European
Historical
Economics
Society

EHES WORKING PAPERS IN ECONOMIC HISTORY | NO. 95

Comparing Income and Wealth Inequality in Pre-Industrial Economies: Lessons from 18th-Century Spain

Esteban A. Nicolini
Economics Department, Universidad Carlos III de Madrid
Universidad del Norte Santo Tomás de Aquino
INVECO, Universidad Nacional de Tucumán

Fernando Ramos Palencia
Economic History Department, Universidad Pablo de Olavide

MARCH 2016

EHES Working Paper | No. 95 | March 2016

Comparing Income and Wealth Inequality in Pre-Industrial Economies: Lessons from 18th-Century Spain

Esteban A. Nicolini*
Economics Department, Universidad Carlos III de Madrid
Universidad del Norte Santo Tomás de Aquino
INVECO, Universidad Nacional de Tucumán

Fernando Ramos Palencia**
Economic History Department, Universidad Pablo de Olavide

Abstract

Most research on the history of inequality in pre-industrial economies has focused on either wealth or income. Characterizing the distribution of wealth (resp., income) is problematic owing to insufficient information about the distribution's low (resp., high) end. Because the sources and methodologies differ between these two approaches, their results are not readily comparable and it is difficult to establish links between the respective distributions that result. In this paper, we shall use a unique data set for different regions of Spain circa 1750 and present results—the first for any pre–20thcentury economy—on both income and wealth distributions for the same sample of households. Information on wealth and income is derived from (respectively) probate inventories and the Ensenada Cadastre. Our main findings are that poor households are not entirely absent from the data set of inventories, that a household's position in the income distribution is strongly correlated with its position in the wealth distribution, and that increases in a household's wealth are associated with less-than-proportional increases in its income.

JEL classification: D31, N33, O15

Keywords: inequality, income, wealth, Spain, probate inventories, Ensenada Cadastre

For helpful comments and discussions on the topic we thank Guido Alfani, Julio Cáceres, Leandro Prados de la Escosura, and Jaime Reis. Esteban Nicolini gratefully acknowledges financial support from the Spanish Ministry of Science and Innovation through Project ECO2011-25713, CIUNT (subsidy 26/F410) and from Ministerio de Ciencia y Tecnología – Argentina through PICT 2429-2013. Fernando Ramos Palencia acknowledges financial support from the Spanish Ministry of Education and Sciences through Project ECO2012-38028. The authors are grateful also for the inestimable help provided by Servicio de Reproducción de Documentos de Archivos Estatales (SRDAE) and Archivos Históricos Provinciales from Guadalajara, Granada, Madrid, and Palencia.

* Esteban A. Nicolini, E-mail.: esteban.nicolini@uc3m.es
** Fernando Ramos Palencia, E-mail.: fcrampal@upo.es or fernando.ramos.palencia@gmail.com

Notice

The material presented in the EHES Working Paper Series is property of the author(s) and should be quoted as such. The views expressed in this Paper are those of the author(s) and do not necessarily represent the views of the EHES or its members

1. Introduction

The significant increase in global income inequality over the last two centuries is one of the most striking and influential aspects of the modern process of economic growth (Bourguignon and Morrison 2002; Moatsos et al. 2014). One leading hypothesis about the evolution of income inequality in Europe is that—over the three centuries preceding the Industrial Revolution—there was a divergence of real wages across countries, which is consistent with the observed increase in across-country inequality (Allen 2001). The evidence on within-country inequality is mixed: in some countries, such as Holland and Italy (van Zanden 1995; Alfani 2010, 2015), inequality during this period increased; in other countries (e.g., Portugal; Reis et al. 2012) inequality decreased or oscillated with no clear trend (e.g., Spain; Álvarez-Nogal and Prados de la Escosura 2013). Some works that estimate economic inequality use income as the relevant variable while others use wealth.

Perhaps the most widely adopted approach to measuring income inequality in pre-industrial times is to examine so-called social tables, wherein the focal population is divided into groups based on occupations and/or social classes and then an average income is assigned to each group; here income is usually inferred based on indirect information from contemporaneous observers (Williamson and Lindert 1980; Milanovic et al. 2007, 2011).[1] This way of assessing household income is relatively straightforward for the low part of the distribution because that is where most income derives from labor, occupations are relatively homogenous, and worker income is easy to infer. For the high part of the distribution, in contrast, there are many sources of income and occupations are quite idiosyncratic; it follows that inferring the income of relatively affluent households based solely on occupation or social class is subject to a wide margin of error.

[1] Atkinson et al. (2009) use tax records to study the historical evolution of "top" incomes.

Research on wealth inequality before the 20th century has relied mostly on data sets based either on tax records (Soltow and van Zanden 1998; Alfani 2010) or on collections of probate inventories (Jones 1978; McCants 2006; Canbakal 2013). However, with these techniques the society's poorer segments are under-represented, which leads to severe selection bias.[2] Some scholars have sought to calculate just how under-represented certain population subgroups are and then to assess the implied selection bias for Colonial America (Main 1974; Smith 1975; Jones 1978).

This paper presents a new data set that we use to calculate economic inequality in Spain based on information, circa 1750, from Palencia, Madrid, Guadalajara and Granada. This data set has some unique characteristics. First, it combines information from two different sources: probate inventories, which contain detailed descriptions of household wealth; and the Ensenada Cadastre, a mid-century government census that contains information about household income. Second, the data set enables us to link the households from the set of inventories with their corresponding records in the Cadastre; this connection makes it possible to analyze the relationship between the *income* of a household when the Cadastre was produced and the *wealth* of that household some years later, when its head passed away. Third, given the practically complete coverage by the Cadastre, we can use the income distribution based on this source to determine the extent of under- or over-representation representation of the different parts of the distribution in the set of inventories. One useful byproduct of that analysis is the ability to weight observations in the set of inventories, thereby reducing the problem of selection bias.

Our main goal is to provide a methodological contribution linking the distributions of income and wealth so that we can propose hypotheses via which their differences can be better understood. The results reported here confirm that the set of surviving inventories is

[2] Soltow and van Zanden (1998, p. 20) warn that "wealth statistics generally are quite deficient in telling us anything about the condition of people below median income."

strongly biased toward the upper part of the distribution; nonetheless, the existence of inventories from the poorest quintile in each of our data set's three regions confirms that the poor are not completely absent from probate inventories. Even though our information on wealth and income come from completely independent sources, there is a remarkably high association between the two variables that suggests they well capture some meaningful dimension of economic affluence. A simple econometric estimation of the relationship between these two variables indicates that, across households: (i) wealth increases more than does income (i.e., the wealth elasticity of income is less than 1); and (ii) if a household head works mainly in the secondary and/or tertiary sectors—that is, in manufacturing and/or trade rather than in agriculture—then that household's income will be higher than predicted by its wealth alone.

The rest of the paper is organized as follows. In Section 2 we discuss previous estimates of economic inequality from the literature. Section 3 describes the Spanish economy's historical context, and the data are presented in Section 4. Our principal findings are summarized in Section 5, and we offer some conclusions in Section 6.

2. Inequality estimation prior to households' surveys

The estimation of economic inequality before the 20th century can only be based on information collected for purposes other than inequality analyses. The field's three leading approaches are constructing social tables, mining tax records, and analyzing probate inventories.[3] The *social tables* are based on dividing the population (or subset of income

[3] An alternative -and indirect- approach was suggested by Williamson (2002) who argued that the ratio of average land rent to average unskilled wages is a good proxy for economic inequality because rents from land (resp., wages from labor) figure largely in the income of the households in the top (resp., bottom) of the distribution. Other authors have modified this idea by using per-capita gross domestic product instead of land rents (Dobado González and García Montero, 2010; Álvarez-Nogal and Prados de la Escosura 2013).

earners) into groups (the usual criterion is occupation or social status) and then assigning an average income to each group. If one assumes that most inequality stems from differences *across* groups rather than *within* a group, then this methodology is similar in spirit to the treatment of modern data sets when populations are divided into (say) quintiles.[4] Once the profile of incomes for a population is constructed in this way, standard measures of inequality can be calculated (Williamson and Lindert 1980; Milanovic et al. 2007; Bertola et al. 2009; Milanovic et al. 2011).[5]

Tax records are another important source of information when seeking to estimate inequality, especially when one considers that the variables of wealth and income have emerged naturally as tax bases in many historical contexts. An already classic example of this approach is in studies by Herlihy (1978) and Herlihy and Klapisch-Zuber (1985) of the Florentine Catasto, which registered households' wealth. Other very well-known example is given by Piketty and co-authors' studies of wealth inequality looking at estate tax returns (Piketty et al. 2006) or income inequality looking at income tax returns (Piketty and Saez 2003). More recently, Alfani (2010) explores wealth inequality in 16th- and 17th-century Ivrea (Italy) based on records of the *estimi*, a tax on the value of real estate owned by households. Other approaches are based on other kinds of fiscal records: Soltow and van Zanden (1998) use the introduction of an income tax in 1749 in the States of Overijssel to

[4] In modern data sets, households are ordered according to their income and so, by definition, the quintiles (or deciles) are nonoverlapping subgroups. In social tables, however, the richest households of one group can be richer than the poorest households of the next highest group. Modalsli (2015) analyzes how this characteristic of social tables affects standard measures of inequality (e.g., the Gini coefficient).

[5] Milanovic et al. (2007) use social tables to calculate Gini coefficients for 14 pre-industrial societies. For Old Castile in 1752, these authors examined groups of households having similar income and also used information from the Ensenada Cadastre (as summarized in Ramos Palencia 2010). The Gini coefficients that Milanovic et al. found for Modern Europe range from 44.9 in England and Wales in 1688 to 63.0 in Holland in 1732; in the middle of this range is Old Castile, with a Gini index of 52.3 circa 1750.

analyze inequality in pre-industrial Holland; and Santiago-Caballero (2011) uses the tithe paid by each grain producer as a proxy for income and then deduces from that information the extent of income inequality in 18th-century Guadalajara, Spain. For the province of Palencia, Nicolini and Ramos Palencia (2015) use a portion of the Ensenada Cadastre that was originally collected with the intention of transforming the fiscal system in Old Castile while gathering information about households' income *and* the detailed sources of that income (mainly land, livestock, and labor). Of course, information from such sources is likely to be far from perfect; it is common for taxes to be based on only a subset of the household's assets (usually land or real estate) or on a specific activity or type of consumption linked in some way to the household's income.[6]

Other important sources for any analysis of historical economic inequality are the scattered but numerous collections of probate inventories from around the world.[7] Although such inventories provide extremely rich and detailed descriptions of the wealth of many households, they are not a priori suitable for the study of inequality owing to selection biases (Lindert 1981; Jones 1982). Two biases in particular are commonly identified: first, the age distribution of deceased household heads differs from the age distribution of all household heads; second, richer households are naturally over-

[6] Soltow and van Zanden (1998, p. 26) analyze income inequality in 16th-century Holland; as a proxy for household income, these authors use the *tiende penning* (the tenth penny), a tax based on the household home's rentable value. Another example is Alfani (2010), who uses the value of real estate to proxy for total wealth in Ivrea during the Early Modern period.

