	**9,508**	**100**	**54,881**	**100**

Source: Boyd-Bowman (1988, 606)

Table 3.3 Nº individual observations by sample and birth decades

BDEC 1580-1750		
Gender	Sample	Ind. Obs.
Age group 23-62		
Male	migrant	3,054
Female	migrant	1,293
Male	non-migrant	17,613
Female	non-migrant	9,072
Age group 33-62		
Male	migrant	663
Female	migrant	489
Male	non-migrant	12,220
Female	non-migrant	5,674
BDEC 1540-1750		
Age group 23-62		
Male	migrant	4,534
Female	migrant	2,686

Note: Individual observations for migrant and non-migrant age group 23-62 (1580-1750 b.d.): 31,032
Individual observations for migrant and non-migrant age group 33-62 (1580-1750 b.d.): 19,046
Individual observations for migrants age group 23-62 (1540-1750 b.d.): 7,220

Table 3.4 Selectivity of migrants (ABCC migrants-ABCC non-migrants)

BDEC 1580-1750	ABCC	Nº Obs.	Age Group
Andalusia			
Migrant	80.5	2,910	23-62
Non-migrant	61.6	13,718	
Selectivity (ABCC migrants — ABCC non-migrants)	**18.9**		
Migrant	80.2	874	33-62
Non-migrant	58.6	9,430	
Selectivity (ABCC migrants — ABCC non-migrants)	**21.6**		
Castile			
Migrant	92.3	767	23-62
Non-migrant	78.7	10,924	
Selectivity (ABCC migrants — ABCC non-migrants)	**13.6**		
Migrant	92.1	146	33-62
Non-migrant	77.1	6,978	
Selectivity (ABCC migrants — ABCC non-migrants)	**15.0**		

Source: see Table 3.1 and table 2.6 in appendix 2.8

Table 3.5 Migrant's skill selectivity in Linear Probability Model (LPM)

	(1)	(2)	(3)	(4)
Samples included	**Males aged 23-62**	**Females aged 23-62**	**Males aged 33-62**	**Females aged 33-62**
Migrant LA	28.39***	8.52*	37.00***	10.71
	(0.000)	(0.087)	(0.000)	(0.127)
Age 2332	1.24	3.61**		
	(0.251)	(0.014)		
Age 4352	-3.05**	-8.54***	-2.31*	-8.49***
	(0.014)	(0.000)	(0.065)	(0.000)
Age5362	-2.19	-22.72***	-0.90	-19.91***
	(0.199)	(0.000)	(0.639)	(0.000)
Constant	33.06	-0.91	-20.11	-0.12
	(0.222)	(0.920)	(0.555)	(0.992)
Observations	20,667	10,365	12,883	6,163
Adjusted R-squared	0.05	0.03	0.06	0.03
Time FE	YES	YES	YES	YES
Region FE	YES	YES	YES	YES

Note: The dependent variable is "not heaped age" (more likely to be numerate). The constant refers to male non-migrants aged 33-42. We multiply the coefficient of the regressions by 125 to report percentages and to adjust them for the 20% of ages that were truly multiples of five, given a normal age distribution. Robust p-Values are given in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 3.6 origins of emigrants in Mexico and Peru by region from the last decades of the 16th century to the first half of the 18th century (%)

	Mexico (%)	Peru (%)
Andalusia	65.8	74.2
Castile	17.6	17
Provinces of the north	13.3	7.4
Other Mediterranean areas (excluding Andalusia)	0.9	0.5

Source: see Table 3.1 and table 2.6 in appendix 2.8

3.8 Appendix

3.8.1 Migrant's skill selectivity in Logit Model (Marginal effects reported)

Table 3.7 Migrant's skill selectivity in Logit Model (Marginal effects reported))

	(1)	(2)	(3)	(4)
Samples included	**Males aged 23-62**	**Females aged 23-62**	**Males aged 33-62**	**Females aged 33-62**
Migrant LA	27.93***	8.65*	35.18***	10.78
	(0.000)	(0.079)	(0.000)	(0.122)
Age 2332	1.34	3.73**		
	(0.235)	(0.015)		
Age 4352	-3.11**	-8.72***	-2.43*	-8.79***
	(0.016)	(0.000)	(0.068)	(0.000)
Age5362	-2.10	-24.22***	-0.91	-21.44***
	(0.244)	(0.000)	(0.661)	(0.000)
Observations	20,664	10,353	12,867	6,148
Pseudo R2	0.0269	0.0269	0.0269	0.0269
Time FE	YES	YES	YES	YES
Region FE	YES	YES	YES	YES

Note: The dependent variable is "not heaped age" (more likely to be numerate). The constant refers to male non-migrants aged 33-42. We multiply the coefficient of the regressions by 125 to report percentages and to adjust them for the 20% of ages that were truly multiples of five, given a normal age distribution. Robust p-Values are given in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

4. Human Capital, schooling and child labour in New Castile (Spain) in the 18th century

No discussion of human capital can omit the influence
of families on the knowledge, skills, values,
and habits of their children (Becker 1964 [1994]:21)

Abstract

This paper analyses schooling and child labour in 22 towns of central Spain circa 1750, using the Cadaster of Ensenada as a source. This study seeks to shed light on the determinants that affected family strategies for subsistence, focusing on child labour and schooling. It examines whether there were behavioural differences depended on family characteristics, such as the occupation of the head of the household and their level of human capital. With a database of 4,204 families from the pre-industrial era in Castile, family size, the birth order of the children, the age of parents, mother's job or a textile factory in the town are shown to be decisive factors in explaining the decision to commit a child to labour. On the contrary, the results suggest that the heads of families with more human capital and greater earnings were more likely to send their children to school. Moreover, the supply of teachers at a municipal level played a positive role on the schooling of boys and girls.

4.1 Introduction

Through the study of the transition of literacy in Spain at a provincial level, Núñez (1992) confirmed that the lack of education and, especially illiteracy, hampers economic development. In a recent work, Beltrán-Tapia et al. (2019) claimed that this relationship between literacy and development in the long run is only detected in the twentieth century. They argued that the literacy process was financed until 1900 by the municipalities, which could have economic, social or geographical different conditions. Therefore, a municipal level study would be required to accurately understand the relationship between both variables. Even on a smaller scale, it is known that families played a role. Reis (2005) proved that from the Modern Age literacy arose (or not) by the decision of the families. In a society where the State did not centralize, regulate or finance a universal educational system, there were some constrains of elementary education as the uneven distribution of schools or the poor quality of many teachers. However, Reis (2005: 204) pointed out: "the most important one was clearly the cost that education entailed for the individual or the family". Sarasúa (2002b) studied for the first time the disaggregated expenses of families on the schooling of their sons and daughters in Spain during the nineteenth century founding that many families were keen on paying for their daughters to attend school. However, the lower public funding of girls' schools was one of the reasons why girls were less educated and literate than boys. Apart from the fact that the acquisition of human capital by the children meant a cost for the family, child labour was a subsistence strategy that not all families could reject. Hence, what we do know about schooling and child labour?

The historiography of childhood made substantial progress in the late 1960s with Philippe Ariès' 1962 book: *Centuries of Childhood: A Social History of Family Life*[51]. Over the subsequent decades, there have been several studies dedicated to deepening this issue, especially since studies on women have gained prominence in historical analysis – the historiography of women and children are highly interrelated. Both historiographic trends suffer from data censoring due to the difficulty in finding sources where women and children are recorded. In this way, the recent surge in studies on women has helped the proliferation of studies on children. Although child labour was most predominant in Britain, it was also essential in countries such as Belgium, France, Prussia and the United States – the countries that also pioneered the industrial process. In sum, children's hands were important in the early phases of industrialisation, especially in textiles and mining (Humphries 2003).

In Spain, childhood history has been studied mainly from historical demographers as well as historians of medicine and education (Borrás Llop 2002b)[52]. More recently, researchers in economic history have become increasingly interested in the study of child labour, mainly focusing on the ages when boys and girls began to work as well as the jobs that they undertook, paying specific attention to gender differentiation. For the eighteenth century, there are studies on Old and New Castile (Hernández 2013; Sarasúa 2013) whereas, for the nineteenth and twentieth centuries, studies on Madrid, Catalonia, rural Spain, Andalusia and La Coruña exist (Borrás Llop 2002c 2002a 2005; Borderías 2013; Campos Luque 2014; Camps 1995; Muñoz Abeledo 2012). These studies generally

[51] Originally published as *L'Enfant et la vie familiale sous l'Ancien Régime* (1960). This book was not recognised until social history started to be important in the late 1960s and early 1970s (Cunningham 2014).

