Working Papers in **History & Economic Institutions**

May 2016 H&EI 16.01

Comparing Income and Wealth Inequality in Pre-Industrial economies. Lessons from Spain in the 18th century.

Esteban Nicolini (Universidad Carlos III de Madrid)
Fernando Ramos Palencia (Universidad Pablo de Olavide)

Keywords: inequality, income, wealth, Spain, probate inventories, Ensenada Cadastre.
JEL Classification Numbers: D31, N33, O15

Área de **Historia e Instituciones Económicas**
Dpto. Economía, Métodos Cuantitativos e Historia Económica
Economic History

Comparing Income and Wealth Inequality in Pre-Industrial economies. Lessons from Spain in the 18th century*

Esteban A. Nicolini[+]
Economics Department, Universidad Carlos III de Madrid – Spain
Universidad del Norte Santo Tomás de Aquino – Argentina
INVECO, Universidad Nacional de Tucumán - Argentina

Fernando Ramos Palencia[++]
Economic History Department, Universidad Pablo de Olavide – Sevilla, Spain

Abstract

Research on the history of inequality in pre-industrial economies has focused mainly on either wealth or income inequality. The most common problem with wealth inequality is the lack information about the bottom of the distribution while the main problem with income inequality is the lack of data to characterize the top of the distribution. Given that in general these approaches are based in different kinds of sources and methodologies, the results are not easy to compare and the links between the two distributions are difficult to establish. In this paper we use a unique data set for different regions of Spain circa 1750 and present results (the first for any pre-20th century economy) of inequality of both income and wealth for the same sample of households. Information of wealth comes from probate inventories while information of income comes from the Ensenada Cadastre. The main results of the paper are that poor households are not completely absent from our data set of inventories, that the position of a household in the distribution of income is closely associated to its position in the distribution of wealth and that an increase of a household's wealth is associated to a less-than-proportional increase in the household's income.

JEL Classification Numbers: D31, N33, O15

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

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

[+] E-mail for correspondence: esteban.nicolini@uc3m.es

[++] E-mail for correspondence: fcrampal@upo.es or fernando.ramos.palencia@gmail.com

Comparing Income and Wealth Inequality in Pre-Industrial economies. Lessons from Spain in the 18th century

1. Introduction

The significant increase in global income inequality in the last two hundred years is one of the most important features of the modern process of economic growth (Bourguignon and Morrison 2002, Van Zanden et al. 2014). One of the main hypotheses about the evolution of income inequality in Europe in the three hundred years before the Industrial Revolution is that there was a divergence of real wages across countries which is consistent with an increase in across country inequality (Allen 2001). Regarding inequality within countries, the evidence is mixed suggesting that inequality increased in some countries like Holland and Italy (Van Zanden 1995 and Alfani 2010 and 2014) while decreased in Portugal (Reis et al. 2012) or oscillated without a clear trend in Spain (Álvarez-Nogal and Prados de la Escosura 2013). Some of the estimations of economic inequality mentioned in this debate have used income as the relevant variable while others have used wealth.

One of the most usual approaches to measure income inequality in pre-industrial times have been to look at social tables in which population is divided into groups based on occupations and/or social classes and an average income is assigned to each group; in these cases, income is usually defined following indirect information from contemporaneous observers (Williamson and Lindert 1980, Milanovic, Lindert and Williamson 2007 and 2011).[1] This way of assessing households' income is relatively straightforward for the bottom part of the distribution because most of income comes from labor, occupations are relatively homogenous and workers' income is easy to infer. However, in the top part of the distribution there are many sources of income and occupations are quite idiosyncratic; hence, inferring incomes of more affluent households from an occupation or social class is subject to a wide margin of error.

Research on wealth inequality before the 20th century has mainly relied on data sets based on tax records (for instance, Soltow and Van Zanden 1998 and Alfani 2010) or collections of probate inventories (for instance, Hanson Jones 1978, McCants 2006 and Canbakal 2013). One limitation of these sources to measure wealth inequality is that in general the poorer segments of the society are underrepresented leading to an important

[1] The historical evolution of top incomes has also been scrutinized trough tax records (for instance, Atkinson, Pikkety and Saez 2009).

selection bias.[2] Some attempts have been made to calculate the level of underrepresentation of some population sub-groups and assess the implied selection bias for Colonial America (Jones 1978, Scott Smith 1975, Main 1974).

In this paper we present a new data set to calculate economic inequality in Spain using information from Palencia, Madrid, Guadalajara and Granada, circa 1750. This data set has some unique characteristics. In the first place we combine information from two different sources: probate inventories that contain a detailed description of households' wealth and a fiscal source –the Ensenada Cadaster- that provides information about households' income. Secondly, we have been able to link the households included in our set of inventories with their corresponding records in the Cadaster, opening the possibility to analyze the relationship between the income of a household when the Cadaster was produced and the wealth of the same household some years later when the head of the household passed away. Thirdly, given that the coverage of the Cadaster was basically complete, we can use the income distribution provided by this source to determine the magnitude of over (or under) representation of the different parts of the distribution in the set of inventories; a byproduct of this analysis is the possibility to weight the observations in the set of inventories in order to reduce the problem of selection bias.