[7] Jones (1980, 1982) uses the available probate inventories to estimate aggregate wealth and wealth distribution of the American Colonies during the second half of the 18th century. Lindert (1981) analyzes wealth inequality in England between 1670 and the 20th century, and McCants (2007) uses probate inventories to asses living conditions of middling and poor households in 18th-century Amsterdam. Canbakal (2013) uses an extensive set of probate inventories to analyze the evolution of inequality in the Ottoman Empire between the 16th and 19th centuries.

represented within the survival inventories. We follow the recommended approach of dealing with selection bias by constructing weights (or multipliers) to correct for the observed bias.[8]

3. Historical context

In Europe, the period between the 16th and 18th centuries witnessed a widening gap—as measured by income per capita—between a group of leading regions (England and Holland) and another group of regions with small or zero growth rates. This phenomenon is known as *small divergence* (Allen 2001), and the outcome in this case was an increase in across-country inequality.[9] For Spain, the outcome of this particular small divergence between the 16th and 18th centuries was to fall clearly behind the European leaders (Álvarez-Nogal and Prados de la Escosura 2007, 2013).

Our knowledge of the evolution of economic inequality within countries or regions during this period has expanded considerably in recent years. In a seminal paper, van Zanden (1995) documented increases in income inequality and wealth inequality for Holland as well as a positive association for Europe between economic inequality and both economic growth and urbanization. Since the publication of that research, other scholars have added evidence for other countries. Alfani (2010, 2015) claims that wealth inequality in northern Italy increased during Early Modern times even though income per capita stagnated; for the Low Countries (Flanders, Brabant, and Holland), Ryckbosch (2016) finds growth in economic inequality during the two centuries before the Industrial Revolution.

[8] Lindert (1981, p. 660) states that, "to derive such multipliers, we need either (a) true wealth distributions for benchmark periods and places or (b) data on other attributes of the probated individuals, primarily attributes linked strongly to their wealth and available for the entire population of adults or household heads." As described in Section 4, our weights use income as an attribute linked to wealth.

[9] Milanovic (2005) discusses in some detail the concepts of global inequality and international inequality.

Reis et al. (2012) argue that income inequality did not increase in a Portuguese economy that stagnated from 1550 to 1700. In the Ottoman Empire, Ergene et al. (2013) find declining inequality but also economic stagnation during the 18th century. More generally, Canbakal (2013) reports that inequality tends to be less pronounced in rural areas than in urban agglomerations. For Spain there are estimates by Santiago-Caballero (2011), whose use of tithe data to proxy for income indicates that inequality was stable during the 18th century in central Castile, and by Álvarez-Nogal and Prados de la Escosura (2013), who use the ratio of per-capita income to unskilled wages and find that—except for the early 17th century—inequality declined (increased) during periods of economic depression (expansion).

Even though the economic evolution of Modern Spain can be described in general terms as a process of relative retardation, the 18th century featured positive changes in Spanish demographic growth, economic expansion, administrative (colonial) reform, and geopolitical relevance. Yet one must bear in mind the series of shadows and light that characterized this period. The vacant throne of the Spanish monarchy triggered the War of the Spanish Succession (1701–1714), which pitted France against England, the Netherlands, and Austria. This war coincided with the beginning of British global hegemony and the arrival in Spain of the French Bourbons. When the conflict finally ended, Phillip V (1700–1746) set out to emulate France—with mixed results—by advocating economic unification and political centralization. Although the diverse coinage circulating in the different territories were withdrawn and replaced with a single currency (following the system in place in Castile), fiscal union was not achieved. The *Nueva Planta* decrees abolished the remaining *fueros* (local privileges and laws) of the Crown of Aragon (encompassing Aragon itself as well as Catalonia and Valencia) because they had mostly supported the Habsburg candidate, Archduke Charles, during the War of the Spanish Succession. At the same time, a single tax (the *equivalente*) was imposed in the form of a

quota levied on rural and urban properties and on the profits deriving from trade, industry, and labor.[10] In contrast, the fiscal prerogatives and exemptions of the Basque Country and Navarre were preserved because the inhabitants of those regions had supported the Bourbons. In Castile, the fiscal reform advocated by the Marquis de Ensenada (and inspired by the Crown of Aragon's reform) was a failure. In 1749, Fernando VI decided to divide the kingdom into provinces. In charge of each provincial capital would be a *corregidor*, which later became that province's *intendente*. This *intendente* was the royal official in charge of tax collection in the province.

From an economic viewpoint, Herr (1960, p. 128) points out that "one could draw a geographical line that separated the North and East—where industry[11] was thriving and the farmers were well-off—from the Centre and South, where industry was backward and the farmers and day labourers in the countryside were exploited by the rural oligarchy."[12]

---

[10] According to Ferrer i Alós (2002, pp. 29–32), the intention was to create a single tax calculated on the basis of personal wealth However, for logistical reasons the Bourbon bureaucracy opted to implement a quota system. That system involved setting a more-or-less fixed amount that was then divided among the localities on the basis of reports gathered in each of them.

[11] In an effort to stimulate industry and finance high internal transportation costs, the Bourbon mercantilist system opted for increasing duties on foreign imports and eliminating the monopoly of Seville and Cadiz in trade with the Latin American colonies. This last measure was immensely beneficial to the merchant navy and to manufacturing industries in Catalonia (paper and cotton), Valencia (silk, linen fabrics, and tiles) and the Basque Country (iron and steel). In the interior, meanwhile, official policy was to support manufacturing (e.g., the textile products of Guadalajara) with subsidies to compete with the luxury goods that were being imported from abroad. For their part, industries that produced essential goods were dominated by the guilds—institutions that monopolized nearly all industrial activity in the cities (exception for Catalonia) and of which the Bourbon politicians were highly critical.

[12] According to the 1797 census, however, 22% of those employed in agriculture were landowners and there were notable exceptions to groupings based on an imaginary line that extended from the northeast (Salamanca) to the southeast (Albacete). For instance, landowners accounted for some 50% of the total in the

With the exception of Madrid and its more than 150,000 inhabitants (c. 1790), Spain's large cities were located near the sea. The populations of Barcelona (Catalonia) and Valencia (Valencia province) rose to 100,000 inhabitants by the end of the 18th century. In the south, the provinces of Andalusia stood out: the city of Cadiz was home to more than 10,000 individuals; Malaga, 50,000; and Seville and Granada, 80,000 each. In contrast, no city in Castile had a population of more than 25,000. Based on the information provided by the Ensenada Cadastre and re-compiled by Matilla Tascón (1947), the provinces of the former Crown of Castile with the highest income per capita were Madrid (1,453 *reales*), Seville (641 *reales*), and Guadalajara (601 *reales*); those with the lowest income per capita were Granada (322 *reales*), Leon-Asturias (278 *reales*), and Galicia (202 *reales*). The average per-capita income in Castile was about 433 *reales*.[13]

---

north (Aragon, Navarre, Biscayan, and Galicia) even as the corresponding figure varied between only 3% and 7% in Andalusia. Landownership was distributed in decreasing order of importance among private owners whose lands were "tied" and could not be sold (*mayorazgos*, which were tied to a particular family, and *señoríos*, landed estates that were a royal privilege granted to the nobility and wealthy commoners); towns and cities (commonweals that were a crucial source of revenue for local councils); the Church; the Crown; and finally private owners of land that was not encumbered in any way. We remark that, more often than not, the "landowner" was the owner of a house and the small piece of land on which it stood. Note also that such landowners might reside on an estate to which rent must be paid. Practically speaking, the aristocratic oligarchy consisted mainly of the *hidalgos* (noblemen) and the urban super-rich. Indeed, more than half of the towns paid tribute to landlords in Extremadura and Western Andalusia and in Valencia; for more details, see Herr (1960, pp. 28–29).

[13] This figure reflects author analysis based on the data in Matilla Tascón (1947, appendices) and on the number of inhabitants in 1752 according to GRUPO 75 (1977, p. 64). To work out the total income in each province, we included income generated by the lay sector (or by the Church) from rural properties, urban properties, livestock, ground rent and other forms of rent, and interest on loans as well as all other revenue derived from industrial or commercial activities and personal work.

For this paper we chose, from north to south, three regions; the first one corresponds to the province of Palencia; the second one, in the Centre of Castile, is comprised by Guadalajara City and some small towns close to Madrid; the third one, in the South, incorporates some in the province of Granada; see Figure 1. Palencia, which is situated in the north of Spain, had a population of about 106,440 distributed among the following *comarcas* (areas): El Cerrato Palentino, Tierra de Campos (the city of Palencia belongs to this *comarca*), Saldaña-Valdavia, Boedo and La Ojeda Valley, Aguilar, and Guardo-Cervera. In the second half of the 18th century, the population of the province of Palencia was distributed irregularly; more than two thirds resided in the province's south (Tierra de Campos and El Cerrato Palentino). These areas were characterized by relatively large population centers and more than a hundred neighbors[14] that were nonetheless geographically distant from each other. In southern Palencia, wheat was the main product while wine and vegetables played a secondary role; livestock was pretty much limited to the animals used for agricultural work or by peddlers. According to Larruga (1787/1995), Palencia was the "most industrious province of Castile". In fact, the *comarcas* of Tierra de Campos and El Cerrato Palentino had significant secondary and tertiary sectors. In northern Palencia (Guardo-Cervera and Aguilar), population density was low and the population was concentrated in many small, closely located nuclei. These areas were characterized by livestock activity, linen production, and mule drivers. The industry of low-quality textiles (domestic production) was of signal importance in the valleys of Boedo and La Ojeda. The most populated towns in northern Palencia, according to the Ensenada Cadastre of 1759, were Palencia (9,639 inhabitants) and Paredes de Nava (3,395), both in Tierra de Campos.[15]

---

[14] The census from Ensenada Cadastre reports the population in neighbors, not inhabitants. Spanish historiography for that period generally uses the following equivalence: 1 neighbor ≈ 4 inhabitants.

