Land inequality and numeracy in Spain during the seventeenth and eighteenth century

By Mari Carmen Pérez-Artés* and Jörg Baten**
University of Tübingen, mcperezartes@gmail.com

University of Tübingen, CESifo and CEPR. Address: Melanchthonstr. 30, 72074 Tübingen,

joerg.baten@uni-tuebingen.de

This is the last working paper version before the study was published by "Historia Agraria", online 2021, https://doi.org/10.26882/histagrar.083e08p

Abstract

We assess the relationship between land inequality and human capital for the later part of the early modern period, focusing on individual-level evidence from Spain. Our main finding is that land inequality already had a significant negative effect on human capital formation in the late seventeenth and eighteenth century. We argue that this reflects the potential role of farmer-family-based social structure (as opposed to the social structure dominated by *latifundia* and day laborers) in the development of numeracy, consistent with earlier studies which argued that farmer households could (1) maintain a relatively favourable nutritional standard – a precondition for cognitive skills, (2) limit child labour and (3) were able to provide basic numerical skills in family farms. Our result is robust to including various controls and potential confounding variables.

Keywords: Land inequality, Numeracy, Early Modern, Spain, Andalusia, Regional development

JEL: N33, N93, O18

1. Introduction

Recently, an agricultural dimension was added to the debate about the determinants of growth and obstacles to development from a long-term perspective. Galor et al. (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 plays 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 (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 Tapia 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 (and we can include some limited evidence on the sixteenth century as well), 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[1]. What was the causal mechanism instead? Building on earlier studies, we argue that farmer families provided a relatively favourable nutritional standard, so that their descendants could acquire human capital (Tollnek & Baten, 2017; Baten et al., 2014). Moreover, farmer parents were able to provide some basic skills at home. This was very relevant for the 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 farmer 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 farmer 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

[1]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).

agricultural labourer parents might often have been lower (Beltrán Tapia and Martinez-Galarraga, 2018)[2].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 (Kagan, 1981). Only a few villages received school donations from a 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 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[3]. This region is the southernmost point of mainland Europe and, with more than 87,000 square kilometers, its area is larger than several European

[2] 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).

[3] 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).

countries (Parejo, 2009, 11).[4] 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).

A new dataset from *padrones* (local nominative population censuses) and Cadastre of Ensenada is analysed here[5]. Table 1 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[6].

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

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

[5] 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).

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

where sources of other education indicators were incomplete. Numeracy, or the 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 introduces the historical context of land inequality and human capital in modern Spain. Section 3 follows with the explanation of the methodology and the data used in this study. In section 4, our empirical results and descriptive analysis are presented. Section 5 adds potential caveats and section 6 presents the conclusions.

2. Land inequality and human capital in modern Spain.

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) [7]. However, Carmona and Simpson (2003: 19) argued that these institutions

[7] *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).

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 farmer 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 Muslim uprising of 1568-70[8]. After this event, the state distributed additional plots to Christian settlers and also prevented the accumulation of *latifundia*[9].

Warlords, nobles and clergy were the main beneficiaries of the Castilian conquest of Al Andalus. From the thirteenth century onwards, the concentration of landownership

[8] This uprising had its precedent in January of 1567, when a royal law obliged all "Moorish" Christians (Muslims forcibly converted to Christianity) to become "real" Christians within a year. This 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 "Moors" from Spain took place (Andújar Castillo, 2004).

[9] 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).

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 and Philip II, 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 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) [10]. 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

[10] 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.

consisted of people coming from northern Spain, whereas the one of the Granada region was administered by the western Andalusians[11]. 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 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 farmer 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[12]. In central and northern Spain, the farmer share was much higher: Navarra and Guadalajara had the highest shares – 84% and 64%. However, the farmer share is not just a correlate of the north-south differences. For example, La Rioja had a relatively low farmer 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 1, provinces are compatible to today's provinces).

2.2 Human Capital in Spain since the sixteenth 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

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

[12] 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).

only direct evidence for measuring education levels. Only in 1797, with the census of Godoy, data about the schooling process became available. Finally, in 1860, a Spanish census included information about the ability of inhabitants to read 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 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 & Bennassar (1978) studied the interior Andalusian regions of Andújar, Iznatoraf, Úbeda and Cordoba using the testimonies of the accused during the inquisition. Vincent (1987) used fiscal sources and assessed the literacy of the Moorish 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, family was the main agent providing education. 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 & 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 and nineteenth centuries, 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 claimed: "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 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[13].