The main goal of this paper is to provide a methodological contribution linking the distribution of income and the distribution and wealth and providing some hypotheses to understand the differences between them. Our results confirm that the set of surviving inventories is strongly biased towards the upward part of the distribution but in the three regions in our data set there are some inventories from the poorest quintile of the population showing that the poor are not completely absent from probate inventories. Even though the information of wealth and income come from completely independent sources, the association between the two variables is remarkable suggesting that they capture quite well some underlying dimension of economic affluence. A simple econometric estimation of the relationship between the two variables suggests that across households wealth increase more than income (i.e. wealth elasticity of income is smaller than one) and households whose head work in the secondary and tertiary sectors have larger incomes than the ones predicted by their wealth.

[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".

The rest of the paper is organized as follows: in section 2 we introduce a discussion of previous estimations of economic inequality; in section 3 we describe the historical context of Spanish Economy. In section 4 we present the data. Section 5 summarizes the main results and finally section 6 suggests some conclusions.

2. The estimation of income and wealth inequality before household surveys

The estimation of economic inequality before the 20th century has to be based on information collected for other purposes than inequality analyses. The three most usual approaches used in the field are the construction of social tables, the exploration of tax records and the analysis of probate inventories.[3] The so called "social tables" are based on dividing the population (or the subset of income earners) in groups (usually associated to an occupation or social status) and assigning an average income to each group. Under the assumption that most of inequality stems from the differences across groups rather than distances within the group, this methodology is similar in spirit to what is done with modern data sets when the population is divided in quintiles.[4] Once the profile of incomes for the population is constructed in this way, standard measures of inequality can be calculated (Lindert and Williamson 1980; Milanovic, Williamson and Lindert, 2007 and 2011; Bertola et al. 2009).[5]

Tax records are another important source to estimate inequality given that wealth and income are two variables which have emerged very naturally as tax bases in many historical contexts. An already classical example of these approaches can be found in the studies of the Florentine Catasto by Herlihy (1975) and Herlihy and Klapisch-Zuber (1985) which registered households' wealth. More recently Alfani (2010) explores wealth inequality in Ivrea in the 16th and 17th centuries based on the records of *estimi*, a tax on the value of

[3] Williamson (2002) has suggested that the ratio between average land rent and average unskilled wages can be used as a proxy for economic inequality given that rents from land are usually part of the income of the households in the top of the distribution while income from labor is more important for households in the bottom of the distribution. Other authors have modified this idea using per capita GDP instead of land rents (Álvarez-Nogal and Prados de la Escosura 2013; Dobado and García Montero, 2010).

[4] In modern data sets, households are ordered according to their income and therefore, by definition, the quintiles or deciles are non-overlapping sub-groups. In social tables the richest households of a group can be richer than the poorest households of the next group. Modalsli (2015) analyses the impact of this characteristics of the social tables in standards measures of inequality like the Gini coefficient.

[5] Milanovic, Lindert and Williamson (2007) calculate Gini coefficients with social tables for 14 pre-industrial societies; they include Old Castile in 1752 looking at groups of households with similar income and taking the information from the Ensenada Cadastre as summarized in Ramos Palencia (2010). The Gini coefficients they found for Modern Europe range from 44.9 in England and Wales in 1688 to 63.0 in Holland in 1732 and Old Castile is in the middle of this range with a Gini of 52.3.

real estate owned by households. Other approaches are based on other kinds of fiscal records: Soltow and, Soltow and Van Zanden (1998) use the introduction of an income tax in 1749 in the States of Overijssel to analyze inequality in pre-industrial Holland while Santiago Caballero (2011) uses the tithe paid by each grain producer as a proxy for income and from that information he deduces the inequality in Guadalajara, Spain, during the 18th century. Nicolini and Ramos Palencia (2015) use a sample of the Ensenada Cadastre in the province in Palencia collected with the purpose of transforming the fiscal system in Old Castile and providing information not only of households' incomes but also a detail of the different sources of income (mainly land, livestock an labor). Sometimes, this kind of sources is far from perfect and it is quite common that taxes were calculated based only on a sub-set of the assets of the households (usually land or real estate) or on a specific activity or kind of consumption which eventually was linked to households' income.[6]

Other important sources for the analysis of economic inequality in the past are the numerous collections of probate inventories scattered across the world.[7] Although PIs provide extremely rich and detailed descriptions of the wealth of many households in the past, their suitability for studying inequality is not obvious given the problem of selection biases (Lindert 1981, Hanson Jones 1982). Two biases are commonly identified: first, the age distribution of the deceased head of households is in general different from the age distribution of all head of households. Second, richer households are usually over-represented within the survival inventories. A recommended strategy to deal with selection bias is to construct weights or multipliers to correct the bias and that would be the approach in this paper.[8]

3. Historical context

[6] Soltow and van Zanden (1998, p. 26) analyze income inequality in the 16th century in Holland, using, as a proxy for households' income, the *tiende* penning (the tenth penny) a tax based on the rentable value of houses. Another example is Alfani (2010) who uses the value of real estate as a proxy for total wealth in Ivrea (Italy) in the Early Modern period.

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

[8] Lindert (1981, p. 660) says 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 we will see in Section 4, our weights will be elaborated using income as an attribute linked to wealth.