[52] For further details see (Borrás Llop 2002b)

conclude that the average age for entering the labour force was ten years, although many children began to work earlier. Girls started to work young, particularly those whose work did not involve physical strength (see Figure 4.1). These studies also show clear evidence that boys were mostly engaged in agriculture whereas girls mostly worked in industry, especially in textile domestic manufacturing and domestic services. In the United Kingdom, child labour has recently been extensively studied, mainly in the context of the industrial revolution. Using the autobiographies of men, Humphries (2013) demonstrated that child labour participation rates increased during the era of industrialisation[53]. Analogous to Spain, evidence from the United Kingdom showed that child labour under the age of ten did occur, and that work was almost universal by the age of fifteen. In this study child labour refers to all those remunerated jobs for the market including those undertook on the family unit, either it was agricultural or manufacturing work, whose production was destined for the market. The children could be paid in money but generally the remuneration was in kind: food or clothing (Sarasúa 2013). As for school children, they also used to work for their families in all kinds of tasks. However, it was self-consumption and is not included in this definition of child labour.

Regarding the decisions that parents made towards the work and the education of their children, for the Spanish case we only have empirical evidence from industrial Catalonia. Borrás Llop (2002c) analyses the occupations of the heads of families and their effects on both schooling and child labour for the Catalan region *Vallés Occidental*, during the late nineteenth and early twentieth centuries. In terms of total time spent at

[53] Humphries (2013) found that the rise in child participation rates occurred during the years 1791 and 1820 when the sons of miners, factory workers, outworkers, casual workers and soldiers started work, on average, below the age of 10.

school, in a public school located in the municipality of *Sant Pere de Terrassa* (Barcelona), the children of the weavers were those who attended school for the greatest number of months, with an average of 61.8, while the children of urban day labourers only attended for 3 months on average (Borrás Llop 2002c: 242). An explanation for this could be that weavers owned several looms that employed family labour and parents could therefore afford to send their children to school. Camps (2002) argued that it was the technological change, the demographic transition and the impact of the modernization of labour markets the reasons behind the decrease of child labour rates between the nineteenth century and the first third of the twentieth in Catalonia.

This paper contributes to the schooling and child labour literature using data on New Castile circa 1750. This study does not simply compare rates of child labour by ages, occupation and gender since, as stated above, there already exists a considerable body of literature on these issues. The aim of this research is rather to analyse the parental decision of whether or not to opt for child labour or schooling according to family background, since there are not studies on this issue for pre-industrial Spain that take explanatory variables into account – such as the occupation of the head of the family, their level of human capital, the size of the family, the birth order of the children and the ratio of school-children to teachers or the cost of school (at a municipal level). Furthermore, also it offers new insights into the connection between the individual, family and their socio-economic situation, a topic of interest among the current economic history's researches. These studies not only take into account the institutional framework and economic development, but also "human capital and education, personal qualifications and skills, age, marital status, individual and family income levels" as factors that affected the strategies to tackle the economic crises of the eighteenth and nineteenth century (Martini and Borderías 2020:7).

The database used in this paper comes from the Cadaster of Ensenada, carried out during the middle of the eighteenth century in the Crown of Castile. Despite the limitation that not all male heads of households provided the occupations of their wives and children in the survey, it was possible to reconstruct the child labour and schooling variables at a family level. The regional units analysed are 22 municipalities corresponding to the provinces of Albacete, Ciudad Real, Toledo, Madrid and Guadalajara (inland Spain). These municipalities have been completely included, avoiding possible bias in the results. The sample extends from small villages with around 200 inhabitants to cities with more than 5,000 inhabitants, resulting in a dataset of 4,204 families (4,005 boys and 3,661 girls aged between 5-14) [54].

The structure of the paper is as follows: the next section contextualises child labour and schooling in Spain during the eighteenth and nineteenth centuries. Section 4.3 describes the methodology and the sources used. In section 4.4, the results are presented and discussed. Finally, section 4.5 summarises the main conclusions of the paper.

4.2 Child labour and schooling in Spain during the 18th century

Child labour was a common practice for family subsistence everywhere and in Spain as well until the early decades of the twentieth century. In the case of Barcelona,

[54] The database used for this paper belongs to professor Carmen Sarasúa to whom I am very grateful. The original sources are located in the AHP (*Archivo Histórico Provincial*) of Ciudad Real, Hacienda section, Catastro de Ensenada; AHP of Guadalajara, Hacienda section, Catastro de Ensenada; AHP of Albacete, Hacienda section, Catastro de Ensenada; AHP of Toledo, Hacienda section, Catastro de Ensenada.

children began to work in the factory, generally, at the age of 10 contributing with additional income to the family economy. It was then when mothers at the age of 30-35 years of age left the paid activity. In other words, when the first child reached the age of contributing with its salary. We have to take into account that women's incomes did not increase throughout their active lives, while those of male children and adolescents were higher than their mothers and increased throughout their working life. Furthermore, even in twentieth-century Barcelona, child labour remained an essential factor in supporting ageing parents. The wages of children constituted the largest part of family income when fathers were older than fifty years of age (Camps 2002)[55].

Likewise, child labour was a way of learning a trade. Until the nineteenth century, factories did not require formal education of their employees. The human capital necessary was as simple as knowing how to use the tools, which could be done on the job. Therefore, for working families, sending children to factories to begin practising a trade could result in a professional career in the textile industry (Camps, 2002)[56]. Also, in Sabadell and Terrassa (Barcelona), the participation of children in labour has been confirmed since the nineteenth century. Although it is difficult to estimate the exact number of working children due to the under-registration, it is known that from the establishment of the Spanish manufacturing industry until the beginning of the twentieth century the presence of children in factories was very common (Borrás Llop 2002c).

[55] A female spinner at the age of 31 earned the same as a male spinner at the age of 16: average income of 6 *pesetas* per day. A female factory worker aged 33 earned an average of 5.63 pesetas a day, while a man aged 27 earned 7.50 pesetas per day doing the same job (Camps 2002: 278-279).

[56] This fact is not only true for the factory of Barcelona during the nineteenth and twentieth century. Also, it could apply to the other jobs and economic sectors.

When did the regulation of child labour begin? The Benot Act of 1873 was the first act to forbid the work of children under ten, or under nine if the child could read and write, in industry and mining. In this context, children aged between 10 and 14, could work six hours in industrial establishments and eight hours in commercial establishments. However, this law never was applied (Borrás Llop 2019). Agricultural child labour continued to take place throughout Spain and was not regulated until 1934, then affecting children under fourteen years of age (Borrás Llop 2002a). In Britain, the first child labour law dated from 1867 and made it one of the few European countries that prohibits agricultural wage labour for children under eight years and limited it of those under ten. In Spain, it was at the age of ten when child labour rates intensified and resulted in the desertion of schools. This was also when a gender gap in labour emerged: males worked full time in agricultural work and females engaged in seasonal agricultural tasks (Borrás Llop 2002b).

In these societies, the opportunity costs of sending the offspring to schools were high. All labour-intensive production benefited from child labour, including industrial activity and especially in textiles (Núñez 1992). In eighteenth-century La Mancha, of the total number of children under fifteen for whom an occupation was reported, 65% of boys worked in agriculture and with livestock while 84% of girls worked in the textile manufactures (Sarasúa 2013). The same pattern existed in Old Castile. Of the total female active population between 6 and 15 years of age, 2.9% worked in the primary sector and 93.2% in the secondary sector. For boys these percentages were 47.0% in the primary sector and 43.7% in the secondary (Hernández 2013: 110). These data emphasise gender

differentiation in the type of labour that the children carried out as well as the importance of their contributions to the agricultural and textile industries[57].

Regarding schooling, the first Spanish census that offers direct global data on the number of first-letter teachers (elementary school for boys), girls' teachers and the number of students is the Godoy Census of 1797 (although the Cadaster of Ensenada allows to partially reconstruct this data). Despite its limitations, it portrays the basic picture of formal elementary education. The main contribution of this census is the demonstration that the number of male schools and male schooling rates was higher than those for females (Laspalas Pérez 1991). The issue with this segregation is that the knowledge taught was also different. Girls' schools were in several cases private houses of women who taught girls to sewing, embroidering, or lace-making, resulting in future occupational segregation between women and men as shows Figure 4.2. The picture portrays a group of girls learning the activities described above outdoors. Therefore, since the eighteenth century, the process of differentiation between boys and girls began in schools although in rural areas many schools were common to both, boy and girls (Sarasúa 2002a). This situation explains why, in the nineteenth century, there were more illiterate women where there were a larger number of female schools, whereas female literacy was higher where there was a larger number of mixed schools. Moreover, despite the fact that the wages of female teachers were lower than those of male teachers and that many families were, in fact, ready to pay for their daughters to attend school, the lack of public financing of girls'

[57] Although there are cases as in Antequera (Málaga) in 1857, where there is evidence of boys and girls working as day labourers, seamstresses, weavers or spinners from four years of age (Campos Luque 2014).

schools was a drawback for the enrolment in schools and literacy rates for girls, as stated above (Sarasúa 2002b).