[15] See Marcos Martín (1985, p. 22) and Camarero Bullón (1990, pp. 231–49).

[[insert Figure 1 about here]]

This paper analyzes two areas from the center of Castile: Las Vegas, which is close to Madrid City; and Guadalajara City. The Las Vegas economy was predominantly agro-pastoral and produced mainly cereals, vegetables, oil, wine, linen, silk, and fruit as well as sheep, goats, cattle, and pigs. In this *comarca*, the most heavily populated towns were Colmenar de Oreja (1,279 neighbors) and Chinchón (1,217 neighbors). The city of Guadalajara (5,218 inhabitants) was an important industrial nucleus during the 18th century because the Bourbons established the *Real Fábrica de Paños* (the former Royal Cloth Mills) there in 1719. This state-owned company organized its production around guilds, was situated in a competitive market, and recruited foreign experts (La Force 1964). The company began to decline during the 1790s, the Napoleonic Wars accelerated that process, and the company closed in 1822.

Finally, we studied two areas in Granada: Lecrín Valley in the southwest and Baza in the northeast. Lecrín Valley (2,398 neighbors, some 9,484 inhabitants) was basically an agricultural economy.[16] The most populated town was Albuñuelas (294 neighbors) and Pinos (260 neighbors). Baza, which exceeds 1,700 km^2, is the largest area in Granada. There are three main zones in that province: the Sierra of Baza, the Meseta, and the Vega. The Meseta and the Vega are flat plains that surround the city of Baza and are dedicated primarly to agriculture. The Sierra of Baza is a rocky massif with deep valleys and escarpments. According to the Ensenada Cadastre, Baza had a population of around 5,366

[16] Many households earned some income apart from that associated with the household head's main job. Some households were engaged in the manufacture and sale of *pleita* (a ring or strip of straw twisted in several branches; sewn *pleita* were used to make mats, hats, pouches, etc.); members of other households were the salesmen, mule drivers, and peddlers who provided a link between the poor villages in these areas and Granada City.

neighbors (more than 20,000 inhabitants); its largest towns were Baza (1,610 neighbors) and Las Cuevas (1,302 neighbors).

4. The data

Two different data sets are used in this paper. The first one includes information contained in 194 probate inventories (PIs); the second one consists of more than 6,000 Ensenada Cadastre records on the characteristics of households from the same areas as the inventories. These areas are situated in the current provinces of Palencia (north), Madrid (center), Guadalajara (center), and Granada (south); see Figure 2.

[[insert Figure 2 about here]]

A *probate inventory* is a comprehensive list of all the goods owned by a deceased individual at the time of death, and it was usually elaborated by a notary or judicial authority within a few days of that time. Although there is some variability in the structure, format, and style of these inventories, one can characterize reasonably well the structure of a typical Castilian post-mortem inventory from the middle of the 18th century until about mid-19th century; thereafter, the rich descriptions previously given in such inventories progressively disappeared. Those descriptions of durable and semi-durable goods (personal clothing, property, and household objects, inter alia) became less necessary over time as the total value of these goods came to be a smaller percentage of the total inventoried assets.[17]

The Ensenada Cadastre (EC) is a census that was undertaken in the middle of the 18th century with the purpose of improving the Spanish monarchy's fiscal organization.[18] The aim of the Marquis of La Ensenada (Secretary of the Treasury from 1743 to 1754) was to establish a single tax (*única contribución*)—universal and also proportional to taxpayer

[17] For a summary of (and historiographical references on) the characteristics of probate inventories as historical sources in Old Castile, see Nicolini and Ramos Palencia (2010, pp. 153–55).

[18] Ruiz Torres (2008, pp. 280–85).

income—that would replace other taxes collected from the provinces (*rentas provinciales*). The *alcabala* (sales tax), *cientos* (hundreds), and *millones* (millions) were prominent indirect taxes that had a regressive effect on income distribution within the economy (Comín and Yun-Casalilla, 2015). The proposed tax reform included direct taxation, recovery of income previously transferred, curtailment of tax exemptions (especially those granted to ecclesiastical institutions), and development of a simplified tax system. The Ensenada's own downfall was due to British pressure on the Spanish Court in 1754 because of his French sympathies during the Seven Years' War (1756–1763). These events greatly reduced political support for his tax reform, which was never implemented although it did generate a large amount of detailed information about the Spanish economy in that period. The Cadastre, which covered the former Crown of Castile (see Figure 2), was carried out between 1749 and 1759 (approximately); it is an excellent source for the study of economic activities in general and economic inequality in particular because it accounted in detail for the income of each household in each locality of Castile and also analyzed the incomes of usually exempted social groups (e.g., members of the Church and the nobility).[19]

The PIs in our data set are all the available inventories between the years 1753 and 1768 from 11 geographic units (GUs) in three different regions of Castile.[20] There were two motivations for the particular geographic coverage of these data. First, we sought to collect all the available inventories in the province of Palencia for the purpose of analyzing several economic aspects of that province. For this purpose we used 116 inventories from 7 of its GUs (Palencia City, Boedo and Ojeda Valleys, Cerrato, Guardo-Cervera, Saldaña-Valdavia,

[19] See Nicolini and Ramos Palencia (2015) for a detailed description of the EC as a source for studying income inequality.

[20] The great challenge (and major difficulty) of this research is finding probate inventories whose household heads are included in the EC—given that only a limited number of each locality's probate inventories have survived.

Tierra de Campos, and Aguilar). Second, we wanted to incorporate two additional regions—one in the Centre of Castile (close to the city of Madrid) and another in the south (in the province of Granada)—so that we could expand on the first data set's number of observations and geographical coverage and then incorporate into the analysis any locality whose productive profile differed from that observed for Palencia. We selected two GUs in each region (Guadalajara City and Las Vegas in the Centre; Baza and Lecrín Valley in Granada) and collected all the available inventories in each of these four geographic units: 49 inventories in the Centre and 29 inventories in Granada. These were combined with our 116 Palentine PIs to create our 194-PI Data Set 1 (DS1).[21]

An important difference between the 116 inventories in Palencia and the 78 inventories in the Centre and Granada is that, in the former, we have full coverage of the available inventories and they include PIs from all seven GUs. Although we collected all available PIs from two GUs each in the Centre and Granada, those areas are likely not representative of the entire province (see the maps in Figure 1).

---

[21] We collected 194 probate inventories from 43 cities, towns, and villages in 11 GUs. In the list that follows, the geographic units are (arranged alphabetically and) italicized: *Aguilar*, 17 probate inventories (3 from Aguilar and 3 each from Bascones de Valdavia, Cordovilla, Corvio, Foldada, Matamorisca, Orbo, Quintanilla de las Torres, Respenda, Revilla de Santullán, San Martín de Perapertú, Valle Espinoso, and Villabellaco); *Baza*, 16 PIs (6 from Baza City and 10 from Cúllar Baza); *Boedo and Ojeda Valleys*, 10 PIs (6 from Prádanos de Ojeda and 4 from Villabermudo); *Cerrato*, 12 PIs (8 from Cevico de la Torre, 3 from Hontoria de Cerrato, and 1 from Soto de Cerrato); *Cervera*, 21 PIs (1 each from Barcenilla, Campo, and Celada, 2 from Cervera, 1 each from Estalaya, Herreruela, Lores, and Muda, 2 from Resoba, 3 from Rueda, 1 each from San Cebrián de Muda and San Martín de los Herreros, 3 from Triollo, and 2 from Verdeña); *Guadalajara City*, 12 probate inventories; *Las Vegas*, 37 PIs (12 from Carabaña, 2 from Colmenar de Oreja, 5 from Orusco, and 18 from Valdaracete); *Lecrín Valley*, 13 PIs (all from Padul); *Palencia City*, 24 probate inventories; *Tierra de Campos*, 32 PIs (16 each from Paredes de Nava and Villarramiel).

We complete DS1 by linking the deceased person named in each PI to the corresponding record in the EC. So for each household we assemble economic information that includes total wealth (from the PI) at the time of the household head's death as well as total household income as recorded (by the EC) several years before.

[[insert Table 1 about here]]

The number of observations in DS1 is 194, and the geographic distribution of these observations is described in column [8] of Table 1. The two main variables in our DSI that are based on the PIs are as follows.

- Wealth: the sum of real estate assets (total of urban and rural properties), financial assets (cash, credits, debts, land rents, advance payments for the funeral service and of estate shares to prospective inheritors), capital assets (farming implements and tools, winemaking and measuring equipment, implements for livestock and for textile production, raw textiles, and livestock) and durable or semi-durable consumption goods (all types of clothes, bed linen, table linen, personal items, articles related to household kitchen equipment and furniture, pictures, books, and jewelry). More details are given in Nicolini and Ramos Palencia (2010).
- Year: from 1753 to 1768.

In addition to these variables, our DS1 also includes information provided by the EC for these 194 households.

The households described by surviving PIs can hardly be other than a biased selection of all households in the population; that is, households with high income and/or wealth are over-represented in the sample of probate inventories. In order to approximate the whole distribution of households in each of the 11 geographical units, we selected one (or two) towns in each GU and recorded all relevant information provided by the EC for every household in those towns; this information is systematized in our Data Set 2 (DS2), which comprises 6,214 households. These data enabled us to approximate the income

distribution for each GU. Table 1 reports the information required for this reconstruction process. In the first three columns we present (respectively) the 11 GUs included in our data sets, the provinces in which each is located, and that province's population. Columns [4]–[6] show, for each GU, the number of towns in each province, the average number of households in each town, and total number of households. Column [7] gives the number of PIs in each locality, while columns [8] and [9] list (respectively) the towns included in our DS2 and each town's number of sampled households.

All information in our DS2 comes from the EC, and it yields the following main variables used in the paper.