The period between 16th and 18th century in Europe is characterized by a process of widening distances in income per capita between a group of leading regions (England and Holland) and another group of regions with small or zero growth rates. This process has been called the small divergence (Allen 2001) and generated an increase in across country inequality.[9] In this small divergence, Spain was an example of relative retardation and between the 16 and the 18th centuries 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 in this period has expanded considerably in the last years. After the seminal paper by Van Zanden (1995), in which an increase in both income and wealth inequality has been documented for Holland and a positive association between economic inequality and economic growth and urbanization has been suggested for Europe, other researches have added new pieces of evidence for other countries. Alfani (2010) claims that wealth inequality in the North of Italy increased in the Early Modern times even though income per capita stagnated; in the Low Countries (Flanders, Brabant and Holland), Ryckbosch (2016) finds growth in economic inequality in the two centuries prior to the Industrial Revolution. Reis et al. (2012) suggest that income inequality do not increase in a Portuguese economy stagnates from 1550 to 1700. In the Ottoman Empire, Ergene et al. (2013) show declining inequality and economic stagnation during the eighteenth century. In addition, Canbakal (2013) finds inequality tended to be smaller in rural areas than in urban agglomerations. For Spain there are estimations by Santiago Caballero (2011) using tithe data as a proxy for income that suggest that inequality stable during the 18th century in central Castile and by Álvarez-Nogal and Prados de la Escosura (2013) that, using the ratio between per capita income and unskilled wages, claim that -except for the early seventeenth century- inequality decreased (increased) in periods of economic depression (expansion).

Even though the economic evolution of Modern Spain is characterized in general terms by a process of relative retardation, the eighteenth century is one of changes in almost in every aspect with a positive final balance: demographic growth, economic expansion, administrative reforms in the colonies and greater international presence and increasing geopolitical relevance. There were all the same a series of shadows and light that we need to bear in mind. The vacant throne of the Spanish monarchy triggered an international war - The War of the Spanish Succession (1701-1714) which pitted France against England, the

[9] Milanovic (2005) presents the concepts of global inequality and international inequality.

Netherlands and Austria. The war saw the beginning of British global hegemony and the arrival of the French Bourbons in Spain. When the conflict had ended, Phillip V (1700-1746) set out to emulate France by pushing for economic unification and political centralization, with mixed results. First of all, the diverse coinage in circulation in the different territories were withdrawn and replaced by a single currency (following the system in place in Castile). However, 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) given that they had mostly supported the opposing side to the Bourbons 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 the profits deriving from trade, industry and labour.[10] Conversely, the fiscal prerogatives and exemptions of the Basque Country and Navarre were preserved. In Castile, the fiscal reform advocated by the Marquis de Ensenada and inspired by the one applied in the Crown of Aragon 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 the *intendente* of the province. This *intendente* was the royal official in charge of tax collection in the province.

From the economic point of view, Herr (p. 128) highlights the fact 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]". With the exception of Madrid and its over 150,000 inhabitants (c. 1790) the

---

[10] According to Ferrer Alós (2002, pp. 29-32) the intention was to create a single tax calculated on the basis of personal wealth, which was derived from the land and from labour. However, the Bourbon bureaucracy, for logistical reasons, opted for a quota system. This meant that a more or less fixed amount was set and this was divided among the localities on the basis of the reports gathered in each of them.

[11] In an effort to stimulate industry and pay for the 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 industry in Catalonia (paper and cotton), in Valencia (silk, linen fabrics and tiles) and also in the Basque Country (iron and steel). In the interior, meanwhile, official policy was to support manufacturing with subsidies (for example, the textile products of Guadalajara) to compete with the luxury goods that were being imported from abroad. For its part, industries that produced essential goods were dominated by the Guilds; institutions that seem to have monopolized all industrial activity in the cities (with the exception of Catalonia) and of which the Bourbon politicians were highly critical.

[12] Although, according to the 1797 census, 22% of the people employed in agriculture were landowners, there were marked disparities from an imaginary line which extended from the north-east (Salamanca) to the south-east (Albacete). Thus, for instance, in the north - Aragon, Navarre, Biscayan and Galicia – landowners accounted for one half of the total, whereas in Andalusia the figure oscillated from 3 to 7%. Land ownership was distributed by 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 which were a royal privilege granted to the nobility and wealthy commoners); the towns and cities (commonweals that were a crucial

large cities were all 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, with the cities of Cadiz (over 10,000 inhabitants), Seville and Granada (80,000 each) and Malaga, with 50,000. In contrast, there was not a single city in Castile with a population of over 25.000. Based on the information provided by the Ensenada Cadastre and re-compiled by Matilla (1947), the provinces of the former Crown of Castile with the highest per capita income were Madrid (1,453 *reales*), Seville (641 *reales*) and Guadalajara (601 *reales*). In contrast, the provinces with the lowest per capita income were Granada (322 *reales*), Leon-Asturias (278 *reales*) and Galicia (202 *reales*). The average per capita income in Castile was in the region of 433 *reales*.[13]

In this article, we have chosen from North to South three provinces: Palencia, Guadalajara city and some small towns of Madrid, and Granada. Palencia -situated in the north of Spain- had a population of about 106,440 inhabitants distributed in the following *comarcas*: 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 eighteenth century the population of the province of Palencia was distributed irregularly, concentrating more than two thirds of the population in the southern: Tierra de Campos and El Cerrato Palentino. In these regions there were relatively large population centers, over a hundred neighbors, but geographically distant from each other. In Southern Palencia wheat production stood out, while the rest (wine, vegetables, etc.) played a secondary role, and also the livestock was limited to animals for agricultural work or for peddlers. Larruga wrote that Palencia was the "most industrious province of Castile". In fact, the *comarcas* of Tierra de Campos and El Cerrato Palentino had quite important secondary and tertiary sectors. In Northern Palencia (Guardo-Cervera and Aguilar), population density was quite low and the population was concentrated in a many small nuclei, very close to each other. These areas were characterized by livestock activities, linen production and mule drivers. Low quality textile industry (domestic production) was

---

source of revenue for the town and city councils); the Church; the Crown; and among private landowners with land that was not tied up in some way. It should be noted that only too often the landowner was the owner of a house and small piece of land. Furthermore, he or she might live on an estate to which rent had to be paid. In practical terms, the great aristocratic oligarchy was made up of the hidalgos and the urban super-rich. Indeed, in both provinces of Castile, in Extremadura and Western Andalusia and in Valencia, more than half of the towns paid tribute to landlords. For more details, see Herr (pp. 28-29).