In 1860, according to the first modern population census, the national male literacy rate was 42 per cent, while the female rate was only 12 per cent, although this varied regionally. At the regional level, the highest female literacy rate was in Madrid, at 35 per cent, while the corresponding male literacy rate was 65 per cent. In Galicia, where female literacy was the lowest in the country at only 5 per cent, the male literacy rate was 44 per cent. In all of Spain, the lowest male literacy rates (25 per cent) were recorded in both Eastern Andalusia and in the provinces of the east. In these regions, female rates were 10 and 8 per cent, respectively (Núñez 1992: 108-111). In all cases, male literacy rates exceeded female literacy rates. This feature is described in the novel *"La Tribuna"* written by Emilia Pardo Bazán in 1882. The countess of Pardo Bazán explained in *"La Tribuna"* the environment for female workers in a cigar factory in La Coruña. In this literary work, only one of these female workers was able to read and was in charge of reading the press aloud to the other women working in the factory.

The time that children spent in school was mostly dependent on the type of job that they had to do, which was usually defined by the municipality where they lived or the occupation of their parents. Child farm labour was the factor that impacted on monthly absenteeism rates the most, due to its seasonality, even in the 1930s (Borrás Llop 2005). It was not until 1837 when the first regulations on education were implemented in Spain, the Someruelos Act and the Montesino Regulation. According to these regulations, all villages should have a public school for elementary education for boys between six and nine years of age and, when the villages could afford it, for girls. However, compulsory schooling was not mentioned (Mallorquí-Ruscalleda 2019). It was two decades later

when the Moyano Act (1857) established compulsory schooling between the ages of six and nine (Núñez 1992). Then, before the nineteenth century when regular school attendance was not required, it was easier to combine schooling with labour in rural areas than in urban areas, because of the seasonal nature of agricultural activities (Borrás Llop 2002b). Likewise, in agricultural municipalities with low demand for manufacturing, school desertion occurred at a later age. The girls who were engaged the most in industry were most affected by this phenomenon. If certain agricultural activities were characterised by promoting absenteeism in line with agricultural cycles, they also allowed children more intense, stable and lasting schooling. As industrial tasks required permanent labour, girls were the most affected in terms of schooling, compounded by the fact that they also had to take over domestic activities (Borrás Llop, 2002c). However, it would be incorrect to associate non-schooling with regions of large farming. As Borrás Llop (2005: 391) has shown:

> *Poor schooling rates occurred in very different farming areas: dry farming (cereals, vineyards and olive trees), mainly in the south of Spain; intensive farming areas (fruit and horticultural) in the east ; and in part of the wet areas of Spain (Galicia and a section of the Cantabrian coast).*

Furthermore, aspects of schooling at that time differ from the model of contemporary schooling. In Old Castile and most of the Northern regions, temporary schools were opened during the winter months, coinciding with the decrease in demand for agricultural work when child labour was not essential for the family. Informal schools also existed and private learning also took place; for example, boys acquired academic instruction with priests and girls did the same in convents. However, while female religious orders focused on primary education, male religious orders mostly engaged in

secondary education, where families were ready to pay high fees for their boys (Sarasúa 2002b).

4.3 Data and Methodology

The source used to reconstruct the child labour data at the family level is the Ensenada Cadaster, a unique source for studying the eighteenth-century society and economy of the Crown of Castile, constituting three-quarters of current territory of Spain. It was carried out between 1750 and 1756 in 90 cities and in more than 15.885 towns and villages. The name of the cadaster was due to its promoter: the first Marquis of Ensenada, Secretary of the Treasury, who had the aim of unifying the fiscal system. This process had two levels of investigation, at the individual level and at the municipal level. At the individual level, households had to declare their properties and incomes as well as their names, marital status, profession and age in a document called *Memorial*[58]. At the municipal level, a survey of 40 questions regarding various aspects of the population was carried out, known as *Respuestas Generales.* The group responsible for the investigation consisted of an intendant, a royal notary, assistants of the royal notary, a geometer, several surveyors, a legal advisor and a bailiff, among others (Camarero Bullón 2002).

[58] The memorials used in this research are as follows: "I belong to the General estate, my trade fuller, married, my family is formed by myself, 46 years old, Ynés López Zamorano, 40 years old. I have four daughters, Agustina, 20, her occupation weaver, Isabel, 13, her occupation spinning, María, 11, her occupation going to sewing school, María Teresa, 2 months" Archivo Histórico Provincial de Ciudad Real, Ciudad Real Ensenada section, Antonio López Rufián, Campo de Criptana, box 502, declaration 77 (Sarasúa 2019: 483).

Although information about others living in the household was requested, "most householders did not declare the occupations of their wives or children because they were not asked to do so, since any subsistence wages earned by wives and children would not be taxed" (Sarasúa 2019: 483). The towns and cities studied in this paper must meet the criterion of at least 15 per cent household's declarations include the occupation of the family members and not exclusively that of the head of household in the census. However, most of them involve at least 40 per cent (Sarasúa 2013: 2019). Figure 4.3 shows the map of the towns included.

The sample consists of 4,005 boys and 3,661 girls aged between 5-14 resulting in a total of 4,204 families analysed. The places studied are 22 towns that belong to the current provinces of Guadalajara, Madrid, Toledo, Ciudad Real and Albacete. The sample extends from small villages with around 200 inhabitants to cities with more than 5,000 inhabitants, such as Guadalajara and Almagro. The only provincial capital included in the analysis is Guadalajara, the other cities shown in bold on the map are used in order to distinguish between the different provinces.

As mentioned, in Guadalajara and Brihuega, there were textile royal factories. The factory of Guadalajara was founded in 1719 and it was closed in 1822. The subsidiary of Brihuega was built in 1750 (López Barahona 2020). The occupational structure by sector also differs among the sample: some towns such as Villamanrique del Tajo, Quintanar or Ajofrín were mostly industrial and others such as Villaviciosa, Bolaños or Pedro Muñoz were largely agricultural, according to the occupations of the inhabitants. The only city where the tertiary sector was the largest by labour force participation was Guadalajara

(45.9%)[59]. This occupational data allows for an analysis of different places with different economic specialisations.

The first analysis performed is based on the occupation of the male head of each household. When women the females were head of households it meant that they were widows (only a few cases and different characteristics). These occupations have been classified following the Armstrong codification scheme (Armstrong 1972). In total there are six different occupational categories. The first one it is the professional group or those with higher education dedicated to local and national government services or professions as a doctor. In second place, semi-professionals or non-manual, are mostly dealers. Individuals with occupation in industry, craftsmen or sellers are in the groups of skilled. The fourth groups contain semi-skilled individuals. Day labourers are included in the fifth group. Finally, farmers are assigned to their own group. Farmers were a special group of primary sector workers because they enjoyed a higher income as well as a certain decision power about how much food would be sold on the market and how much food could be consumed by the farmer family. In previous research, it has been shown that these two factors could contribute to farmers having a higher level of human capital (numeracy) than other primary sector workers in agriculture and livestock (Tollnek and Baten 2017).

The ratio of first-letter teachers (primary school teachers) to school-age children has been estimated, differentiating between the teachers of boys and girls. In addition to the self-reported occupations from the *memorials*, the number of male first-letter teachers (primary school) in the *Respuestas Generales* have been confirmed. For example, question number 25 asked about the public spending that the town council had to pay, while questions 32 and 33 provided a summary of the number of professions that were

[59] Data calculated from Sarasúa (2019: 495) Table 4.

carried out in each town, whether professional or manual and their earnings. Using this information, it was possible to double check the number of first-letter teachers in each town, comparing them with the data obtained through the responses from the head of each household[60]. Furthermore, using these sources, it has been possible to know the salary they received. For the ratio of female teachers to school-aged girls, only women reported as teachers in the memorials have been considered since they were not recorded in the *Respuestas Generales*. To calculate this ratio, one finds that the minimum and maximum ages of children enrolled in elementary school had not yet been established. Rather, it was the family's decision to send or stop sending their children to school when they thought it was appropriate. In addition, as noted above, there was also no rule regulating school attendance. In this paper, it is assumed that the school-age children were between five and ten years of age. Therefore, the number of teachers in each town has been divided by the number of children aged five to ten, distinguishing between boys (male teachers) and girls (female teachers).

4.4 Determinants of child schooling and child labour

Before analysing the family's decision towards their children, the rates of working and schooling children in our sample are shown in table 4.1. These are minimum rates as calculated from the declarations of householders, who often unrecorded the paid occupations of their wives and children. As explained before, as schooled child is

[60] Although the boys enrolled in high school and university are not the object of this study, the Cadaster also gives information about them, including the names of the institutions where they studied.

considered the age range between 5 and 10. The ages for working children, male and female, have been established following Borrás Llop (2005), namely ages between 6-14.

Table 4.1 shows that the rates of girls and boys working was quite similar: 17.7% and 16% respectively. However, the rates of schooling were 6.5% higher for boys in comparison with the rate of girls. To analyse the logic in the use of child labour, one should look for explanations in the family economies and the factors that condition them. One of these factors is the occupation of the head of the household. Álvarez and Ramos Palencia (2018) had proven for Castile eighteenth century that male workers with greater skills had higher wages. Table 4.2 illustrates the six most common occupations (including female and male students) of children depending on the occupation of the male head of their household.