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 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).

[13] 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 affected agricultural production (Simpson and Carmona, 2017).

3. Methodology and data

The regions considered in this research are illustrated in Figure 1. Table 1 specifies the number of observations by province and period. Table A1 in the appendix contains a description of the sources[14].

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 & 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[15]. Although hortelanos were obviously not farmers, we included them for simplicity in the variable "farmer share" (justified by their small number). In order to assess the plausibility of the farmer shares based on our sample, we can calculate a similar farmer share for the Floridablanca census (even if the Floridablanca census was recorded somewhat later, in 1785-87). The correlation is very strong (Figure 2, aggregated on province level). A large share of both our-sample-based farmer shares and the Floridablanca-based farmer shares are in the 20 to 40 percent range. Our sample is slightly more urban (hence a lower farmer share for Sevilla, for example) and more Andalusian. This difference is mostly removed by our weighting procedure.

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

[15] 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 & Martín Chicano, 2007).

In order to assess numeracy, we employ the "age heaping" methodology using the ABCC index[16]. 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:

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

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

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

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

(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 & Baten, 2010).[17]

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 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 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 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 Cadastre; that might have reflected a special sub-population (such as the nuns in a monastery or the merchant quarter of a city, for

[17] 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.

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 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)[18]. 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 A2 in the appendix)[19]. 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 occupation 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.

4. Descriptive analysis and regression results

Table 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

[18] 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.

[19] 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.

5. The mean farmer 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 farmer 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 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)[20].

To measure the effect of farmer shares on numeracy, based on the occupational information for 17,145 cases, we calculated the farmer share of each local community and period. Our inequality data provide 117 observations combining local communities and birth centuries. We then assigned this farmer share in a given local community and century to all

[20] See note Table 3.

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 3 shows the results of the effect of farmer 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[21]. Interestingly, if we include the inequality proxy "farmer share" the farmer coefficient by itself does not show a significant difference, relative to persons who are not farmers[22]. Consequently, we conclude that the social structure in regions with high farmer 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.

[21] Following Reher (1994), 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.

[22] 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 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.

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 "farmer share" varies by local community and century. Nevertheless, in all specifications, our equality measure farmer share had a large positive impact on numeracy. The variable city never appears significantly correlated. In this analysis, females do not have a significant disadvantage once we control for farmer share. It should be taken into account that mothers had a very important role in farmer households (Tollnek and Baten, 2017). Table 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 farmer share has a substantial influence.

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 FE; figure 3). In a second step, we regress the main explanatory variable of interest, the farmer share, on all of these variables except numeracy. In both steps we saved the residuals, of numeracy and farmer 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 farmer 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.[23] 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

[23] 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.

fully captured by the large-city-dummy variable (which was also assigned to smaller urban centres)[24]. 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, 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 farmer share.

We also considered endogeneity and a potentially confounding role of skill-selective migration (Appendix B). 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 include already 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 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

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

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 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 occupations 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 farmer 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 farmer share fell from around 18 percent to 2 percent between the sixteenth and eighteenth century (Figure 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 farmer 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[25].

5. Conclusions

We conclude that the land equality indicator "farmer 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

[25] 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.

relatively good nutrition, some farmer children were not burdened with child labour, whereas day labourer households also depended on child labour, inhibiting schooling (Tollnek & 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 farmer 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 farmer 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 farmer 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).

Do our findings have implications for the debate about the backwardness of the Spanish economy compared to other European economies until the first half of twentieth century? Our results suggest that land inequality could have played a role because it hindered numeracy formation. This is consistent with the views of Pujol (2001) and Gallego (2001) who argued that there was a lower development due to income inequality and the persistence of a traditional organic-based agriculture. Our research is also in line with the authors who found that land inequality had consequences for economic development. Pascual and Sudrià (2002), Llopis (2002), Pinilla (2004) and others found that an unequal distribution of land did not encourage large landowners to invest in technology in the countryside, due in part to the

existence of a cheap labour force in the rural world during the nineteenth and twentieth centuries (see also Palafox, 2002; Simpson, 2002; Clar and Pinilla, 2009).

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 farmer 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 farmer shares on numerical characteristics of the population. This certainly provides intriguing insights for Spain, but also more general conclusions about the role of farmer shares in human capital formation throughout world economic history.