- Income: this variable is measured in *reales*. It includes income derived from: land; buildings and non-land property (e.g., houses in the city, mills in the countryside); livestock; taxes, fees, credits, and/or debts; and personal earnings. Those earnings include labor income deriving from the household head's main activity, which was imputed by census officials while assuming a daily income and a certain number of days per year (120 days for agricultural laborers, 180 days for workers in secondary and tertiary sectors, and 360 days for shepherds). Personal earnings also include income obtained from trade associated with the household head's main job or with other activity (e.g., a shoemaker who is also in charge of brandy distribution), labor income from a second occupation, and income derived from agro-pastoral activities on land that is rented from others. For additional details, see Nicolini and Ramos Palencia (2015).
- Urban: this is a dummy variable that takes the value 1 if the household is in the city of Palencia or Guadalajara (and 0 otherwise).

- Economic sector: the sector in which the household head's main income-generating activity is performed.[22]

5. Estimation and results

An unusual characteristic of our data sets is the coexistence of a *complete* set of a given area's available inventories and a *representative* set of household incomes in that area.[23] This characteristic makes it possible to identify the specific location (within the income distribution) of households with inventories and to discuss the resulting implications for whether surviving PIs are affected by selection bias. Before exploring that topic, however, we must take into account that this link—between probate inventories and the income distribution—is different in Palencia than in the other two regions. In Palencia, we have a representative sample of incomes of the entire province; that sample was built by collecting information from all the relevant geographic units in the province and then calculating the proper weights for each GU (cf. Nicolini and Ramos Palencia 2015). For each of the other two regions we collected information in only two areas, neither of which are statistically representative (overall) of their respective province. Hence any conclusion we draw

[22] When the occupation recorded in the EC was not indicative of a particular economic sector, we assumed that the focal household head worked in the primary (resp. tertiary) sector if more than half of that head's total income derives from rural (resp. urban) properties. After this procedure, there remained observations for which an economic sector could not be reliably assigned; these instances usually involve the poor, the disabled, and women. For more details on how households are assigned to an economic sector, see Nicolini and Ramos Palencia (2010).

[23] McCants (2007) uses a set of inventories from 912 poor to lower-middling citizen households in 18th-century Amsterdam. She assigns more than half of these households to a position in the income distribution by using only the monthly rent on their dwellings.

concerning these latter two regions applies only to those GUs about which we have information.

For the observations from Palencia, we can compare the distribution of inventories (recall that out PI data is complete for that region) with the income distribution among the related population. This latter distribution is approximated by first compiling a complete list of the households of one town in each GU and then using population-based weights (in column [11] of Table 1) to compensate for the quantity of households in the towns of each locality *not* being proportional to that GU's total population.[24] In the province of Palencia, only 2.6% of the inventories come from households in the first (lowest) income quintile and another 7.8% come from the second quintile; thus the 40% of households in the distribution's bottom part account for only 10.4% of all inventories, which means that the size of the selection bias is considerable. Despite this evident strong selection bias of PIs, our matched data reveal that households whose income is below the median are not completely absent from the records of wealth. In particular: the weighted median income in Palencia in DS2 is 698 *reales*, and in DS1 altogether 23 (19.8%) of Palentine households

[24] Two kinds of weightings are applied in this paper. The first one is based on population and is devised for the purpose of constructing, for each region, a corrected income distribution from the data in the EC. The bias that we seek to minimize here is that the sample size of each GU in each region is not proportional to that region's population. In Palencia, for example, Guardo-Cervera has 63 of the observations in DS2 (1.5% of the province's total) but comprises 9.3% of the households in Palencia. In this case we weight each observation in DS2 by the ratio of relative population to relative sample size (for more details, see Nicolini and Ramos Palencia 2015). The second weighting strategy is income based and is intended to construct a corrected wealth distribution using each GU's income distribution (obtained via the first strategy). The problem to be solved here is that wealthier households are over-represented in the surviving probate inventories; we would therefore like to give more weight (in the wealth distribution) to PIs coming from relatively poorer households. Section 5 explains this weighting strategy in more detail (when we present the econometric estimations).

have income below that level; the implication is that nearly a fifth of the PIs are from households whose income is below the median.[25]

As for *intragroup* selection bias in the province of Palencia, the weighted average income of households below the median in our DS2 is 450 *reales* whereas the average income of those households with PIs (our DS1) and income below 698 *reales* is 545.4 *reales*. If we look at the bottom of the distribution, the weighted mean income of the first quintile is 122.5 *reales* while the three surviving inventories in that group report respective incomes of 251, 266, and 333.5 *reales*.[26] So below the median, the average income of PI households is 21% higher than the population's average income; in the first quintile, the average income reported by PIs is 131% higher than the population's average. Thus the intragroup selection bias not only is large but also increases for lower parts of the distribution.

For each household in DS1 we have information on both income (from the EC) and wealth (from the PI). We can therefore examine, case by case, the relationship between income (when the Cadastre was produced) and wealth (when the household head died). In Figure 3, Graphs 1 and 2 plot this relationship when using variables given in (respectively) levels and logs. The association revealed in each graph—and especially the second—is remarkable.

[[insert Figure 3 about here]]

It is difficult here to infer a causal relationship between income and wealth because, on the one hand, wealth can be viewed as the accumulation of past income streams and, on the other hand, income is determined in part by the returns on wealth. In the case of

---

[25] Table 2 (appearing later in this section) gives the share of inventories in each income group for the 194 observations in DS1. In this case, the bottom 40% of the income distribution accounts for only 5.1% of the sample's inventories.

[26] In the first (lowest) income quintile, the upper limit is 360 *reales*. Thus the average income of the two households in our first quintile is closer to that quintile's upper limit than to its average.

societies such as the Ancien Régime, the tendency is to suppose that wealth is largely predetermined; because wealth does not change significantly within a generation, it would seem that wealth drives income more so than income drives wealth (Alfani 2010, p. 514). In any event, our econometric approach is more in the line of a descriptive aid to understanding the association between variables—and possibly to inferring how one variable's distribution is affected by that of some other variable. That is, we do not attempt to explain one variable's behavior by identifying exogenous variation in the other variable.

For the purpose of devising a simple theoretical framework in which to discuss the relationship between these two variables, we assume that the deceased household head's wealth equals that of an individual still thriving and that income is a function of wealth. We then hypothesize that

$$X_i = Y_i + r\,\mathrm{PW}_i\,, \tag{1}$$

$$\mathrm{TW}_i = N\,\mathrm{PW}_i + \mathrm{PW}_i\,. \tag{2}$$

In these equations, X_i is the income of household i, Y_i is the wage, TW_i is total wealth (as recorded in inventories), PW_i is productive wealth or wealth that produces a flow of economic returns (as with land, for instance), NPW_i is nonproductive wealth (e.g., durable consumption goods), and r is the average rate of return on productive wealth.

In the simplest case, where the rate of return on productive assets is the same for every household and where both wages and nonproductive wealth are zero, income will be a constant proportion of total wealth and so inequality of income or wealth (as measured by standard indices, such as the Gini index), will be equal.[27] However, the empirical evidence is that wealth inequality always exceeds income inequality. One possible reason for this difference is that wages are nonzero and also larger (as a share of total income) in

[27] One of the Gini index's useful properties is independence of scale, whereby the Gini coefficient does not change when all incomes within a given distribution increase by the same proportion.

the lower part of the distribution. In their model of changes in income inequality, Soltow and van Zanden (1998, pp. 49–54) employ a simplification of this hypothesis under which all households have the same labor income. Nicolini and Ramos Palencia (2015) use the Shorrocks (1982) decomposition to show that, for a sample of households in Palencia (circa 1750) labor incomes have a positive contribution to total income inequality and therefore they are positively correlated with total income (i.e. more affluent households have larger labour incomes than households in the bottom of the distribution). The positive correlation with total income means that also the labor income of relatively affluent households is greater than that of low-quartile households.

Another possible explanation for why wealth inequality is greater than income inequality is that nonproductive wealth features more prominently in the top part of the distribution—in view, for example, of the larger share of their wealth that affluent households hold in nonproductive assets (books, jewelry, etc.). From this observation it follows that, for various segments of the distribution, differences in wealth will be greater than differences in income. This account seems plausible enough but has not been explored in the literature.

After characterizing both income and wealth in terms of levels, one can estimate

$$X_i = a + \beta\, \mathrm{TW}_i + \gamma\, Z_i + e_i \qquad (3)$$

As before, X_i and TW_i denote household i's income and total wealth (respectively); here Z_i is a set of control variables linked to the focal household's observable economic characteristics (e.g., economic sector, place of residence). The estimated a would be the income level when wealth is equal to zero (in our framework, having labor income only), and the estimate β would be the rate of return on wealth. Because inventories incorporate total wealth, β is a downward-biased estimate of the rate of return on productive wealth.

We mentioned previously that the PIs available in archives (which are the base for our DS1) are not a random sample of the households in their geographic area because the more affluent households are over-represented. In order to correct for this selection bias, we first constructed a distribution of household income for the 11 GUs in our DS2.[28] Then, given that distribution, we calculated the percentage of households in our DS1 that belong to the corresponding income group in DS2; see Table 2. Suppose we consider the 194 observations of the three regions together; then there are only 4 households (out of those 194 with PIs) in the first quintile, which implies that the poorest 20% of households accounted for only 2.1% of the inventories. At the other extreme, the richest 10% of households accounted for 39.7% of the inventories; this means that households in the highest *decile* are 37.8 times more likely to be included in the collection of probate inventories than are households in the lowest *quintile*. Our econometric model addresses this problem by weighting each observation in DS1 with the following ratio: the percentage of household in its quintile divided by the percentage of inventories in that quintile in DS2.[29]

[[insert Table 2 about here]]

[28] Because the number of observations in each GU is not proportional to its population, we weight each observation in DS2 so as to make the relative size of each GU proportional to its relative population.

[29] This is the strategy suggested by Lindert (1981) and applied by previous researchers using the information from PIs to calculate average wealth or wealth inequality (Jones 1982; Lindert 1986; Roine and Waldenström 2015). Yet it is not clear that weights are necessary in our case because we are trying not to *estimate* those variables but only to *link* wealth and income. Selection bias under these circumstances would be a problem if, for instance, some variables—related to both income and wealth—were influential *only* in the top (or bottom) of the distribution. As it turns out, the econometric results presented in this section are not much affected when the estimations are instead based on unweighted regressions.