[13] This is the author's own analysis based on the data provided by Matilla (1947, appendices) and the number of inhabitants in 1752 provided by the 75 Group (1977, p. 64). To work out the total income in each province, we have included the income generated both by the lay sector and the Church from rural properties, urban properties, livestock, ground rent and other forms of rent, interest on loans and all of the other revenue derived from industrial, commercial activities and personal work.

of great importance in the valleys of Boedo and La Ojeda. The most populated towns according to Ensenada Cadastre (1759) were: Palencia (9,639 inhabitants) and Paredes de Nava (3,395) both from *Tierra de Campos*.[14]

The areas (*comarcas*) from the center of Castile analyzed in this paper are Las Vegas, very close to Madrid City, and Guadalajara City. With the exception of Guadalajara City, the *economy was predominantly agro-pastoral* producing mainly cereals, vegetables, oil, wine, linen, silk, fruit and also sheep, goats, cattle and pigs. 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) in 1719. This company was characterized by state ownership, organization of production around guilds, competitive markets, and finally recruitment of foreign experts (La Force, 1964). Unfortunately the company began to decline during the 1790s due to the crisis of the Ancien Régime. The process speeded up during the Napoleonic Wars and finished with the closure of the company in 1822. The next most populated towns were Colmenar de Oreja (1,279 neighbors) and Chinchón (1,217 neighbors).

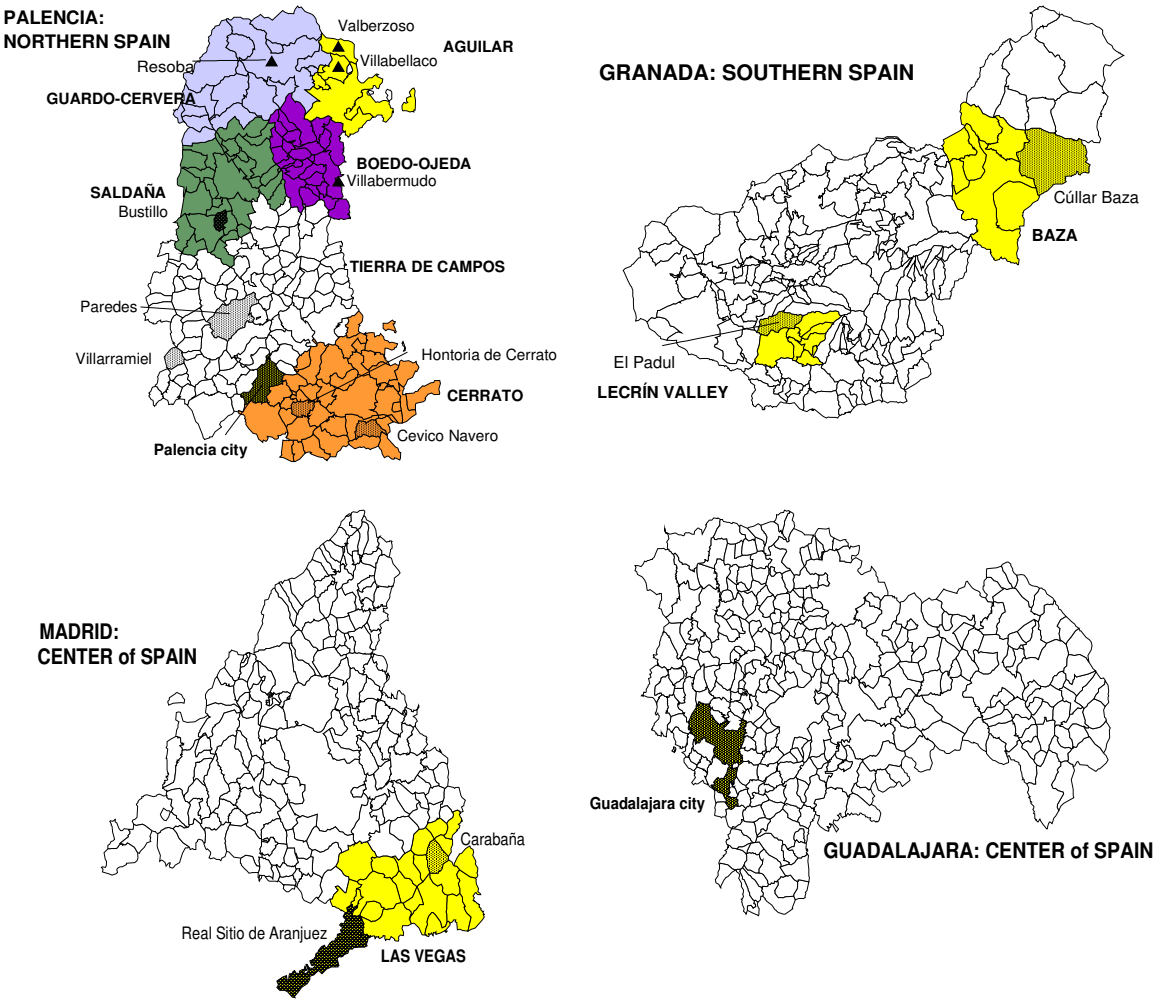
Finally in Granada we have studied two areas: Lecrín Valley in the southwest and Baza in the northeast. Lecrín Valley (2,398 neighbours, approximately 9,484 inhabitants) was basically an agricultural economy.[15] The most populated town was Albuñuelas (294 neighbours) and Pinos (260). Baza is the largest area (greater than 1,700 km2) of the province of Granada. There are three main zones: the Sierra of Baza, the Meseta and the Vega. The Meseta and the Vega are flat plains surrounding the city of Baza, mainly dedicated to agricultural uses. The Sierra of Baza is a rocky massif with deep valleys and escarpments. According to Ensenada Cadastre, Baza had a population of around 5,366 neighbors (approximately greater than 20,000 inhabitants). The largest towns were Baza (1,610 neighbors) and Las Cuevas (1,302 neighbors).