Acknowledgements

The authors would like to thank the anonymous reviewers of *Historia Agraria* for the 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.

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. DOI:10.1017/S0022050709001120

Àlvarez-Nogal, C. & Prados de la Escosura, L. (2013). "The Rise and Fall of Spain (1270-1850)", *Economic History Review*, 66(1), 1-37. DOI:10.1111/j.1468-0289.2012.00656.x

Álvarez, B. & Ramos Palencia, F. (2018). Human capital and earnings in eighteenth-century Castile. *Explorations in Economic History*, (67), 105–133. DOI: 10.1016/j.eeh.2017.10.005.

Andújar Castillo, F.(2004). La cuestión morisca: de la General Conversión a la guerra y el destierro, *Andalucía en la Historia*, Nº4, 16-21. https://www.centrodeestudiosandaluces.es/publicaciones/descargar/374/documento/804/AH04.pdf

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. DOI: 10.1007/s10887-017-9151-1

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. https://www.mitpressjournals.org/doi/pdf/10.1162/REST_a_00403

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* (2017), 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. https://doi.org/10.1093/ereh/hew017

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. DOI: 10.1016/j.eeh.2017.12.003

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. DOI:10.12795/rea.1987.i08.04

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* (pp. 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. http://www.catastro.meh.es/documentos/publicaciones/ct/ct46/06.%20Concepcion%20Camarero.pdf

Carmona, J. & Simpson, J. (2003). *El Laberinto de la agricultura española. Instituciones, contratos y organización entre 1850 y 1936*. Prensas Universitarias de Zaragoza. http://seha.info/FILE/monografias/seha_monografia_2_juan_carmona_y_james_simspon el_laberinto_de_la_agricultura_espanola_instituciones_contratos_y_organizacion_entre_ 1850_y_1936.pdf
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, p. 669–690
Clar, E.; Pinilla, V. (2009). The contribution of agriculture to Spanish economic development. *In:* Lains, P. y Pinilla, V. (Eds.*). Agriculture and Economic Development in Europe since 1870.* Routledge, Londres, 311-332.
Clark, G. & Gray, R. (2014). Geography is not destiny: geography, institutions and literacy in England, 1837–63. *Oxford Economic Papers*, 66 (4), 1042–1069. DOI: 10.1093/oep/gpu020.
Clayburn la Force, J. (1964). Royal textile factories in Spain, 1700-1800. *The Journal of Economic History*, 24 (3), 337-363. https://www.cambridge.org/core/journals/journal-of-economic-history/article/royal-textile-factories-in-spain-17001800/8AF03CCB5C574D7D5AED3ABBDC98513B
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. https://www.sciencedirect.com/science/article/pii/S0014498309000357
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. https://revistas.uca.es/index.php/trocadero/article/view/1251
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.
Gallego, D. (2001). Historia de un desarrollo pausado: integración mercantil y transformaciones productivas de la agricultura española (1800-1936). In J. Pujol et al. (Eds.), *El pozo de todos los males. Sobre el atraso en la agricultura española contemporánea*, Barcelona: Crítica, 147-214.
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. DOI: 10.1111/j.1467-937X.2008.00506.x
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. https://EconPapers.repec.org/RePEc:ahe:dtaehe: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. DOI 10.1007/s10887-012-9081-x
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. https://doi.org/10.1016/j.eeh.2012.12.002
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.
Llopis, E. (2002). Otras caras "menos amables" de la agricultura española contemporánea. *Historia Agraria*, 28, 179-198.
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. https://www.mapa.gob.es/ministerio/pags/biblioteca/revistas/pdf_ays/a032_04.pdf
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. DOI: 10.1007/s10887-016-9132-9.
Palafox, J. (2002). Las agriculturas españolas en los siglos XIX y XX: ¿el mejor de los mundos posibles?. *Historia Agraria*, 28, 199-205.
Parejo, A. (2009). *Historia económica de Andalucía contemporánea*. Síntesis.
Pascual, P.; Sudrià, C. (2002). Notas sobre El pozo… *Historia Agraria*, 28, 207-216.
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. http://www.ub.edu/geocrit/geocritica1-100/geo58.pdf
Pinilla, V. (2004). Sobre la agricultura y el crecimiento económico en España (1800-1935). *Historia Agraria*, 34, 137-162.
Ponsot, P. (1986). *Atlas de Historia Económica de la Baja Andalucía*. Granada: Editoriales Andaluzas Unidas.
Pujol, J. (2001). La historiografía del atraso o el atraso de la historiografía. In Pujol et al. (Eds.): *El pozo de todos los males. Sobre el atraso en la agricultura española contemporánea*, Barcelona: Crítica, 14-42.
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*, Salvat-Centre de Cultura Contemporàni de Barcelona: Barcelona, 1-30.
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*, Oxford University Press, 195–225.
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. https://www.persee.fr/doc/carav_00080152_1978_num_31_1_2150
Santiago-Caballero, C. (2011). Income inequality in central Spain, 1690–1800. *Explorations in Economic History*, 48 (1), 83–96. DOI: 10.1016/j.eeh.2010.09.004.
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*, Universidad de Alicante, 549-609.
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. (2002). "El pozo", y el debate sobre la agricultura española. *Historia Agraria*, 28, 217-228.
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. http://www.historiaagraria.com/FILE/articulos/HA72_Simpson_Carmona.pdf
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. DOI: 10.1111/ehr.12382
Viñao Frago, A. (1999): Alfabetización y primeras letras (siglos XVI-XVII). In A. Castillo (Ed.): *Escribir y leer en el siglo de Cervantes*. Barcelona: Gedisa Editorial, 39–84.