The results from our weighted regression of equation (3) are presented in Table 3. The value of the constant is 493 *reales*; this is the annual wage that those with hardly any wealth would earn and that many *jornaleros* (day laborers) did earn.[30]

[[insert Table 3 about here]]

The β-value of 0.029, which is equivalent to r in equation (1), means that an increase in wealth of 100 monetary units is associated with an increase in income of 2.9 monetary units. If one assumes that wealth is exogenous and that causality runs from wealth to income, then we can use this information to conclude (howsoever provisionally) that the average rate of return was 2.9%. In Castile during this period, the rates of return on financial investments range from 3 to 10%.[31] [32] The specification of equation (3) implies that wealth's elasticity is a function of its level.[33] The wealth elasticity of income at the means of our two variables is 0.347 (when all dummies are set equal to 0).[34]

Although analysis in terms of a linear relationship is useful, Figure 3's Graph 1 and the pattern of residuals both suggest that a linear specification is likely suboptimal in light of the data's concavity. If the regression incorporates a third-order polynomial in wealth, then the model's fit improves and the implied function is clearly increasing and concave. If the aim is to model nonlinear relationship between the variables, then a log-in-log

---

[30] Values more frequently imputed to the occupational category of *jornalero* are 360 *reales* or 480 *reales*.

[31] Official interest rates in Castile declined to 3% in 1705; the rate fell that low also in the former Crown of Aragon upon the *Real Pragmática* (Royal Decree) of July 6, 1750 (Yun Casalilla 1987, p. 357). There was no integrated public debt market for all Castile, so there was much diversity in nominal interest rates (ranging from 5% to 10%) offered by the *juros de alcabalas* in such cities as Burgos (which offered 11 different interest rates in the mid-18th century), Cadiz, and Murcia. See Álvarez-Nogal (2009, pp. 129–30).

[32] If we exclude the two outliers on the right in Graph 1, then the slope changes to 0.049.

[33] If $X_i = Y_i + r\mathrm{PW}_i$ then the elasticity of income is $\varepsilon_i = \beta[\mathrm{PW}_i/(Y_i + r\mathrm{PW}_i)]$, an increasing (and concave) function of PW_i.

[34] If the two outliers are excluded then the elasticity increases to 0.551.

specification is theoretically appealing on two counts:[35] (i) it uses the log of income, which is consistent with empirical observations documenting that the income distribution is well approximated by a log-normal distribution; and (ii) it presumes (implicitly) that the wealth elasticity of income is constant, which allows for straightforward comparisons of this variable across different estimations.[36] Furthermore, if we assume that wages are well described by a log-normal distribution, then we can use the approach advocated by Davidson and McKinnon (1981) to assess the relative merits of a semi-log approximation (with wages in logs and wealth in levels) and a log-log approximation (with both variables in logs). That test clearly rejects the semi-log in favor of the log-log approximation.[37] We therefore estimated the following alternative specification, in which the two main variables are logs:

$$x_i = a' + \beta' \, \mathrm{tw}_i + \gamma' Z_i + e'_i \, ; \qquad (4)$$

here x_i is the log of household i's income, tw_i is the log of that household's total wealth, and Z_i is a set of other household characteristics.

In this case, the parameter β' can be interpreted as the percentage change in income associated with each percentage-point increase in wealth. (We assume that this elasticity is constant for the different levels of income.) Results of the regression with variables in logs are reported in Table 4.

[[insert Table 4 about here]]

---

[35] Wooldrige (2002, p. 279) states that, "in many cases, using logarithms of certain variables and adding quadratics is sufficient for detecting many important nonlinear relationships in economics."

[36] If we take the correct econometric model to be log-in-log, where $x_i = a + \beta w_i^{\beta}$ for $x = \log(X)$ and $w = \log(W)$, then we are implicitly assuming that $X_i = \gamma r \mathrm{PW}_i^{\beta}$. The problem with this specification is that it implies income is zero whenever productive wealth is zero, which does not accord with the evidence.

[37] Results are available from the authors upon request.

Here the elasticity of 0.579 suggests that an increase of 1 percentage point in wealth is associated with an increase of slightly more than half a percentage point in income. Coefficients for the dummies indicating secondary and tertiary sectors are positive; this result suggests that, for a given level of wealth, income is higher for households whose head is working in those sectors. We can offer no empirically based explanation for this finding, yet it is consistent with the notion that a greater share of wealth is "productive" in the secondary and tertiary sectors (because e.g. rural aristocracy tends to accumulate wealth in non-productive status goods).[38]

Recall that we have observations from Palencia in the north of Spain, from Guadalajara and Las Vegas in the center, and from Granada in the south. As a result, we can check for the existence of systematic differences in the estimated relationship across regions. Table 5 presents results from our regression of specification (2) for each region.[39]

[[insert Table 5 about here]]

For the province of Palencia (Region 1), wealth elasticity is close to that for the whole sample. Also, the dummy variable for urban households is not significant (though negative for the whole sample) whereas the dummies for secondary and tertiary sectors are both positive and significant.

Elasticity is less in Region 2 than in Region 1, and the parameters associated with the secondary sector or the tertiary sector are positive. All the estimated parameters are statistically significant at the 5% level. The effect of being located in an urban context is negative and statistically significant in Region 2 but is not significant in Region 1, which

[38] An alternative explanation that may be more in line the data's characteristics is that wealth in these sectors is easier to conceal and so is systematically under-reported in PIs, which makes that wealth seem less productive. On the possibility of under-reporting during elaboration of a probate inventory, see Nicolini and Ramos Palencia (2010, pp. 156–59).

[39] In these "regional" regressions, the weights we use are specific to each region.

suggests that the dynamics of Guadalajara City (the only city in Region 2) differ from those of Palencia City (the only city in Region 1).

In Region 3, the only variable included in the regression is the log of wealth. The reason is that the Region 3 localities in our data set are mostly rural and so most household heads in this region are engaged in the primary sector: only one household head (a blacksmith) is engaged in a secondary activity; and only three—an apothecary, a carriage driver, and a *regidor perpetuo* (former mayor town)—are engaged in tertiary activities. Here our estimation of equation (2) yields a wealth elasticity that is greater than that in the other two regions; although the point estimate is less than 1, the confidence interval includes values that are greater than 1.[40] In this region, we cannot reject the hypothesis that income and wealth increase proportionally. A remarkable finding reported in Table 5 is that in Granada, for which we have but 29 observations, the association between our two variables is extremely high: the R-squared is 0.71 and the adjusted R-squared is 0.70.

Several aspects of our results merit additional comment. First, it is worth remarking the close association between our two variables, which are collected from completely independent sources and so provide alternative perspectives on household affluence. This association suggests that—notwithstanding any doubts about the accuracy of historical sources as regards the material well-being of past populations—these variables do a good job of capturing some underlying dimension of such well-being.

A second feature of our results is that, on average and in the two regions with more observations, a household's income increases more slowly than its wealth. Data limitations preclude a convincing empirical explanation for this pattern, but its existence suggests several hypotheses. For instance, one could posit that labor income (or, more generally,

[40] The difference in our results for Region 3 is not driven by most of that region's households being rural. When equation (2) is estimated only for the rural households of the other regions, elasticity increases slightly (as compared with the estimate for rural and urban households combined) but is still less than 1.

income unrelated to physical assets) is relatively larger in the bottom part of the distribution; if that were true then households with more wealth would derive relatively less of their income from labor, resulting in a less-than-proportional increase in income (vis-à-vis wealth). The three graphs in Figure 4 plot the share of personal income in total income (vertical axis) against total income (horizontal axis). These graphs show some households for which that proportion is 0 or 1 (the points marked on the horizontal lines corresponding to those *y*-axis values) and also confirm that a higher total income corresponds to a smaller proportion of labor income.[41] Graphs 2 and 3 indicate that this pattern is far more pronounced when the head of household is engaged mainly in agricultural activities than when the head is linked to activities in the secondary and tertiary sectors.

[[insert Figure 4 about here]]

Another explanation for the differential increase in household income versus household wealth is that the latter is more concentrated in income-producing assets in the bottom of the distribution but more concentrated in other kind of assets (status goods, luxury consumption, etc.) in the top of the distribution. Preliminary analysis of the share of different kinds of goods in PIs supports this account. As shown in Table 6, the share of real estate (land and buildings, which presumably qualify as income-generating assets) in total wealth is clearly greater in the first (lowest) quintile whereas the shares of perishable goods and money in cash (neither of which is likely to produce income) are greater in the upper two quintiles (i.e., 4 and 5). The top quintile features greater shares of "debts in favor" (financial assets that presumably yield some earned interest income) and shop assets

[41] The household for which this proportion *exceeded* 1 is that of Antonio de Laya, a lawyer with personal income of 3,300 *reales* but also with the "negative income" represented by 297 *reales* from *censos en contra*—the annual interest rate paid on financial liabilities. For more details, see Nicolini and Ramos Palencia (2015, p. 13).

(which we assume are related to income-generating activities in the tertiary sector) than do the distribution's other quintiles, but these differences are not large.

[[insert Table 6 about here]]

Finally, a third highlight of our findings is that the relation between wealth and income differs depending on whether the focal household's head is engaged in agriculture or instead works in the secondary or tertiary sector(s). We also find that urban households exhibit income/wealth ratios similar to those of rural households in some regions (viz., Palencia) but not in others (Guadalajara), where urban dwellers have less income per wealth than do rural dwellers. Such differences should be born in mind when comparing levels of economic inequality across sectors or regions, since some previous approaches have used real estate (or particular subsets of certain assets) to proxy for wealth or income (see e.g. van Zanden 1995; Alfani 2010).

6. Conclusions

Estimating economic inequality in pre-industrial economies requires that one make creative use of indirect data. The income distribution, which is usually reconstructed using social tables, is often imprecisely demarcated at the top because of widely varying income within upper-quintile occupational categories. Since the distribution of wealth is usually deduced from fiscal sources or probate inventories, it can be poorly estimated in the distribution's bottom part because of selection bias due to the general under representation of poor households in such records. In this paper we combine information on income and wealth for the same set of households, which helps shed light on how the distributions of these two variables are related in the context of a pre-industrial, medium-size, semi-urbanized Spanish population.