[14] See Marcos (1985, p. 22) and Camarero Bullón (1990, pp. 231-249).

[15] Many households earned some income apart from the one associated to the main job of the head of the household. Some of them were engaged in the manufacture and sale of "pleita" (ring or strip of straw twisted in several branches, sewn "pleita" was used to make mats, hats, pouches, etc.) and others were salesmen, mule drivers and peddlers, who provided a link between the very poor villages from these areas and Granada city.

FIGURE 1

GRANADA, GUADALAJARA, MADRID AND PALENCIA, 1753-1768.

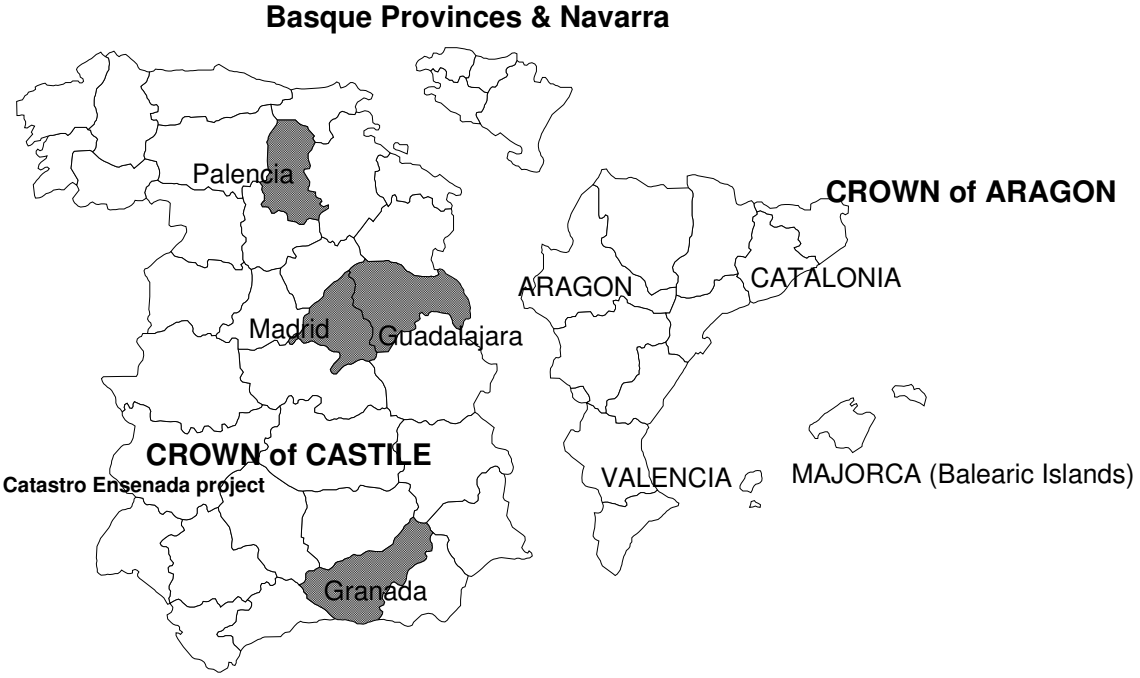


GUs	Area (km²)	Province current boundaries
▪ Urban areas:		
Guadalajara city	151	Guadalajara
Palencia city	n.a.	Palencia
• Rural areas (*Comarcas*):		
Aguilar	476	Palencia
Lecrín Valley	461	Granada
Baza	1,732	Granada
Boedo and Ojeda Valleys	613	Palencia
Cerrato	1,389	Palencia
Cervera	858	Palencia
Las Vegas (excluded Aranjuez)	1,189	Madrid
Saldaña-Valdavia	347	Palencia
Tierra de Campos (included Palencia city)	2,171	Palencia

4. The data

Two different data sets are used in this paper. The first one includes information contained in 194 probate inventories (PIs hereafter) and the second one is composed by more than 6,000 records on household characteristics from the Ensenada Cadaster from the same areas than the inventories. These areas are situated in the current provinces of Palencia (north), Guadalajara (center), Madrid (center) and Granada (south).