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.

Table 1 Nº Observations by province and birth century

Province	Nº Total Observations 16th	17th	18th	Province	Nº Observations with occupations 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		

Table 2 Descriptive statistics

Variable	Obs.	Mean.	Std. Dev.
Numerate	26,851	0.57	0.50
Farmer 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 this variables are coded as 0 or 1.

Table 3 The effect of land equality indicator "farmer share" on individual numeracy (the likelihood of individuals not to report a rounded age) using a linear probability model (LPM)

	(1)	(2)	(3)
Farmer 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 4 The effect of the land equality indicator "farmer share" on individual numeracy (the likelihood of individuals not to report a rounded age) using a Logit model (Marginal effects reported)

	(1)	(2)	(3)
Farmer 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 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.

Figure 1 Location and sample (birth decade 1580-1760)

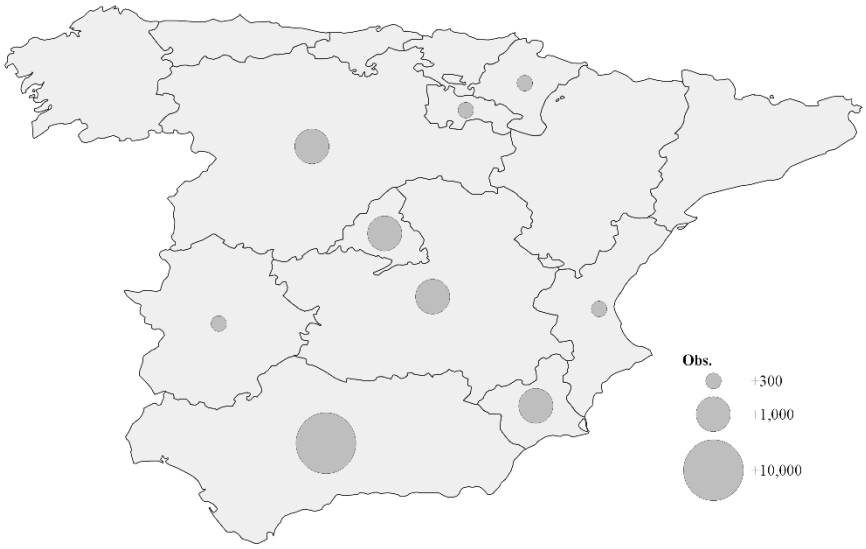
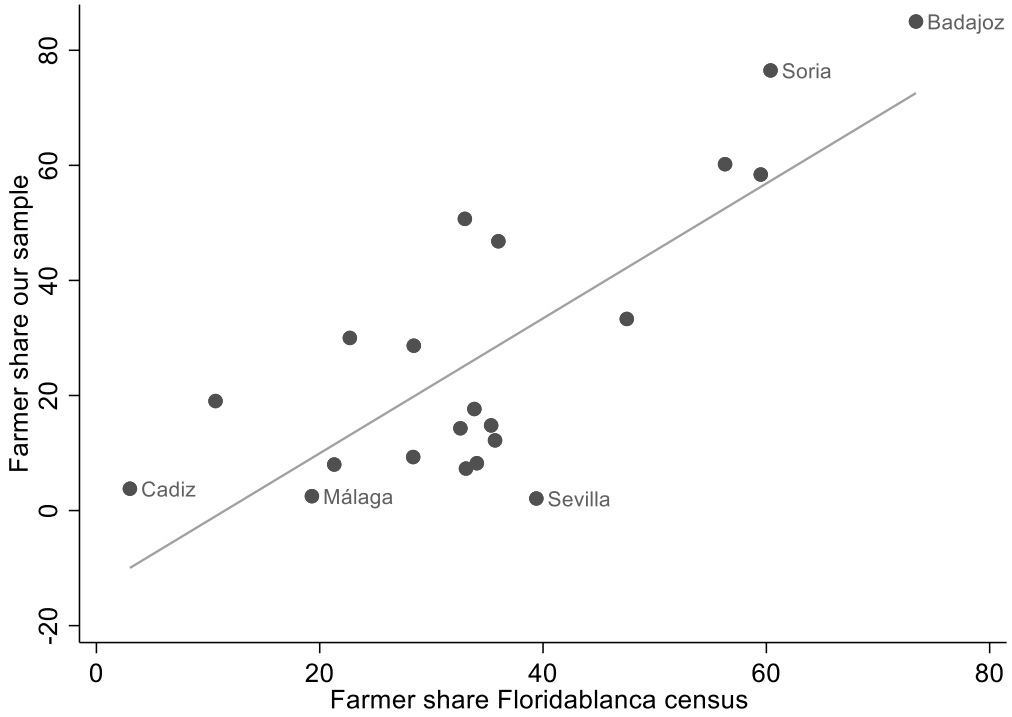
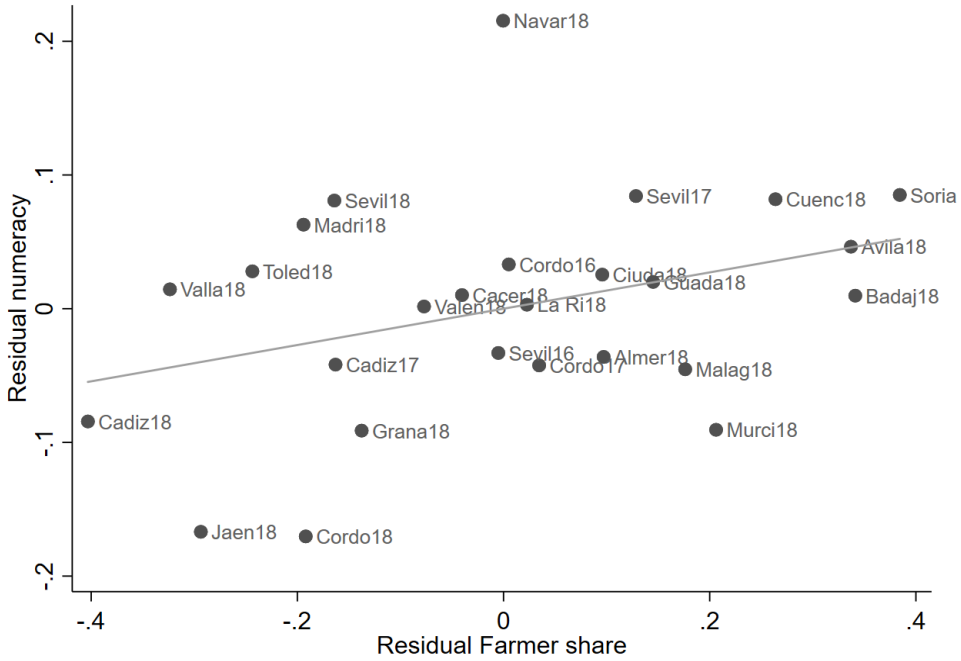


Figure 2 Comparison of the farmer share in the Floridablanca census and in our sample