In the towns of La Mancha region, there seems to be a correlation between the main occupation of the head of the household and the percentage of boys attending to school. Among the unskilled and semiskilled workers, 12% of those who declared the occupation of their children, stated that their male children were schooled. For skilled head of household this percentage increases up to 19%. Among the intermediate qualified workers 45% declared to have a male child studying and among the professionals 56%. Finally, farmers are between unskilled/semiskilled workers and skilled workers since those who stated to have a male schooled child were 17%. For girl students this correlation is not clear. Also, it is interesting that only intermediate and professional workers declared to have three male students and not only two as in the rest of the categories. According to the occupation also a pattern is observed: only if the fathers were unskilled and semiskilled, they stated to have a child working in an agriculture field (excluding farming).

For female householders (widows), due to the number of observations, only those classified as "skilled" has been taken into account (table 4.3)[61]. Most of the occupations declared by widows were related to textiles. The strategy followed by widows is clear: boys should go to school while girls should work in textiles, probably supplemented by domestic work. This result may suggest that widows had to choose in whom they invested their human capital. Since this region was an area with high demand for textiles, this could have influenced the decision to use female child labour to sustain a family.

To test whether there were a correlation between the occupation of the male head of the family (following the Armstrong classification) and the decision to send their children to work or to school, I run two different linear probability models (LPM), using child labour and schooling as dependent variables. They are described in the following equation.

$$\text{Working Child/Schooled Child}_{ir} = \alpha_i + \beta_1 \text{ Semiskilled}_{ir} + \beta_2 \text{ Skilled}_{ir} + \beta_3 \text{ Intermediate}_{ir} + \beta_4 \text{ Professional}_{ir} + \beta_5 \text{ Farmer}_{ir} + \mu_r + \varepsilon_{ir}$$

$_i$ indicates the respective child of each head of household and $_r$ denotes the town in which the individual was born. The explained variables are *working child* and *schooled child*, respectively, coded as 1 when the child of the family (boy or girl) was working or in school (aged 6-14 if working and 5-10 if schooled), and 0 otherwise. *Semiskilled* is a dummy for those heads of the household categorised as semiskilled. It applies similarly for the dummies *Skilled, Intermediate, Professional* and *Farmer*. The constant refers to the unskilled group, mostly comprised for agriculture day labourers. The model includes region fixed effects (μ_r) that reflect the towns from Figure 4.3 where the family was born.

[61] There are only a maximum of 4 widows in the other categories.

Controlling for time fixed effects is not necessary since all the data is from the Cadaster of Ensenada circa 1750. Finally, the equation allows for a constant term (α) and an error term (ε). Column 1 from table 4.4 analyses the probability of having a male working child and column 2 analyses the probability of having a female working child. Likewise, columns 3 and 4 describe the probabilities for schooled children. The reference group are the sons or daughters of unskilled workers.

Regarding child labour, the results in table 4.4 suggest that families where the male head of the household with an intermediate or professional occupation had a lower probability of having a male working child. This coefficient was larger among the professionals. In other words, a male child with a professional father had 10% less probability of working that the male child of unskilled workers. On the other hand, daughters of semiskilled and skilled workers had a slightly higher probability of working that the daughters of unskilled workers.

Concerning schooling, in all cases, the sons of semiskilled, skilled, intermediate and professional workers and farmers were more likely to attend school than the sons of unskilled workers. This coefficient is larger among the "professionals" (being 23.74% more likely to attend school than the sons of unskilled workers), "intermediates" (21.80%) and farmers (11.84%). For girls, this coefficient appears only significant correlated for the daughters of professional workers and farmers, with a 11.27% and 5.45% more likelihood to attend school respectively.

In the previous analysis it is proven that *exists a correlation between the occupational group of fathers and the decision of sending their children to work or to school.* As family's characteristics shed light on the rationality of their work (and schooling) decisions (Camps 2002), I focus on these characteristics in the next step. First, I performed a regression for working children distinguishing between boys and girls.

$$\text{Working Child}_{ir} = \alpha + \beta_1 \textit{ Log Nº Children}_{ir} + \beta_2 \textit{ Log Birth Order}_{ir} + \beta_3 \textit{ Father numerate}_{ir} + \beta_4 \textit{ Mother numerate}_{ir} + \beta_5 \textit{ Father 50}_{ir} + \beta_6 \textit{ Mother 50}_{ir} + \beta_7 \textit{ Mother\_job}_{ir} + \beta_8 \textit{ Royal Factory}_{ir} + \mu_r + \varepsilon_{ir}$$

$_i$ indicates the respective child of family and $_r$ indicates the town in which the family was born. The variable to be explained is *working child*, coded as 1 when the head of the family responded that they have working children (aged 6-14), and 0 otherwise. *Log Nº Children* is the logged number of children under twenty years of age in the family[62]. *Log Birth Order* is the logged birth order of the children depending on the year that they were born. In this case, the older children in the family had the higher numbers in the database. *Father numerate* is a dummy for the level of human capital of the fathers used to proxy for numeracy, coded as 0 when age is stated as a multiple of five and 1 if not. *Mother numerate* is applied in the same way. *Father 50* and *Mother 50* are dummies coded as 1 when the father or mother reported an age of more than 50, respectively. *Mother_job* is a dummy coded as 1 when the head of the household reported the occupation of the wife. *Royal Factory* is a dummy that refers those towns where there was a textile factory (Guadalajara and Brihuega). The model includes region fixed effects at a municipal level (μ_r), a constant term (α) and an error term (ε). Finally, I included occupational group fixed effects in order to control for the characteristics of the workers across occupations. These occupational groups are the same as those of Table 4.4: unskilled, semiskilled, skilled, intermediate, professional and farmer workers.

Table 4.5 shows the results of the determinants of having a male or a female working child. For the determinants of child labour for boys, birth order, the age of their

[62] I have chosen twenty years of age because there are very few children living with their parents after 20 years in this sample.

parents and the presence of a textile factory were the family lives are positively correlated with male child labour. That means older brothers had more probability of working. As the literature has shown, older children were more likely to work than their younger siblings (Emerson et al. 2008; Orraca 2014). Larger families have a higher likelihood of being impoverished and there was therefore a need for additional income having a first-born son working. It has been argued that older sons could earn higher wages than their younger brothers (Emerson et al. 1998). Also, sons of fathers older than 50 years of age and those who live in a town with a textile factory, had a higher probability of working. However, the significance of the age of the parents might be partly explained by the correlation with the age of the children. Fathers and mothers older than 50 years of age had more often older children who were more likely to work. Apart from the age of the parents, the existence of a textile factory has the third highest coefficient. The determinants that are negatively correlated with male child labour are the number of children and if the mother is numerate.

Among the girls, the number of children (or the overall size of the family), the age of the parents, to have a mother working and the presence of a factory in their town, are correlated with the decision of the families of having a female working child. In the case of the daughters, the highest coefficient that affect positively the decision of sending them to work is if their mother also had an occupation. The explanation is because most of the girls worked in the textile, and the textile was not inherited through the fathers but from the mothers[63]. As Borrás Llop (2002a: 182) has stated: "... Girls, with exceptions, were not oriented to take responsibility for family farming".

[63] 62.2% of the women were engaged in the secondary sector (mostly textile industry) while fathers were engaged in a similar percentage in the primary sector (Sarasúa 2019:495).

Furthermore, girls mostly worked in the secondary sector, with a rate of 84% compared to boys with only 12% (Sarasúa 2013). In Almagro, for example, lace making employed hundreds of women, and girls learned how to produce lace from a very young age, even before the age of seven (Sarasúa 2002b). On the other hand, to have a father and a mother numerate is correlated negatively with having a daughter working.

In the next step, the determinants that play a role in the families on the decision on sending their children to school are tested, also separately for boys and girls. In this case, the regression is as follows.

$$\text{Schooled Child}_{ir} = \alpha + \beta_1 \textit{ Log NºChildren }_{ir} + \beta_2 \textit{ Log Birth Order }_{ir} + \beta_3 \textit{ Father numerate }_{ir} + \beta_4 \textit{ Mother numerate }_{ir} + \beta_5 \textit{ Teacher's wage }_{ir} + \beta_6 \textit{ Royal Factory }_{ir} + \beta_7 \textit{ Town 1000 }_{ir} + \beta_8 \textit{ Log Teacher/Boys }_{ir} + \beta_9 \textit{ Log Teacher/Girls }_{ir} + \mu_r + \varepsilon_{ir}$$

$_i$ indicates the respective child of family and $_r$ indicates the town in which the family was born. The variable to be explained is *schooled child*, coded as 1 when the head of the family responded that they have schooled children aged 5-10, and 0 otherwise. As in the previous regression, *Log NºChildren* , *Log Birth Order*, *Father numerate*, *Mother numerate* and *Royal Factory* determinants are included. *Teacher's wage* is the salary of teachers of the first letter male schools and *Town 1000* is a dummy for those towns with less of 1,000 inhabitants. *Log Teacher/Boys* is the ratio of male teachers to boys of school-going age. Likewise, *Log Teacher/Girls* is the same for female teachers and girls. The model includes region fixed effects at a municipal level (μ_r), a constant term (α) and an error term (ε). Finally, I included the occupational group fixed effects as in Table 4.5.