FIGURE 2
SPAIN, CIRCA 1750
(GUADALAJARA, GRANADA, MADRID AND PALENCIA FROM FORMER CROWN OF CASTILE)



A probate inventory is a comprehensive list of all the goods owned by the deceased at the moment of his/her death and it was usually elaborated by a notary or judicial authority a few days after the person passed away. Although there is some variability in the structure, the format and the style of the inventories, it is possible to summarize the structure of a Castilian post-mortem inventory, from the middle of the 18th Century. The rich description given in the post-mortem inventories remained in Castile until around the middle of the 19th century. After this date, the descriptions of durable and semi-durable goods (personal clothing, property and other household objects, amongst other things)

progressively disappeared, their total value becoming an increasingly smaller percentage in relation to the total inventoried assets. [16]

The Ensenada Cadastre (EC hereafter) is a census carried on in the middle of the 18th century with the purpose to improve the fiscal organization of the Spanish monarchy.[17] The purpose of the Marquis of La Ensenada (Secretary of the Treasury between 1743 and 1754) was to establish a single tax (*Única Contribución*) -universal and proportional to the income of the taxpayers-, which would replace other taxes collected from the provinces (*Rentas Provinciales*). The "*Alcabalas*", "Hundreds" (*Cientos*) and "Millions" (*Millones*), among others, were indirect taxes and they were supposed to have a regressive impact on the overall distribution of income in the economy (Comín and Yun, 2015). The tax reform planned by Ensenada proposed the direct taxation, the recovery of income previously transferred, the option of eliminating tax exemptions (especially in the ecclesiastical institutions) and the development of a simplified tax system. The downfall of Ensenada by pressure from the British lobby of the Court in 1754 -because of his sympathy towards France in The Seven Years' War, 1756-63- tax reform greatly reduced the political support for the tax reform which was never implemented but generated a huge 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 approximately between 1749 and 1759 and is a very good source for studying economic activities in general and economic inequality in particular because it collected a detailed accounting of the income of each household in each locality in Old Castile and incorporated an analysis of incomes of usually exempted social groups like the members of the church and the nobility. [18]

The probate inventories in our data set are all the available inventories between the years 1753 and 1768 from 11 geographic units (each geographic unit will be called GU hereafter) in three different regions of Castile.[19] The geographic coverage of this data set is the result of two motivations: initially we collected all the available inventories in the province of Palencia to analyze several economic topics in that province. There were 116 inventories from 7 different GUs (Palencia City, Boedo and Ojeda Valleys, Cerrato,

[16] See a summary and historiographical references on the characteristics of probate inventories as historical sources in Old Castile in Nicolini and Ramos (2010, p. 153-155).
[17] Ruiz Torres (2008, p. 280-285.)
[18] See Nicolini and Ramos Palencia (2015) for a detailed description of the EC as a source for studying income inequality.
[19] The great challenge -and at the same time one of the major difficulties of this research- is to find probate inventories whose heads of household are included in the Ensenada Cadastre. Unfortunately it is not an easy task mainly because only a limited number of probate inventories in each locality have come down to us in the archives.

Guardo-Cervera, Saldaña Tierra de Campos and Aguilar). Afterwards, we incorporated two other regions, one in the Centre of Old Castile (close to the city of Madrid) and another one in the South (in the province of Granada) in order to (i) expand the number of observations and the geographical coverage of the first data set and (ii) to incorporate into the analysis localities with a productive profile different to the one observed in Palencia. In these two cases, we selected two GUs in each region (Guadalajara City and Las Vegas in the Centre and Baza and Lecrín Valley in Granada) and collected all the available inventories in each of these four geographic units: 50 inventories in the Centre and 30 inventories in Granada. These makes the 194 PIs that we included in our Data Set 1 (DS1 hereafter). [20]

A very important difference between the 116 inventories in Palencia and the 78 inventories in the Centre and Granada is that in Palencia we have a full coverage of the available inventories in the Province and we have inventories for all the six natural traditional demarcations in the province. In each of the other two regions we have collected all the available PIs from two GUs but these GUs are not necessarily representative of the whole province (see the maps in figure 1).

To complete DS1 we have linked the name of the deceased person in each PI to the corresponding record in the EC; in this way we produced information for each household on -among other things- total wealth at the moment of the death of the head of the household (from PI) and the total income of the household as recorded by the EC some years before.

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

TABLE 1

GEOGRAPHIC DISTRIBUTION OF POPULATION AND WEIGHTS IN THE SOURCE DATA SET

Geographic Unit (GU)	Province	Population	Towns, villages and *lugares*	Average number households	Households	Probate Inventories	Towns surveyed	Households sampled	Average size of city or towns in sample	*Freq*
• Aguilar	Palencia	7,168	68	26	1,795	17	Valberzoso and 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 and 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	Palencia	3,652	29	36	1,044	0	Bustillo de la Vega	34	34	31
• Tierra Campos (excl. Palencia city)	Palencia	45,869	75	150	11,220	32	Paredes and Villarramiel	1,099	550	10
TOTAL Data Set		165,728	356	5,196	41,001	194		6,214		
TOTAL CASTILE		6,570,499			1,685,832					

Notes: Values reported in the third data column for the province of Palencia are from Marcos Martín (1985, pp. 21–29); for the rest of the provinces, authors's calculations from the EC; the data for total Castile (inhabitants c. 1752) are from Grupo 75 (1977, p. 64). The size of the listed towns differs from their 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 that widows are counted as 0.5 and also (and mainly) because the *Libros de Hacienda* includes anyone—not always a head of household—who derived income from any kind of property and/or employment.