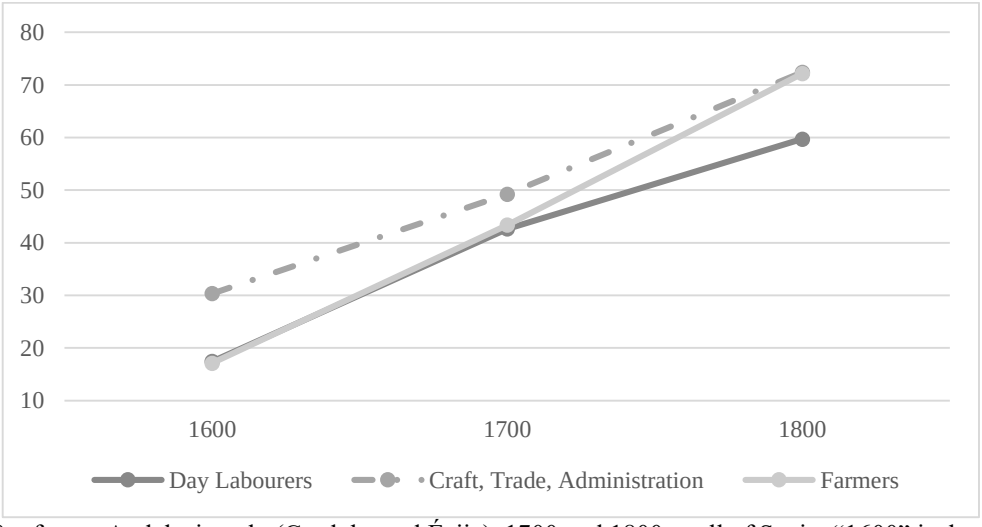
Note: we aggregate the farmer share here for only the local communities for which we have numeracy data. For example, Murcia is only represented by Lorca, Valencia only by Sueca. Consequently, this comparison does not aim at representativeness for the provinces.

Figure 3 Relation of residual farmer share and residual numeracy, on a provincial aggregate level



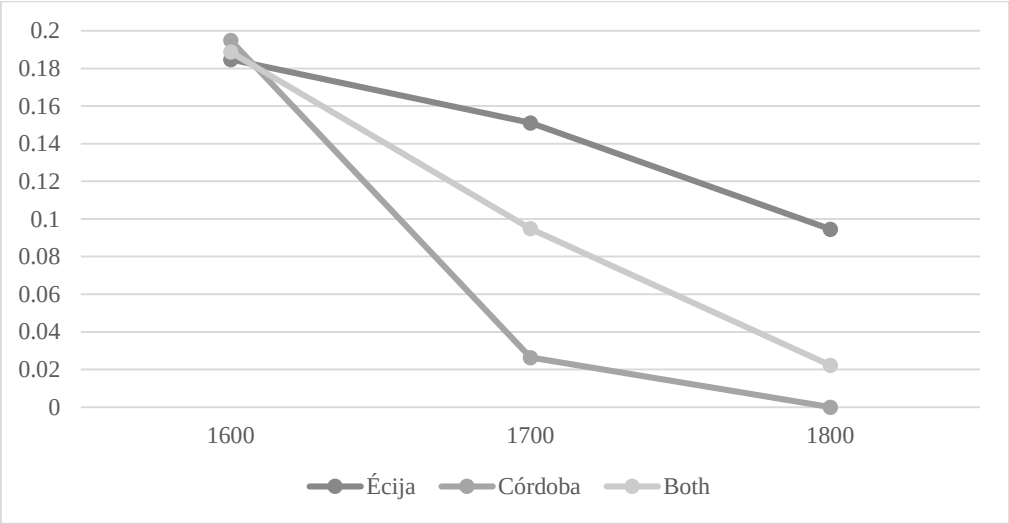
Note: in the regression analysis, we used 117 local community-birth century units. Here we aggregated by province and birth century, in order to make the figure more easily understandable.

Figure 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 16th century etc.

Figure 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 16th century etc.