Table 4.6 shows the results of the determinants of having a male or a female school-going child. Starting with boys (column 1) the determinants that play a positive role in their schooling decision by their family are if their father was numerate and, with

a highest coefficient, the ratio of male teachers to boys of school-going age. As Reis (2005) claimed, although demand side factors contributed to the spread of literacy, supply side factors (as the location of schools) also had importance. In the case of Castile La-Mancha in the eighteenth century, a teacher was equal to a school. On the contrary, even controlling for the family size (the number of children under twenty years of age), birth order is negatively correlated with having a schooled son. It is important to control for family size in order to know the exact impact of birth order. Family size has been negatively correlated with developmental effects "due to the fact that family resources are spread more thinly the larger the family is" (Emerson et al. 2008: 1648). Once again, our result is in line with the literature (at least for boys): younger siblings had more opportunities to attend to school (Emerson et al. 2008; Orraca 2014). The other determinant negatively correlated with the decision of the family of schooled their son, although not in a high coefficient but significant, is the wage of the teachers, or in other words, the cost of education.

The wages of the male teachers depended largely on the municipality and family's financial resources or whether the teacher was graduate. These salaries used to be very low therefore teachers used to have other occupations (Sarasúa 2002b). In our sample only 3 teachers of 25 stated to be qualified. Pablo Sánchez Barburdo of Ajofrín declared to be "examined teacher of the art of first letters, it is my exercise to teach this art" and Juan Francisco Vega Maldonado of Almagro pointed out that he had a certificate of examination since 1716. As well, Félix García Rico from El Carpio states to be examined. According to a complementary occupation, Juan Castillo from Alanchete reported to be also a farmer; Manuel Martínez Puga from Alcaraz to be scribe; Alonso Brihuega from Brihuega, Juan Moracho Sanz from Guadalajara and José López Román from Torre de Juan Abad were also sacristans; José Ramos Anay from Valenzuela was cartwright apart

from teacher and Francisco Ayuso Rico declared to sing the memorials of those that did not know to write. Table 4.8 in the appendix shows teachers and their wages by town[64].

For girls, the determinant that plays a positive role on their schooling is if their father was numerate, and as in the case of the boys, the numbers of female teachers in the town according to girls of school-going age. In this case, we do not have more information about the teachers since in the *Respuestas Generales* these data are not given. However, the size of the family, the fact that there was a royal factory in the town, and if the town had less than 1,000 inhabitants, contributed negatively to the decision of their schooling by their families. It has been established that a decrease in the average size of families indicates that quality replaces the number of descendants. In other words, parents would decide to have fewer children and invest more resources in their health and education (Núñez 1992). It seems that those families with a larger number of children refused the education of their daughters instead of the education of their sons. In this case, if the town had less than 1,000 inhabitants was the largest coefficient with a negative correlation of 19%. This could be related to the fact that, except Guadalajara, the textile industry of New Castile had its basis on peasant domestic units that used to combine these works with those of the countryside or tertiary sector (López Barahona 2020).

Focusing on the human capital of the fathers, measured through numeracy proxy, it has been found correlated positively with having sons and daughters schooled, fact analysed below in a more detail [65]. The level of numeracy of the head of household can

[64] In the *respuesta general* number 32 of Puebla del Príncipe, it is reported that the wage of the teacher is 0 because he did not have any disciple (student). However, it has been counted from the memorials that in Puebla del Príncipe there were 9 boys attending school.

[65] In terms of child labour has been found that numeracy level of the father is negatively correlated with the decision of sending their daughter to work and numeracy level of mother is likely

be used as an indicator for human capital and can be estimated by occupation groups, offering insights into the decisions that families made for their children. The age-heaping technique used to measure numeracy levels is based on the accuracy of self-reported age data in historical documents (death and marriage registers or censuses, for example). In this way, it has been established that individuals who did not know their exact age tended to respond in numbers ending in 0 or 5, allowing researchers to use this method as a proxy for basic mathematical and numerical skills (A'Hearn et al. 2009).

Table 4.7 demonstrates that there is a relationship between the ABCC levels, as a proxy for numeracy (basic mathematical skills), and the percentage of son students grouped by the occupation of the heads of their households: a higher ABCC level implies higher percentages of male children studying.

From table 4.7 we could interpret that indeed parents with greater human capital may have invested more in their children's education, specifically in the education of their sons. There is also evidence that these workers with greater skills were better paid and had the ability to diversify their earnings through secondary employment (Álvarez and Palencia 2018). These privileged groups could afford to pay a teacher, when this was not a service provided by the municipality, as well as forego income from child labour. According to farmers, this result reinforces the previous idea about the ease at which farmers could combine the work of their sons with schooling, despite being absent from school at times that were critical for agriculture. For girls, this pattern is not followed since the daughters of all occupational groups were schooled by 6-7% percent.

correlated with this decision for both, boys and girls. However, we do not know if there were instead in the school.

4.5 Conclusions

This research has aimed at examining the effects that family circumstances had on the child-labour and schooling decisions for children in eighteenth-century Spain, being the first study that performed this analysis on a family level in pre-industrial economies. Using the Memorials of the Cadaster of Ensenada it has been possible to reconstruct child labour and schooling statistics at a family level, a topic that often suffers from data unrecording. The findings from this study make several contributions to the current literature of child labour, schooling, human capital and family's economies (Humphries 2003; Hernández 2013; Sarasúa 2013; Beltrán et al. 2019; Martini and Borderías 2020).

In the case of New Castile, 17.7% of girls and 16% of the boys aged 6-14 had a remunerated job. Moreover, although the decision of household 'heads was logic, we need to add here that from the national perspective this clearly hindered human capital formation and subsequent economic growth.

Regarding child labour, it has been proven that *for boys, the size of the family and the human capital of the mother was negatively correlated with the fact that they were working*. On the contrary, *their birth order, the age of the parents and the existence of a textile factory in the town that they lived, were correlated positively*.

For girls, the human capital of their parents was negatively correlated with their likelihood of working. However, if their mother had a job and they stayed in a town with a textile factory, the were more likely to work. Moreover, older parents and larger family size were positively related with the decision of the families on sending their daughters to work. These two factors changed with the demographic transition in the twentieth

century and were responsible for ending child labour in developed countries, along with legal amendments and technological progress (Camps 2002).

When analysing the determinants of schooling I have been able to prove that *the older boys had less likelihood to attend school than their younger* brothers. As well, the wages of the teacher of the town are correlated negatively since it was the family and the municipality who had to pay these wages. However, even controlling for the occupational structure of the male head of households, I found *a significant effect of numeracy on the decision of schooling their sons*. Similarly, the supply of teachers for boys played a positive role on this decision.

For girls, it *was the family size rather than their birth position that was negative related with their schooling*. Living in a town with a textile factory or in a town of less than 1,000 inhabitants, also affected them negatively in this regard. On the opposite, and *as we just saw for boys, the human capital of their fathers and the supply of teachers for girls positively affected their parents' decision to send them to school.*

In a double check analysis, I demonstrate that there was *a relationship between numeracy levels by the occupations of the head of the family and the decision on sending their sons to school*. Those who belonged to the occupational group with the lowest numeracy level (unskilled workers, mostly day labourers) appear to be less likely to invest in the education of their children. Conversely, the occupational groups with a higher human capital level (intermediate and professional workers) implied a higher percentage of school-going children. As has been shown by the literature, here I prove at a family level that farmers, due to their peculiarity of owning a land, were more interested in investing education than day labourers (Tollnek and Baten 2017; Beltrán Tapia and Martínez Galarraga 2018; Pérez Artés and Baten forthcoming).

Finally, I have been able to conclude that *boys and girls were treated differently as there were different factors affecting their schooling and work decision by their families, such as family size, birth order or the fact that their mother was working*. However, *there were also factors that had an influence on both, boys and girls, for instance the age and the human capital level of their parents, the supply of teachers or the existence of a textile factory in their town*.

4.6 References

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

Álvarez, B., & Ramos Palencia, F. 2018. Human capital and earnings in eighteenth-century Castile. *Explorations in Economic History*, 67: 105–133.

Ariès, P. 1962. *Centuries of Childhood: A Social History of Family Life*. New York: Vintage.

Armstrong, A. 1972. The use of information about occupation, in E. A. Wrigley, ed., *Nineteenth-century society: essays in the use of quantitative methods for the study of social data*: 191-310. Cambridge.

Becker, G. 1964 [1994]. *Human capital revisited. In Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. The University of Chicago Press (Third edition).

Beltrán Tapia, F. J. & Martinez-Galarraga, J. 2018. Inequality and education in pre-industrial economies: Evidence from Spain. *Explorations in Economic History,* 69: 81–101.