The number of observations in DS1 is 194 and the geographic distribution of these observations is described in column 7 of Table 1. The main variables coming from the PI in our DS1 are as follows:

- Wealth: the sum of real estate assets (total of urban and rural properties), financial assets (cash, credits, debts, land rents, expenditures for the funeral service and shares of the estate received by the inheritors in advance), capital assets (implements and tools -farming implements, winemaking and measuring equipment, implements for livestock and for producing textiles-, 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 equipment -kitchen and furniture-, pictures, books, and jewellery). More details in Nicolini and Ramos (2010).
- Year: from 1753 to 1768.

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

It is well known that the households described by the surviving PIs available in archives are usually a biased selection of total households of the underlying population: households with high income or wealth are over-represented in the sample of PIs. In order to have an approximation of the whole distribution of households in each of the 11 GUs we have selected one (or two) towns in each GU and recorded all the relevant information provided by the EC for every household in those towns; this information is systematized in our Data Set 2 (DS2) and involves 6,214 households. In this way, we have been able to generate an approximation of the whole distribution of income for each GU. In Table 1 we show the basic information behind this process of reconstructing the income distribution in each GU. In the first three columns we present the 11 GUs included in our data sets, the provinces in which each of them is located and the population of the province. Columns 4 to 6 show the number of towns in each province, the average number of households in each town and total number of households in the corresponding GU. Column 7 reports the number of PI in each locality while columns 8 and 9 show the towns included in our DS2 in each GU and the quantity of households in each town.

All the information in DS2 set comes from the EC and the main variables used in this paper are as follows:

- Income: It is measured in *reales*. It includes income derived from land; buildings and non-land properties (e.g., houses in the cities and mills in the countryside); livestock; taxes, fees, credits and/or debts; and finally, personal earnings. This last part includes labour income of the main activity of the household head (this income is imputed by census officials and assuming a daily income and a certain number of days per year: 120 days for agricultural labourers, 180 days for workers in secondary and tertiary sectors, and 360 days for shepherds); additional income obtained from trade associated with the main job or with other activity (e.g. shoemaker is also in charge of the distribution of brandy); labour income from a second occupation; and income derived from agro-pastoral activities on land that is rented from others. More details in Nicolini and Ramos (2015).

- Urban: it is a dummy variable with value 1 if the household is in the cities of Palencia and Guadalajara and 0 otherwise.

- Economic Sector: we have distinguished the economic sector at which the head of the household performs his or her activity. [21]

5. Estimation and Results

An unusual characteristic of our data sets is the coexistence of a set of all the available inventories in a given area and a representative set of all the household incomes in that area.[22] This characteristic opens the possibility to discuss the specific location of the households with inventories in the whole distribution of income and its implication for the selection biases affecting the surviving PIs. Before advancing in that topic, it is necessary to take into account that this link between the inventories and the whole distribution of income in Palencia is different from the one in the other two regions. In Palencia we have a representative sample of incomes of the whole province built by collecting information

[21] When the occupation recorded in the EC was not indicative of the economic sector, we have assumed that if more than 50% of the total income comes from rural properties they belong to primary sector while when more than 50% of their total income comes from urban properties they belong to tertiary sector. After this procedure, there are still observations for which it was impossible to define an economic sector; the typical categories here are poor, disabled and women. For more details about the way in which the economic sector is assigned to each household see Nicolini and Ramos (2010).

[22] McCants (2007) uses a set of inventories from 912 poor to lower-middling citizen households from the city of Amsterdam and she is able to link the position of over half of the households in the whole income distribution using information on the monthly rental cost of their dwellings.

from all the relevant geographic units in the province and calculating the proper weights for each GU (see Nicolini and Ramos 2015). For each of the other two regions we have collected information in only two areas and they are not statistically representative of the whole provinces (and they are not intended to be). When we draw some conclusion of what happens in regions 2 and 3, we are talking only about what happens in those GUs for what we have information.

Taking the observations from Palencia, we can compare the distribution of inventories (recall that they are all the available inventories of that particular geographic region) with the distribution of income of the related population (which is the complete list of households in a subset of localities which are weighted in order to obtain an approximation of the income distribution of the whole province). In the province of Palencia, only 2.6% of the inventories come from households in the first quintile and another 7.8 % come from the second quintile implying that the 40% in the bottom of the distribution contributes with only 10.4% of the total of inventories: the size of the selection bias is considerable. However, and despite the quite large selection bias of PIs, our matched data shows that people below the median income are not completely absent from the records of wealth; the weighted median income in DS2 is 698 *reales* and 23 households in that province in DS1 (19.8%) have income below that level implying that almost a fifth of the PIs come from households whose income is below the median.[23]

Regarding the intra-group selection bias, and still focusing in the province of Palencia, the weighted average income of households below the median in our DS2 is 450 *reales* while the average income of those households with PIs (our DS1) and income below 698 *reales* is 545.4. If we look at the bottom of the distribution, the weighted mean income of the first quintile is 122.5, while the three survival inventories in that group have incomes of 251 and 266 and 333.5 *reales*.[24] So below the median, the average income of the PIs is 21 % higher than the average income in the whole population while in the first quintile the average income of the PIs is a 131 % higher than the average income in the whole population. The intra-group selection bias is not only large but it also increases as long as the focus moves to the bottom part of the distribution.

[23] The share of inventories in each income group for the 194 observations in DS1 can be observed in table 1. In this case the bottom 40 % of the income distribution represents only the 5.1 % of the inventories of the sample.