Appendix A Tables

Table A1 Description of the sources

Local community	Year of Source	Source
Abanilla	1756	Cadastre of Ensenada, Family Search
Abarán	1756	Cadastre of Ensenada, Family Search
Abenójar	1752	Cadastre of Ensenada, Family Search
Adanero	1752	Cadastre of Ensenada, Family Search
Adradas	1752	Cadastre of Ensenada, Provincial Historical Archive Soria
Aguaviva de la Vega	1752	Cadastre of Ensenada, Provincial Historical Archive Soria
Aguilar de Montuenga	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Albudeite	1756	Cadastre of Ensenada, Family Search
Alcubilla del Marques	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Aldea de San Esteban	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Aldeasenor	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Alentisque	1752	Cadastre of Ensenada, Provincial Historical Archive Soria
Alía	1752	Cadastre of Ensenada, Family Search
Almaluez	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Almarza	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Almazán	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Almería	1753	Cadastre of Ensenada, Family Search
Almuñecar	1752	Cadastre of Ensenada, Provincial Historical Archive Granada
Alovera	1751	Cadastre of Ensenada, National Historical Archive Madrid
Andaluz	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Arcos de Jalon	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Arevalo	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Arganda	1753	Cadastre of Ensenada, Regional Archive Madrid
Atauta	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Berchules	1752	Cadastre of Ensenada, Provincial Historical Archive Granada
Bubion/Capileira	1750	Cadastre of Ensenada, Provincial Historical Archive Granada
Cavanillas	1751	Cadastre of Ensenada, National Historical Archive Madrid
Colomera	1752	Cadastre 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	Cadastre of Ensenada, National Historical Archive Madrid
Estepona	1752	Cadastre of Ensenada, Provincial Historical Archive Granada
Fuente El Sol	1752	Cadastre of Ensenada, Family Search
Granada	1752	Cadastre of Ensenada, Provincial Historical Archive Granada
Inestrillas	1752	Cadastre of Ensenada, Family Search

Iznalloz	1753	Cadastre of Ensenada, Provincial Historical Archive Granada
Jaen	1771	Padrón, Family Search
Laujar de Andarax	1751	Cadastre of Ensenada, Provincial Historical Archive Almería
Logroño	1751	Cadastre of Ensenada, Family Search
Loja	1750	Padrón, Family Search
Loja	1753	Cadastre of Ensenada, Provincial Historical Archive Granada
Lorca	1756	Cadastre of Ensenada, Family Search
Málaga	1751	Cadastre of Ensenada, Municipal Archive Málaga
Málaga	1776	Padrón, Municipal Archive Málaga
Marchamalo	1752	Cadastre of Ensenada, National Historical Archive Madrid
Montilla	1752	Cadastre of Ensenada, Family Search
Móstoles	1753	Cadastre of Ensenada, Regional Archive Madrid
Navas de San Juan	1752	Cadastre of Ensenada, Family Search
Olite	1786	Floridablanca, Municipal Archive Olite
Ontalvilla de Almazán	1753	Cadastre of Ensenada, Provincial Historical Archive Soria
Pinto	1753	Cadastre 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	Cadastre of Ensenada, Family Search
Sueca	1794	Padrón, Family Search
Toledo	1752	Cadastre of Ensenada, Regional Archive Madrid
Torreandaluz	1752	Cadastre of Ensenada, Provincial Historical Archive Soria
V. de la Torre	1751	Cadastre of Ensenada, National Historical Archive Madrid
Valdecaballeros	1753	Cadastre of Ensenada, Family Search
Villanueva del Rey	1750	Cadastre of Ensenada, Family Search
Yunquera de Henares	1751	Cadastre of Ensenada, National Historical Archive Madrid
Ziria	1753	Cadastre of Ensenada, Provincial Historical Archive Soria

Table A2 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
Castilla La Mancha	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
Castilla La Mancha	Toledo	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	

Appendix B 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 farmer 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)[26]. As a result, during the repopulation process, a small number of aristocratic families and ecclesiastical entities emerged as owners of large

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

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 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[27].

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. DOI: 10.1007/s10887-017-9151-1

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. DOI: 10.1016/j.eeh.2017.12.003

Bernal, A.M. (1987). Latifundios, jornaleros y paro agrícola, *Revista de Estudios Andaluces*, 8, 67-86. DOI:10.12795/rea.1987.i08.04

Carmona, J. & Simpson, J. (2003). *El Laberinto de la agricultura española. Instituciones, contratos y organización entre 1850 y 1936*. Prensas Universitarias de Zaragoza. http://seha.info/FILE/monografias/seha_monografia_2_juan_carmona_y_james_simspon el_laberinto_de_la_agricultura_espanola_instituciones_contratos_y_organizacion_entre_ 1850_y_1936.pdf

Cinnirella, F. & Hornung, E. (2016). Landownership concentration and the expansion of education. *Journal of Development Economics*, 121, 135-152. DOI: 10.1016/j.jdeveco.2016.03.001

De Albornoz de la Escosura, A. (1963). El "milagro" de la agricultura española. La distribución de la propiedad. *Comercio Exterior*, 764–767 http://revistas.bancomext.gob.mx/rce/magazines/676/11/RCE_11.pdf

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. https://hdl.handle.net/11441/81207

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. DOI: 10.1007/s10887-016-9132-9

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

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