The method we use to identify the position of each household (in the data set of surviving probate inventories) within the distribution of all households (reconstructed

using the Ensenada Cadastre) makes it possible to assess the magnitude of selection bias.[42] Surviving PIs are a biased slice of society in that poorer households are clearly under-represented; even so, our PI data set does include some households from the bottom of the income distribution. Hence once should *not* accept the widespread but unfounded notion that households of below-median income or wealth are entirely absent from the archives of probate inventories.

We find also that the income assigned by the EC and the wealth registered in the PIs are closely associated not only within the whole sample of 194 inventories but also when the three regions that we consider (Palencia; Madrid and Guadalajara City; Granada) are analyzed separately. From this result we can conclude that, even though income inequality seems to be consistently less than wealth inequality, a given household's location in one distribution depends strongly on its location in the other.

Given the data set used in this paper, the best econometric specification for estimating the association between variables is one in which both income and wealth are stated in terms of their logarithms. Under this specification, the elasticity of income with respect to wealth varies between 0.4 and 0.9 (depending on the GU). These values imply that a 10% increase in the wealth of a household is associated with its income being from 4% to 9% higher.

Elasticity that is less than 1 is consistent with observations that wealth inequality is greater than income inequality. Our paper does not explain this relation between the distributions of those two variables in Modern Spain, but we do document some highly relevant results. In particular: labor income is slightly larger (relative to other income sources) in the bottom part of the distribution; at the same time, changes in relative importance of the different components of wealth (land, livestock, buildings, urban

[42] In this case, the "selection" is affected by (a) the choices and possibilities related to the 18th-century elaboration of a PI and (b) the *absence* of inventories that failed to survive until the present.

properties, financial assets, money, consumption goods) across different parts of the distribution indicate that the assets more clearly related to generating income (e.g., land) are more important in the bottom of the distribution.

Finally, the parameters associated with our *Secondary* and *Tertiary* dummy variables are positive and significant for the whole sample and also for two of the three regions estimated separately. This result suggests that, for a given level of wealth, households with a head who works in one of those sectors tend to have more income than households with a head who works in the primary sector. These systematic differences in the relationship between income and wealth across urbanization levels and economic sectors call for caution when using land or real estate as a proxy for wealth or when using wealth as a proxy for income. This is because, for example, a household whose head is engaged in manufacturing or trade will have more wealth than would be predicted by the value of its real estate and also more income than would be predicted by that household's wealth.

References

Alfani, G. "Wealth inequalities and population dynamics in early modern northern Italy." *Journal of Interdisciplinary History* 40 (2010): 513-49.

Alfani, G. "Economic inequality in Northwestern Italy: A long-term view (fourteenth to eighteenth centuries)." *The Journal of Economic History* 75, no. 4 (2015): 1058-1096.

Allen, R. C. "The Great Divergence in European wages and prices from the Middle Ages to the First World War." *Explorations in Economic History* 38 (2001): 411-47.

Álvarez-Nogal, C. *Oferta y demanda de deuda pública en Castilla. Juros de alcabalas (1540-1740).* Madrid: Banco de España, Estudios de Historia Económica, no. 55, 2009.

Álvarez-Nogal, C. and Prados de la Escosura, L. "The decline of Spain (1500–1850): conjectural estimates." *European Review of Economic History* 11 (2007): 319–66.

Álvarez-Nogal, C. and Prados de la Escosura, L. "The rise and fall of Spain (1270-1850)." *Economic History Review* 65 (2013): 1-37.

Atkinson, A., Piketty, T. and Saez, E. "Top Incomes in the Long Run of History." *National Bureau of Economic Research (NBER)* Woking Paper no. 15408. Cambridge: Massachusetts, October, 2009

Bertola, L., Castelnovo, C., Rodríguez, J., and Willebald, H. "Income Distribution in the Southern Cone during the First Globalization and Beyond." *International Journal of Comparative Sociology* 50, no. 5-6 (2009): 452-485.

Bourguignon, F. and Morrisson, C. "Inequality Among World Citizens: 1820-1992." *American Economic Review* 92, no. 4 (2002): 727-744.

Camarero Bullón, C. "El vecindario de la provincia de Palencia realizado en 1759 con datos del Catastro de Ensenada." Separata del Tomo III, volumen I, de las *Actas del II Congreso de Historia de Palencia,* (1990): 231-249.

Canbakal, J. "Wealth and Inequality in Ottoman Bursa, 1500-1840." Unpublished paper presented at the *Economic History Society Annual Conference*, York (United Kingdom), 5-7 April, 2013.

Comín, F. and Yun-Casalilla, B. "Spain: from composite monarchy to nation state, 1492–1914. An exceptional case?" In O'Brien, P. and Yun-Casalilla, B. (eds.): *The Rise of Fiscal States. A Global History, 1500–1914*. Cambridge: Cambridge UP, 2015.

Davidson, R. and MacKinnon, J. G. "Several Tests for Model Specification in the Presence of Alternative Hypotheses". *Econometrica* 49, no. 3, (1981): 781-93.

Dobado González, R. and García Montero, H. "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, no. 2 (2010): 253-277.

Ergene, B. A., Kaygun, A. and Cosgel, M. M. "A temporal analysis of wealth in eighteenth-century Ottoman Kastamonu." *Continuity and Change* 28, (2013): 1–26.

Ferrer i Alós, Ll. "¿Modernización fiscal? La implantación del Catastro en Cataluña." *CT/Catastro* 46, (2002): 27-36.

GRUPO 75. *La economía del antiguo régimen: la 'Renta Nacional' de la Corona de Castilla*. Madrid, 1977.

Herlihy, D. "The distribution of wealth in a Renaissance community: Florence 1427." In Abrams, P. and Wrigley, E. A. (eds.): *Towns in societies: essays in economic history and historical sociology*. Cambridge: Cambridge University Press, 1978, pp. 131-57.

Herlihy, D. and Klapisch-Zuber, C. *Tuscans and their families: a study of the Florentine Catasto of 1427*. Yale University Press: New Haven and London, 1985.

Herr, R. *España y la Revolución del siglo XVIII*. Madrid: Aguilar, 1988 (English edition *The Eighteenth Century Revolution in Spain*, Princeton, NJ: Princeton University Press, 1960).

Jones, A. H. *American Colonial Wealth. Documents and Methods*. New York: Arno Press, 1978.

Jones, A. H. *Wealth of a nation to be: The American colonies on the eve of the Revolution*. New York: Columbia University Press, 1980.

Jones, A. H. "Estimating Wealth of the Living from a Probate Sample." *Journal of Interdisciplinary History* 13, no. 2 (1982): 273-300.

La Force, J.C. "Royal Textile Factories in Spain, 1700-1800." *Journal of Economic History* 24, no. 3 (1964): 337-363.

Larruga, E. *Memorias políticas y económicas sobre los frutos, fábricas, comercio y minas de España*. Zaragoza: Institución Fernando el Católico, Instituto Aragonés de Fomento, vols. XXXII and XXXIII, 1995 (first edition 1787).

Lindert, P. "An algorithm for probate sampling." *Journal of Interdisciplinary History* 11, no. 4 (1981): 649-668.

Lindert, P. "Unequal English Wealth since 1670." *Journal of Political Economy* 94, no. 6 (1986): 1127-1162.

Main, G.L. "Personal wealth in colonial America: Explorations in the Use of Probate Records from Maryland and Massachusetts." *Journal of Economic History* 34, no. 1 (1974): 289-294.

Marcos Martín, A. *Economía, sociedad y pobreza en Castilla: Palencia, 1500-1814*. Palencia, 1985.

Matilla Tascón, A. *La única contribución y el catastro de la Ensenada*. Madrid: Ministerio de Hacienda, 1947.

McCants, A. "After-death inventories as a source for the study of material culture, economic well-being, and household formation among the poor of eighteenth-century Amsterdam." *Historical methods* 39, no. 1 (2006): 10-23.

McCants, A. "Inequality among the poor of eighteenth century Amsterdam." *Explorations in Economic History* 44, no. 1 (2007): 1-21.

Milanovic, B. *Worlds Apart: Measuring International and Global Inequality*. Princeton, NJ: Princeton University Press, 2005.

Milanovic, B., Lindert, P. H., and Williamson, J. G. "Ancient Inequality", revised version of "Measuring Ancient Inequality." *National Bureau of Economic Research (NBER)* Woking Paper no. 13550. Cambridge: Massachusetts, October, 2007.

Milanovic, B., Lindert, P. H., and Williamson, J. G. "Pre-industrial Inequality." *The Economic Journal* 121 (2011): 255–72.

Moatsos, M., Baten, J., Foldvari, P., van Leeuwen, B. and van Zanden, J. L. "Income inequality since 1820." In J. L. van Zanden, J. Baten, M. Mira d'Ercole, A. Rijpma, C. Smith and M. Timmer (eds.), *How Was Life? Global Well-being since 1820*, OECD Publishing, 2014, pp. 199-215.

Modalsli, J. "Inequality in the very long run: inferring inequality from data on social groups." *Journal of Economic Inequality* 13, no. 2 (2015): 225-247.

Nicolini, E. and Ramos Palencia, F. "A new method for estimating the Money demand in pre-industrial economies: probate inventories and Spain and the eighteenth century." *European Review of Economic History* 14, no. 1 (2010): 145-177.

Nicolini, E. and Ramos Palencia, F. "Decomposing income inequality in a backward pre-industrial economy: Old Castile (Spain) in the middle of the eighteenth century." *The Economic History Review*, article first published online: 10 SEP 2015 | DOI: 10.1111/ehr.12122, (2015).

Piketty, T., Postel-Vinay, G., and J.L. Rosenthal, J.L. "Wealth Concentration in a Developing Economy: Paris and France, 1807-1994." *American Economic Review* 96, no. 1 (2006): 236-256.

Piketty, T. and Saez, E. "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics* 118, no. 1 (2003): 1-39.

Ramos Palencia, F. *Pautas de Consumo y Mercado en Castilla, 1750-1850. Economía familiar en Palencia al final del Antiguo Régimen*. Madrid: Sílex, 2010.

Reis, J., Pereira, A., and Martins, C. ("How unequal were the Latins? The strange case of Portugal, 1550-1770." In *Wellbeing and inequality in the long run: measurement, history and ideas*. Scientific Meeting: Fundación Ramón Areces, Madrid, May 31 and June 1, 2012.