Beltrán-Tapia F.J.; Díez-Minguela, A.; Martínez-Galarraga J. & Tirado-Fabregat, D. 2019. Capital humano y desigualdad territorial. El proceso de alfabetización en los municipios españoles desde la ley Moyano hasta la guerra civil. *Estudios de Historia Económica*, 74, Banco de España.

Borderías, C. 2013. Salarios infantiles y presupuestos familiares en la Cataluña Obrera, 1856-1920. In J. M. Borrás Llop (Ed.), *El trabajo infantil en España, 1700-1950*: 371–408. Barcelona: Icaria.

Borrás Llop, J. M. 2002a. Aprender trabajando. La actividad de niñas y niños en tierras de regadío (la Vega del Tajuña a comienzos del siglo XX). In C. Sarasúa & L. Gálvez (Eds.), *¿Privilegios o eficiencia? Mujeres y hombres en los mercados de trabajo*: 157–183. Alicante: Universidad de Alicante.

Borrás Llop, J. M. 2002b. El trabajo infantil en el mundo rural español, 1849-1936. Género, edades y ocupaciones. In J. M. Martínez Carrión (Ed.), *El nivel de vida en la España rural, siglos XVIII-XX* : 497–547. Alicante: Universidad de Alicante.

Borrás Llop, J. M. 2002c. Mercado laboral, escolarizacion y empleo infantil en una comarca agricola e industrial (el Valles Occidental, 1881-1910). *Cuadernos de Historia Contemporánea / Departamento de Historia Contemporánea*.

Borrás Llop, J. M. 2005. Schooling and child farm labour in Spain, circa 1880-1930. *Continuity and Change*, 20(3): 385-406.

Borrás Llop, J. M. 2019. Introducción. In María Jesús Fraga (Ed.). *Elena Fortún. Lo que cuentan los niños. Entrevistas a niños trabajadores (1930-1931)*: 25-103. Biblioteca Elena Fortún: Renacimineto.

Camarero Bullón, C. 2002. El Catastro de Ensenada, 1749-1759: diez años de intenso trabajo y 80.000 volúmenes manuscritos. *CT: Catastro*, 46: 61–68.

Campos Luque, C. 2014. La tasa de actividad femenina a mediados del siglo XIX en Andalucía: el caso de Antequera. *Investigaciones de Historia Económica - Economic History Research*, 10(3): 191–201.

Camps, E. 1995. *La formación del mercado de trabajo industrial en la Cataluña del siglo XIX*. Ministerio de Trabajo y Seguridad Social.

Camps, E. 2002. Trabajo infantil y estrategias familiares durante los primeros estadios de la industrialización catalana (1850-1925). Esbozos a partir del estudio de un caso. *Cuadernos de Historia Contemporánea*: 263–279.

Chiswick, B. R., Lee, Y. L., & Miller, P. W. 2003. Schooling, literacy, numeracy and labour market success. *Economic Record*, 79:165-181.

Crayen, D., & Baten, J. 2010. New evidence and new methods to measure human capital inequality before and during the industrial revolution: France and the US in the seventeenth to nineteenth centuries. *The Economic History Review*, 63(2): 452–478.

Cunningham, H. 2014. *Children and Childhood in Western Society Since 1500*. Hoboken: Taylor and Francis (1995).

Emerson, Pat & Souza, A. 2008. Birth order, child labor, and school attendance in Brazil, *World Development*, 36 (9): 1647-1664.

Hernández, R. 2013. La mano de obra infantil en la Castilla rural del siglo XVIII. El trabajo del niño es poco, pero el que lo desprecia un loco. In J. M. Borrás Llop (Ed.), *El trabajo infantil en España, 1700-1950:* 91-115. Barcelona: Icaria.

Humphries, J. 2003. Child Labor: Lessons from the Historical Experience of Today's Industrial Economies. *The World Bank Economic Review.*

Humphries, J. 2013. Childhood and child labour in the British industrial revolution. *The Economic History Review,* 66(2): 395–418.

Laspalas Pérez, F. 1991. La escolarización elemental en España según el censo de Godoy (1797). *Historia de La Educación: Revista Interuniversitaria*.

López Barahona, V. 2020. El trabajo de las mujeres en la real fábrica de Guadalajara durante el siglo XVIII. *Historia Social*, 96: 97-112.

Mallorquí-Ruscalleda, N. 2019. School Act and Elementary Education in Nineteenth-Century Spain. In J. Westberg, L. Boser, & I. Brühwiler (Eds.), *School Acts and the Rise of Mass Schooling. Education Policy in the Long Nineteenth Century*: 171–193. Palgrave Macmillan.

Martini, M. & Borderías, C. 2020. Coping with crisis: labour markets, institutional changes and household economies. An introduction. *Continuity and change*, 35: 1-9.

Muñoz Abeledo, L. 2012. Women in the Rural and Industrial Labor Force in Nineteenth-Century Spain. *Feminist Economics*, 18(4): 121–144.

Núñez, C. E. 1992. *La fuente de la riqueza: educación y desarrollo económico en la España contemporánea*. Alianza.

Orraca, P. 2014. El trabajo infantil en México y sus causas. *Revista Problemas del Desarrollo*, 178(45): 113-137.

Pérez Artés, M.C. & Baten, J. Forthcoming. Land inequality and numeracy in Spain during the seventeenth and eighteenth century. *Historia Agraria. Revista de Agricultura e Historia Rural.*

Reis, J. 2005. Economic growth, human capital formation and consumption in Western Europe before 1800, in R. C. Allen, T. Bengtsson y M. Dribe (Eds.), *Living standards in the past: new perspectives on well-being in Asia nd Europe*: 195-227. Oxford: Oxford University Press.

Sarasúa, C. 2002a. Aprendiendo a ser mujeres: las escuelas de niñas en la España del siglo XIX. *Cuadernos de Historia Contemporánea / Departamento de Historia Contemporánea*.

Sarasúa, C. 2002b. El acceso de niñas y niños a los recursos educativos en la España rural del siglo XIX. In J. M. Martínez Carrión (Ed.), *El nivel de vida en la España rural, siglos XVIII-XX* : 549–609. Alicante: Universidad de Alicante.

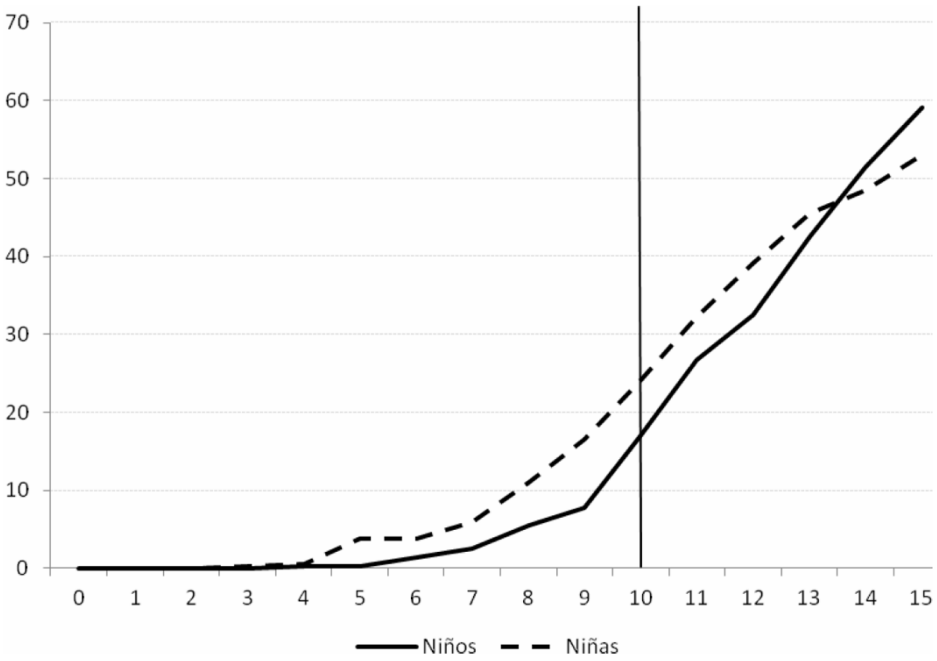
Sarasúa, C. 2013. ¿Activos desde cuándo? La edad de acceso al mercado de trabajo en la España del siglo XVIII. In J. M. Borrás Llop (Ed.), *El trabajo infantil en España, 1700-1950*: 61–80. Barcelona: Icaria.

Sarasúa, C. 2019. Women's work and structural change: occupational structure in eighteenth-century Spain. *Economic History Review,* 72: 481-509.

Tollnek, F., & Baten, J. 2017. Farmers at the heart of the 'human capital revolution'? Decomposing the numeracy increase in early modern Europe. *The Economic History Review,* 70(3): 779–809.