Roine, J. and Waldenström, D. "Long-run trends in the distribution of income and wealth." In Atkinson, A.B. and Bourguignon, F. (eds.): *Handbook of Income Distribution*, vol. 2A, North-Holland, Amsterdam, 2015.

Ruiz Torres, P. *Reformismo e Ilustración. Volumen 5*. In Fontana, J. and Villares, R. (eds.): *Historia de España*. Barcelona-Madrid: Crítica-Marcial Pons, 2008.

Ryckbosch, W. "Economic inequality and growth before the Industrial Revolution: the case of the Low Countries (fourteenth to nineteenth centuries)." *European Review of Economic History* 20, no. 1 (2016): 1-22 doi:10.1093/ereh/hev018.

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

Shorrocks, A. F. "Inequality decomposition by factor components." *Econometrica* 50, (1982): 193–211.

Smith, D. S. "Underregistration and bias in probate records: An analysis of data from Eighteenth-Century Hingham, Massachusetts." *William and Mary Quarterly* 32, no. 1 (1975): 100-110.

Soltow, L. and van Zanden, J. L. *Income and wealth inequality in the Netherlands, 16th to 20th century*. Amsterdam: Het Spinhuis, 1998.

van Zanden, J. L. 'Tracing the Beginning of the Kuznets Curve: Western Europe during the Early Modern Period." *Economic History Review* XLVIII, (1995): 643-664.

Willamson, J. (2002). Land, Labor, and Globalization in the Third World, 1870–1940. *Journal of Economic History* 62, no. 1 (2002): 55-85.

Williamson, J. and Lindert, P. *American Inequality. A Macroeconomic History*. New York: New York Academic Press, 1980.

Wooldrige, J. M. *Introductory Econometrics. A modern approach*. South-Western, 2002.

Yun Casalilla, B. *Sobre la transición al capitalismo en Castilla. Economía y Sociedad en Tierra de Campos 1500-1814*. Junta de Castilla y León, 1987.

PRIMARY SOURCES

Probate Inventories (section Protocolos Notariales)

Archivo Histórico de Protocolos de Madrid (years: 1753-1765)

Books: 29874, 29590, 29591, 29907, 29908 and 32550

Archivo Histórico Provincial de Granada (years: 1753-1760)

Books: 1097, 1127, 1117, 1118, 1119 and 1178.

Archivo Histórico Provincial de Guadalajara (years: 1754-1761)

Books: 919, 920, 921, 922, 930, 931, 932, 996, 996, 998 and 1002.

Archivo Histórico Provincial de Palencia (years: 1753-1768)

Books: 97, 98, 99, 379, 433, 434, 435, 642, 643, 675, 1000, 2656, 2733, 3731, 3732, 3793, 3795, 6277, 6278, 6962, 6964, 7814, 7816, 7817, 10617, 10618, 10619, 10834, 11383 and 11384.

Archivo Regional de la Comunidad de Madrid (years: 1755-1766)

(microfilms) MC 007268, MC 007270 and MC7271

Servicio de Reproducción de Documentos de Archivos Estatales (SRDAE), Madrid. Catastro de Ensenada.

Section (http://pares.mcu.es/Catastro/), Respuestas Particulares (Private Answers) from provinces of Granada, microfilms no. 0001-0071; Guadalajara, microfilms no. 0001-0154; Madrid, microfilms no. 0001-0031; Palencia, microfilms no. 0001–0170; and Toledo, microfilms no. 0001-0332.

TABLES

TABLE 1

GEOGRAPHIC DISTRIBUTION OF POPULATION AND WEIGHTS IN THE SOURCE DATA SET

Geographic unit [1]	Province [2]	Population [3]	Towns, villages, and *lugares* [4]	Average number of households [5]	Total number of households [6]	Probate inventories [7]	Towns surveyed [8]	Households sampled [9]	Average size of cities & towns in sample [10]	*Freq* [11]
Aguilar	Palencia	7,168	68	26	1,795	17	Valberzoso, Villabellaco	62	31	29
Baza	Granada	20,918	8	648	5,366	16	Cúllar Baza	678	678	8
Boedo and Ojeda Valleys	Palencia	9,484	45	53	2,385	10	Villabermudo	77	77	31
Cerrato	Palencia	19,372	41	105	4,313	12	Cevico Navero, Hontoria	201	101	21
Guadalajara City	Guadalajara	5,238	1	1,333	1,333	12	Guadalajara City	1,301	1,301	1
Guardo-Cervera	Palencia	11,000	49	48	2,372	21	Resoba	63	63	38
Las Vegas (excl. Aranjuez)	Madrid	23,904	22	284	6,401	37	Carabaña	182	182	35
Lecrín Valley	Granada	9,484	17	139	2,398	13	El Padul	258	258	9
Palencia City	Palencia	9,639	1	2,374	2,374	24	Palencia City	2,259	2,259	1
Saldaña-Valdavia	Palencia	3,652	29	36	1,044	0	Bustillo de la Vega	34	34	31
Tierra de Campos (excl. Palencia City)	Palencia	45,869	75	150	11,220	32	Paredes, Villarramiel	1,099	550	10
Data Set Total		165,728	356	5,196	41,001	194		6,214		
Castile Total		6,570,499			1,685,832					

Notes: In column [4], values reported for the province of Palencia are from Marcos Martín (1985, pp. 21–29) and those for the other provinces are from author calculations based on the EC; data for "Castile Total" (inhabitants c. 1752) are from GRUPO 75 (1977, p. 64). The size of a listed town may differ from its size in our data set because the number of household heads included in the *Libros de Cabeza de Familia* need not coincide with the quantity of households included in *Libros de Hacienda*, which is our source for information on individual households. The reason is because the *Libros de Hacienda* includes any household member—and not just the household head—who derived income from any kind of property and/or employment.

TABLE 2

WEIGHTS AND INCOME RANGES
IN PROBATE INVENTORIES, c. 1750–1770

				Income ranges	
Weights	Inventories	% Inventories	% Population	From	To
9.70	4	2.1	20	0.0	264.0
6.47	6	3.1	20	264.5	450.0
1.49	26	13.4	20	451.5	662.5
0.78	25	12.9	10	663.0	897.5
1.29	15	7.7	10	898.0	1,122.5
0.47	41	21.1	10	1,123.0	1,808.0
0.25	77	39.7	10	1,810.0	61,350.0
	194	100.0			

Source: Author calculations.

TABLE 3

LINEAR REGRESSION ESTIMATES. OLS

	Dependent variable:
	Income
Wealth	0.029***
	(0.003)
Secondary	426.604**
	(199.071)
Tertiary	734.746***
	(274.611)
Urban	-177.453
	(185.655)
Constant	493.616***
	(78.091)
R-squared	0.426
Adj R-squared	0.414
F-statistic	35.05
N	194

Source: Author calculations.
Note: Standard errors are reported in parentheses. Significant at *10%, **5%, ***1%.

TABLE 4

LOG REGRESSION ESTIMATES. OLS

	Dependent variable: *Log Income*
Log Wealth	0.579***
	(0.054)
Secondary	0.843***
	(0.154)
Tertiary	0.810***
	(0.212)
Urban	-0.448***
	(0.143)
Constant	1.169**
	(0.474)
R-squared	0.482
Adj R-squared	0.471
F-statistic	44.01
N	194

Source: Author calculations.
Note: Standard errors are reported in parentheses. Significant at *10%, **5%, ***1%.

TABLE 5

REGIONAL ANALYSIS OF LOG REGRESSION ESTIMATES. OLS

	Dependent variable: *Log Income*		
	Region 1 Province of Palencia	Region 2 Madrid and Guadalajara city	Region 3 Rural Granada
Log Wealth	0.553***	0.407***	0.914***
	(0.058)	(0.111)	(0.112)
Secondary	0.320**	1.411***	- - -
	(0.149)	(0.461)	
Tertiary	0.598***	1.609**	- - -
	(0.200)	(0.671)	
Urban	0.128	-0.746**	- - -
	(0.165)	(0.289)	
Constant	1.488***	2.759***	-1.389
	(0.518)	(0.971)	(0.945)
R-squared	0.501	0.497	0.712
Adj R-squared	0.483	0.451	0.701
F-statistic	27.91	10.88	66.63
N	116	49	29

Source: Author calculations.
Note: Standard errors are reported in parentheses. Significant at *10%, **5%, ***1%.

TABLE 6

PERCENTAGE OF ASSETS IN PROBATE INVENTORIES BY QUINTILE

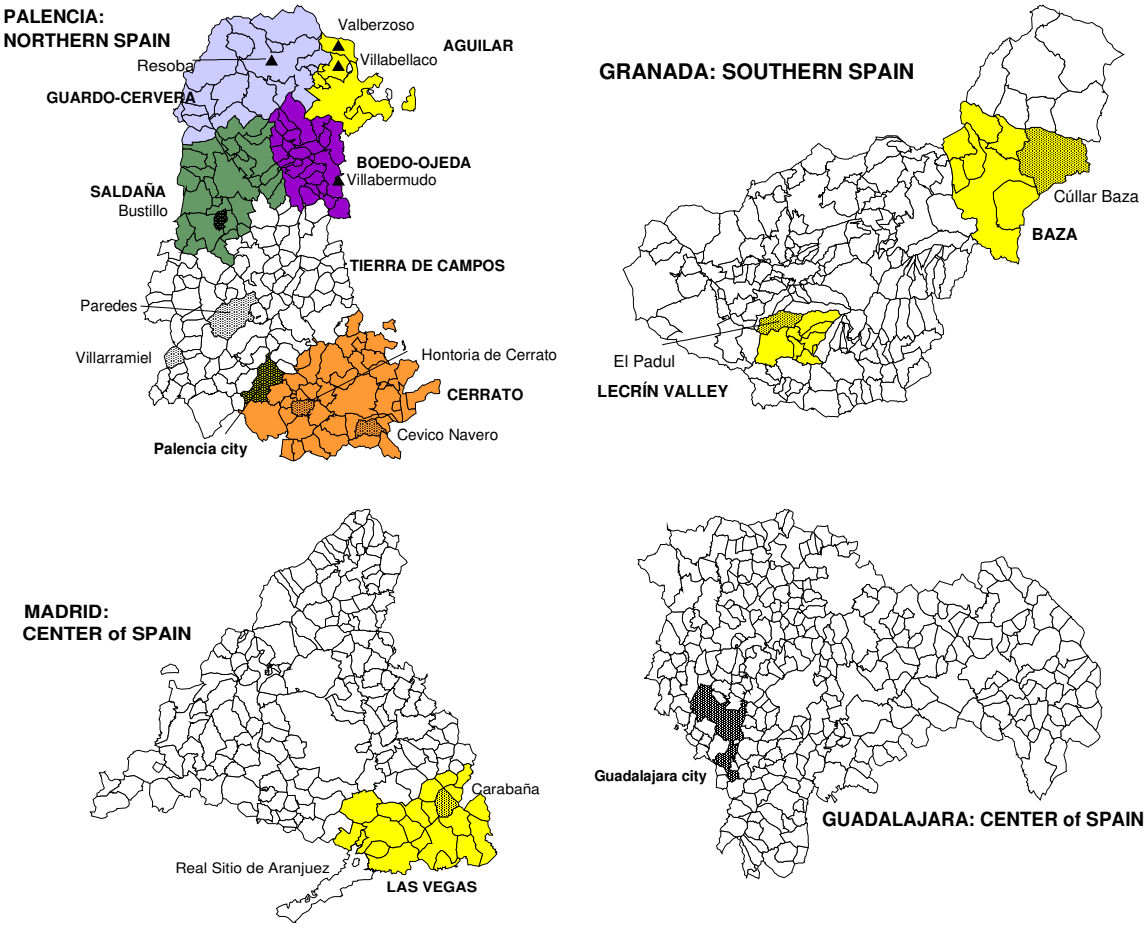
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Non-land properties	24.78	18.19	16.61	20.31	13.67
Rural properties	34.23	30.84	26.96	26.96	27.60
Money	1.38	2.96	1.79	2.65	4.63
Debts in favor (financial assets)	5.02	1.34	3.59	3.62	6.57
Inheritance in advance	4.02	5.59	9.08	7.77	4.17
Shop assets	0.00	0.97	0.23	0.00	4.55
Tools	3.38	7.18	8.53	5.09	4.72
Perishable goods	4.87	7.77	5.84	11.01	8.33
Raw textiles	2.18	7.30	5.68	3.15	5.61
Livestock	7.19	9.10	9.58	10.00	7.24
Consumption goods	12.93	8.86	12.10	9.44	12.92
Wealth (mean) in *reales*	6,560.08	10,397.93	14,765.52	20,992.48	44,416.04
N	24	23	23	23	23

Source: Author calculations of probate inventories from Archivo Histórico Provincial of Palencia.
Note: These percentages are calculated based only on 116 probate inventories from the province of Palencia; see Ramos Palencia (2010).

FIGURES

FIGURE 1

PALENCIA, GRANADA, MADRID, AND GUADALAJARA (1753–1768)



Geographic units	Area (km²)	Province's current boundaries
Urban areas		
Guadalajara City	151	Guadalajara
Palencia City	n.a.	Palencia
Rural areas (comarcas)		
Aguilar	476	Palencia
Lecrín Valley	461	Granada
Baza	1,732	Granada
Boedo and Ojeda Valleys	613	Palencia
Cerrato	1,389	Palencia
Cervera	858	Palencia
Las Vegas (excl. Aranjuez)	1,189	Madrid
Saldaña-Valdavia	347	Palencia
Tierra de Campos (incl. Palencia City)	2,171	Palencia

FIGURE 2

SPAIN, CIRCA 1750
(PALENCIA, MADRID, GUADALAJARA, AND GRANADA)

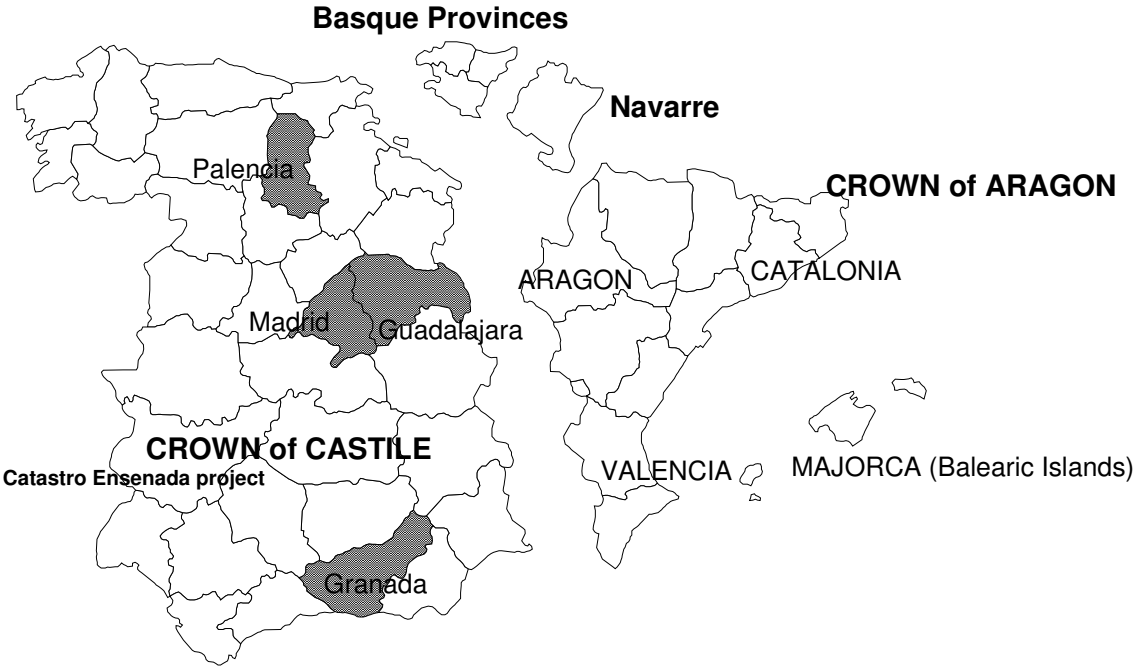
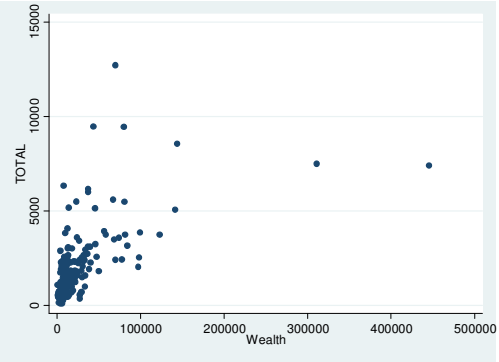


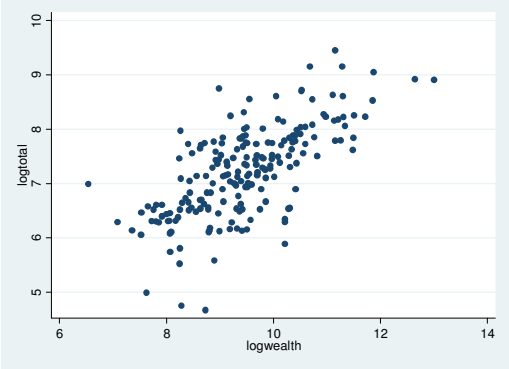
FIGURE 3

WEALTH AND TOTAL INCOME IN THE 194 INVENTORIES

Graph 1
Variables in levels



Graph 2
Variables in logarithms

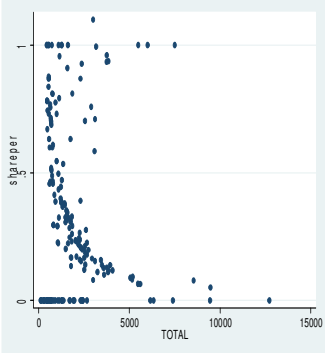


Source: Author calculations.

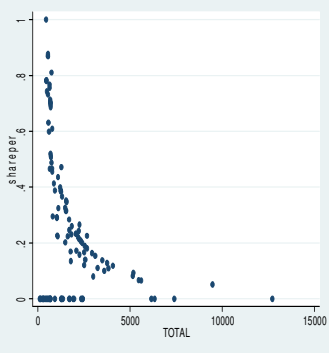
FIGURE 4

PERSONAL INCOME AS SHARE OF TOTAL INCOME

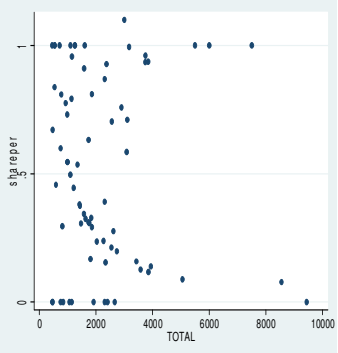
Graph 1: Total



Graph 2: Primary sector



Graph 3: Secondary and Tertiary sectors



Source: Author calculations.

European
Historical
Economics
Society

EHES Working Paper Series

Recent EHES Working Papers

2016

EHES.94 Reconstruction of annual money supply over the long run: The case of England, 1279-1870
Nuno Palma

EHES.93 World trade, 1800-1938: a new data-set
Giovanni Federico and Antonio Tena-Junguito

EHES.92 Capital shares and income inequality: Evidence from the long run
Erik Bengtsson and Daniel Waldenström

2015

EHES.91 The Rise of the Middle Class, Brazil (1839-1950)
María Gómez-León

EHES.90 Spanish Land Reform in the 1930s: Economic Necessity or Political Opportunism?
Juan Carmona, Joan R. Rosés and James Simpson

EHES.89 Risen from Chaos: What drove the spread of Mass Education in the early 20th century China
Pei Gao

EHES.88 A city of trades: Spanish and Italian Immigrants in Late Nineteenth Century Buenos Aires. Argentina
Leticia Arroyo Abad and Blanca Sánchez-Alonso

EHES.87 A closer look at the long-term patterns of regional income inequality in Spain: the poor stay poor (and stay together)
Daniel A. Tirado, Alfonso Díez-Minguela and Julio Martínez-Galarraga

All papers may be downloaded free of charge from: www.ehes.org
The European Historical Economics Society is concerned with advancing education in European economic history through study of European economies and economic history. The society is registered with the Charity Commissioners of England and Wales number: 1052680