4.7 Figures and tables

Figure 4.1 Activity rate of boys and girls under 15 years in New Castile (1753)

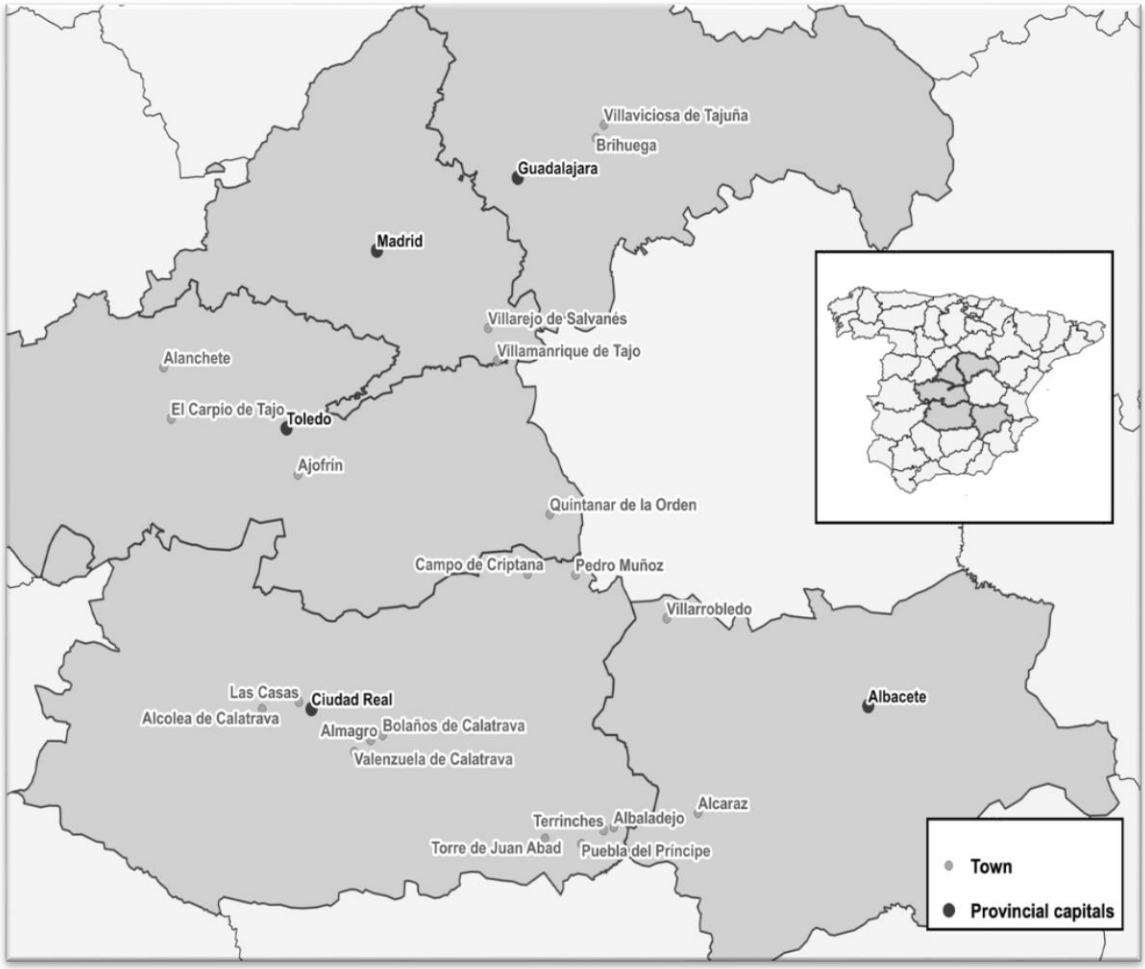


Source: Sarasúa, 2013, 72

Figure 4.2 Girl's School (Maestra de niñas) circa 1750

Source: Domingo Martínez (1688-1749), Carro de la Tierra (hacia 1748). Real Fábrica de Tabacos, Sevilla, Donación de Estado (1896).

Figure 4.3 Figure 2 Towns included



Source: see footnote 54

Table 4.1 rates of schooled and working children

Girls	Working (aged 6-14)	Schooled (aged 5-10)	% Working	% Schooled
Yes	642	159	17.7	5.5
No	2,976	2,711	82.3	94.5
Total	3,618	2,870	100	100

Boys	Working (aged 6-14)	Schooled (aged 5-10)	% Working	% Schooled
Yes	628	367	16.0	12.0
No	3,296	2,698	84.0	88.0
Total	3,924	3,065	100	100

Source: see footnote 54

Table 4.2 Six most frequent responses for child occupation (boys and girls) by the male head of the household depending on his occupation (Armstrong category)

Occupation Child 3	*Occupation Child 2*	*Occupation Child 1*	*Nº of Families*	*% of Families*	*Nº Total Families*
		Unskilled			
	Textiles/Wool manufacture	Textiles/Wool manufacture	61	13	454
	Students/male	Students/male	56	12	454
	Agriculture/Animal husbandry	Agriculture/Animal husbandry	39	9	454
	Agriculture/Agriculture Labourer	Agriculture/Agriculture Labourer	35	8	454
	Textiles/Lace manufacture	Textiles/Lace manufacture	32	7	454
	Students/female	Students/female	26	6	454
		Semiskilled			
	Students/male	Students/male	13	12	110
	Textiles/Wool manufacture	Textiles/Wool manufacture	12	11	110
	Industries producing products from fibres	Industries producing products from fibres	8	7	110
	Students/female	Students/female	6	5	110
	Clothing/Clothing manufacture	Clothing/Clothing manufacture	5	5	110
	Agriculture/Agriculture Labourer	Agriculture/Agriculture Labourer	3	3	110
		Skilled			
	Textiles/Wool manufacture	Textiles/Wool manufacture	51	20	254
	Students/male	Students/male	48	19	254
	Students/female	Students/female	14	6	254

	Textiles/Lace manufacture	Textiles/Lace manufacture	14	6	254
	Textiles/Wool manufacture	Textiles/Wool manufacture	11	4	254
	Footwear	Footwear	5	2	254
		Intermediate			
	Students/male	Students/male	11	41	27
	Textiles/Wool manufacture	Textiles/Wool manufacture	8	30	27
	Students/male	Textiles/Lace manufacture	2	7	27
	Textiles/Lace manufacture	Textiles/Lace manufacture	2	7	27
Students/male	Students/male	Students/male	1	4	27
		Sellers of food	1	4	27
		Professional			
	Students/male	Students/male	25	42	60
	Students/female	Students/female	5	8	60
	Students/male	Students/female	4	7	60
Students/male	Students/male	Students/male	4	7	60
		Students/female	2	3	60
	Textiles/Wool manufacture	Textiles/Wool manufacture	2	3	60
		Farmer			
	Students/male	Students/male	45	17	260
	Agriculture/Farming	Agriculture/Farming	31	12	260
	Students/female	Students/female	18	7	260
	Textiles/Wool manufacture	Textiles/Wool manufacture	18	7	260
	Textiles/Lace manufacture	Textiles/Lace manufacture	12	5	260
	Students/male	Students/female	10	4	260

Source: see footnote 54

Table 4.3 shows the six most common occupations of children if the head of household was a women (widow).

Occupation Child 1	*Occupation Child 2*	*Nº of Families*	*% of Families*	*Nº Total Families*
Skilled				
Textiles/Wool manufacture	Textiles/Wool manufacture	28	33	85
Students/male	Students/male	12	14	85
Textiles/Linen manufacture	Textiles/Linen manufacture	4	5	85
Agriculture/Agricultural labourer	Agriculture/Agricultural labourer	4	5	85
Textiles/Lace manufacture	Textiles/Lace manufacture	3	4	85
Clothing/Hats, gloves, stockings	Clothing/Hats, gloves, stockings	2	2	85

Source: see footnote 54

Table 4.4 Probability of having a working child by occupation of the head of household – Linear Probability Model (LPM)

	(1)	(2)	(3)	(4)
	Male working child	Female working child	Male schooled child	Female schooled child
Semiskilled	-1.00	4.40*	4.70**	0.02
	(0.644)	(0.093)	(0.044)	(0.991)
Skilled	1.11	2.93*	9.64***	1.40
	(0.508)	(0.090)	(0.000)	(0.197)
Intermediate	-7.80**	1.18	21.80***	-0.27
	(0.018)	(0.792)	(0.000)	(0.701)
Professional	-10.09***	-3.74	23.74***	11.27***
	(0.000)	(0.174)	(0.000)	(0.000)
Farmer	0.40	-0.53	11.84***	5.45***
	(0.818)	(0.777)	(0.000)	(0.000)
Constant	8.92***	4.09**	-5.34***	-0.15
	(0.000)	(0.025)	(0.000)	(0.876)
Observations	3,479	3,176	2,769	2,571
Adjusted R-squared	0.06	0.09	0.11	0.11
Region FE	YES	YES	YES	YES

Robust pval in parentheses *** p<0.01, ** p<0.05, * p<0.1
The dependent variable is 1 if there was a working child or child in school, 0 otherwise. The constant refers to sons or daughters aged 6-14 if working and 5-10 if schooled of unskilled fathers. Region fixed effects refer to the towns included in Figure 2.

Table 4.5 Determinants of having a child working by family -Linear Probability Model (LPM)

	(1)	(2)
	Male working child	Female working child
Log NºChildren	-3.00**	2.79*
	(0.045)	(0.094)
Log BirthOrder	9.55***	1.62
	(0.000)	(0.335)
Father Numerate	0.20	-4.13***
	(0.872)	(0.002)
Mother Numerate	-2.79**	-3.64***
	(0.024)	(0.006)
Father50	11.09***	7.74***
	(0.000)	(0.000)
Mother50	12.32***	10.27***
	(0.000)	(0.004)
Mother_job	2.28	14.81***
	(0.255)	(0.000)
Royal Factory	10.31***	13.61***
	(0.001)	(0.000)
Constant	-2.61	3.04
	(0.471)	(0.492)
Observations	3,479	3,176
Adjusted R-squared	0.08	0.12
Armstrong Groups FE	YES	YES
Region FE	YES	YES

The dependent variable is 1 if there was a working child, 0 otherwise. Column 1 included sons and column 2 daughters. The constant refers to sons or daughters aged 6-14. Region fixed effects are the provinces included in Figure 1. Occupational group FEs correspond to the occupational sector in which the head of the family worked (unskilled, semiskilled, skilled, intermediate, professional and farmer). Robust p-Values are given in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.6 Determinants of having a schooled child by family -Linear Probability Model (LPM)

	(1)	(2)
	Male schooled child	Female schooled child
Log NºChildren	2.13	-2.46*
	(0.180)	(0.064)
Log BirthOrder	-4.08**	0.16
	(0.011)	(0.890)
Father Numerate	2.43*	3.46***
	(0.059)	(0.000)
Mother Numerate	1.59	-0.69
	(0.218)	(0.514)
Teacher's Wage	-0.03***	
	(0.000)	
Royal Factory	-2.44	-12.20***
	(0.301)	(0.000)
Town 1000	-0.91	-19.31***
	(0.764)	(0.009)
Log Teachers/Boys	6.62***	
	(0.000)	
Log Teacher/Girls		5.76***
		(0.009)
Constant	60.46***	48.70***
	(0.000)	(0.000)
Observations	2,584	1,422

Adjusted R-squared	0.09	0.09
Armstrong Groups FE	YES	YES
Region FE	YES	YES

The dependent variable is 1 if there was a child at school, 0 otherwise. Column 1 included sons and column 2 daughters. The constant refers to sons or daughters aged 5-10. Region fixed effects are the provinces included in Figure 1. Occupational group FEs correspond to the occupational sector in which the head of the family worked (unskilled, semiskilled, skilled, intermediate, professional and farmer). Robust p-Values are given in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 4.7 ABCC index, percentage of boys and girls studying by the occupation of the male head of the household

	Male HH		
Head of the household	ABCC levels	% son students	% daughter students
Unskilled	70.4	12	6
Semiskilled	71.7	12	6
Farmer	73.2	17	7
Skilled	78.9	19	6
Intermediate and Professional	85.7	47	6

Source: see footnote 54

4.8 Appendix

4.8.1 Teachers and wages by town

Table 4.8 Teachers of elementary school and wages by town

Name and Surnames of Teachers			Town	Wages (Reales)
Pablo	Sánchez	Barbudo	Ajofrín	1,100
Juan	Castillo		Alanchete	108
Gregorio	Sánchez	Olmo	Albaladejo	300
Manuel	Martínez	Puga	Alcaraz	432
Juan	Funes		Almagro	660
Juan Francisco	Vega	Maldonado	Almagro	660
Juan	García	Segura	Almagro	660
Isidro	Estúñiga		Brihuega	200
Alonso	Brihuega		Brihuega	200
Felipe	Sánchez		Brihuega	1,500
Alfonso	Sánchez	Berenguillo	Campo de Criptana	600
Francisco Antonio	Sánchez	Alarcos	Campo de Criptana	600
Félix	García	Rico	El Carpio	1,400
Juan	Moracho	Sanz	Guadalajara	1,100
Sebastián	Bravo	Delgado	Guadalajara	905
Juan de la Cruz	Alcoholado		Pedro Muñoz	800
José Tomás	Pisa		Puebla del Príncipe	0
Manuel	Díaz	Romeral	Quintanar	800
Juan	Hurtado	Mendoza	Terrinches	500
José	López	Román	Torre de Juan Abad	1,300
José	Ramos	Anay	Valenzuela	350
Francisco	Ayuso	Rico	Villarejo de Salvanés	660
Diego	Rentero		Villarobledo	875
Francisco José	González		Villarobledo	875
Julián	Sáiz	de la Morena	Villarobledo	875

5. Summary and Outlook

Literature on human capital has been broadly studied over the past century (Barro and Lee 1996; Becker 1962; Cipolla 1969; Mincer 1974; Núñez 1992; Romer 1989; Schultz 1961). However, empirical evidence prior to nineteenth century is still scarce. Through numeracy and the ABCC indicator, it has been possible to conduct research from sixteenth century to the eighteenth century in Spain. A number of conclusions can be drawn from the results of this thesis.

First, we found that land equality (as opposed to regions with *latifundistas* and many day labourers) played a role on the formation of human capital among pre-industrial societies: the higher share of farmers had a positive effect on regional numeracy. We argue that farmers were more willing to invest in the skills of their children since they would need them to run their farms in future, they had relatively high levels of nutrition and were not as burdened with child labour as day labourers were; as shown in chapter two (Baten et al. 2014; Tollnek and Baten 2017). These findings are in line with the literature of nineteenth Spain century. Beltrán Tapia and Martínez Galarraga (2018) demonstrated that there was a negative relationship between the fraction of farmers and male literacy rates un mid-nineteenth century.

Moving to the self-selection of migrants, chapter three demonstrates that those who migrated from Spain to colonial Hispanic America were positively selected in terms of human capital, using the age-heaping method. Although this has been established for the nineteenth and twentieth centuries, this is the first estimation of migrant human-capital selectivity for pre-independent Hispanic America (Sánchez Alonso 2007). Among migrants, those who were more numerate, on average, went to Mexico instead of Peru. The role of educational institutions, the viceroyalty structure and the presence of religious

orders could have influenced the decision of people with greater human capital to emigrate to Mexico. According to the numeracy inequality among different ethnic groups I found that levels of numeracy in Mexico between people of Spanish and native origins indicated a relative high inequality during the late seventeenth century, reducing the gap circa 1710. This result is consistent with the decrease differences in heights between ethnic groups in Mexico since 1730 (Dobado González and García Montero 2010). In an international comparison with the Cape colony (the only research focused on the numeracy differences between settlers and colonised) is proved that Hispanic America was less unequal in terms of human capital (using the abcc index as a proxy) (Baten and Fourie 2015).

Chapter four tests which determinants influenced the family decisions regarding the child-labour and schooling of their children in eighteenth-century Spain. For boys, the size of the family and the human capital of the mother played a negative role on the fact that they were working while their birth order, the age of the parents and the existence of a textile factory in the town that they lived played a positive role. On the other hand, the older boys were less likely to attend school than their younger brothers while the numeracy levels of their fathers or the supply of teachers for boys contributed positively to attend school. According to girls, the human capital of their parents was negatively correlated with their likelihood of working. On the opposite, if their mother had a job, they stayed in a town with a textile factory, they had older parents and larger family size, they were more likely to work. According to their schooling, the size of the family or the fact of living in a town with a textile factory were the factors that hampered their schooling. As for boys, the human capital of their fathers and the supply of teachers for girls played a positive role.

5.1 References

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

Baten, J., & Fourie, J. 2015. Numeracy of Africans, Asians, and Europeans during the early modern period: new evidence from Cape Colony court registers. *The Economic History Review*, *68*(2): 632–656.

Baten, J.; Crayen, D., & Voth, HJ. 2014. Numeracy and the Impact of High Food Prices in Industrializing Britain, 1780–1850. *Review of Economics and Statistics*, 96 (3): 418–430.

Becker, G. 1962. Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5, Part 2): 9 – 49.

Beltrán Tapia, F. J. & Martinez-Galarraga, J. (2018). Inequality and education in pre-industrial economies: Evidence from Spain. *Explorations in Economic History*, 69: 81-101.

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

Dobado González, R., & García 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.

Mincer, J. 1974. *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research.

Núñez, C. E. 1992. *La fuente de la riqueza: educación y desarrollo económico en la España contemporánea*. Alianza.

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

Sánchez Alonso, B. 2007. The Other Europeans: Immigration into Latin America and the International Labour Market (1870–1930). *Revista De Historia Económica / Journal of Iberian and Latin American Economic History*, 25(03): 395–426.

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

Tollnek, F. & Baten, J. (2017). Farmers at the heart of the 'human capital revolution'? Decomposing the numeracy increase in early modern Europe. *The Economic History Review,* 70 (3), 779–809.