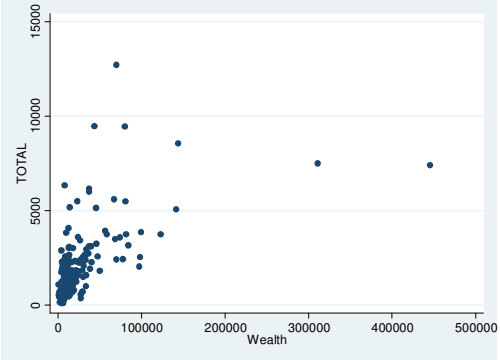
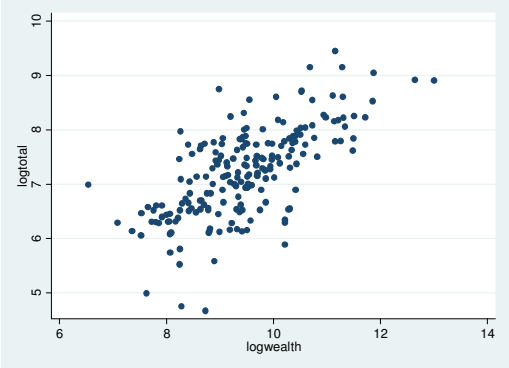
[24] The upper limit income of the first quintile is 360 *reales*. So the average income of the two households in the first quintile is closer to the upper limit than to the average income of the quintile.

Another characteristic of the comparison of the two distributions is that the wealth of those very close to the bottom of the distribution is not negligible in terms of their own income: in Palencia the ratio wealth-to-income of households in the first two quintiles (14.16) is actually *larger* than that ratio in the whole distribution of the province of Palencia (11.39). If the three GU are analyzed together (without weights) this pattern is even stronger, the picture is that in the first two quintiles the ratio wealth-to-income is 30.46 and in the whole distribution is 12.67. We have to be very cautious in drawing conclusions from this regularity because this hypothesis can be strongly affected by within-group selection bias: it is highly probable that, in the bottom of the distribution, only those households with relatively large wealth produce an inventory and therefore the average wealth of the inventories in that group is larger than the average wealth of all the households in that group.

Each observations in our DS1 has information on both income (from the EC) and wealth (from the PI) so it is possible to explore, on an individual basis, the relationship between income when the Cadaster was produced and wealth in the moment of the death of the head of the household. Graphs 1 shows this relationship with the variables in levels while in Graph 2 these variables are in logs. The association between them is quite remarkable, in particular in the second graph.

FIGURE 3

WEALTH AND TOTAL INCOME IN THE 194 INVENTORIES

Graph 1 Variables in levels	Graph 2 Variables in logarithms
	

Source: Authors' calculations

Trying to infer a causal relationship between income and wealth in this case is complex because on the one hand wealth can be regarded as the accumulated result of past streams of income and on the other hand income is determined, at least partially, by the returns of wealth. In a society like the *Ancien Regime* we tend to think that wealth is rather

predetermined, it does not change significantly within a generation and therefore the direction of causality would go more from wealth to income than from income to wealth (Alfani 2010, p. 514). Anyway, we present our econometric approach more as a descriptive device to understand the association between the variables and the possibility to infer what happens with the distribution of one of the variables from the information about the other variable rather than an attempt to explain the behavior of one variable through identification of an exogenous variation of the other variable.

In order to have a simple theoretical framework to discuss the relationship between the two variables, we can assume that wealth of the deceased is equal to the death of the living and that income is a function of wealth, we can hypothesize that

$$X_i = Y_i + rPW_i \quad (1)$$

$$TW_i = NPW_i + PW_i \quad (2)$$

where X_i is income of household i, Y is wage, r is the rate of return TW is total wealth (as recorded in inventories), PW is productive wealth or wealth that produces a flow of economic returns (land for instance), NPW is non-productive wealth (like durable consumption goods) and r is the average rate of return of the productive wealth.

In the simplest case in which the rate of return of productive assets is the same for every household and wages and non-productive wealth are zero, income will be a proportion of total wealth and inequality of income and wealth, measured by the standards indices (Gini among them), will be equal.[25] Empirically, however, wealth inequality is always larger than income inequality and this could be due to several facts. One of them is that wages are non-zero and larger, in relative terms, in the lower part of the distribution. A simplification of this hypothesis is that all the households have the same labour incomes and is presented by Soltow and Van Zanden in their model for changes in income inequality (Soltow and Van Zanden 1998, pp. 49-54). Nicolini and Ramos (2015), using the Shorrocks decomposition (Shorrocks 1982) show that, in sample of households in Palencia in circa 1750, labour incomes have a positive contribution to total income inequality and therefore they are not equal across the distribution and they are positively correlated with total income (i.e. more affluent households have larger labour incomes than households in the bottom of the distribution).

[25] A property of the Gini index is the independence of scale which implies that if all the incomes in a given distribution increase by the same proportion, the Gini does not change.

Another possibility is that non-productive wealth is relatively larger in the top part of the distribution, i.e. more affluent households have a larger share of their wealth in non-productive assets (books, jewellery, etc.); in this case, differences in wealth across different parts of the distribution will be larger than differences in income. This explanation, although plausible has not been explored in the literature.

If we put both income and wealth in levels and estimate

$$X_i = \alpha + \beta TW_i + \gamma Z_i + e_i \quad (3)$$

where X_i and TW_i are income and total wealth respectively while X_i is a set of control variables linked to observable economic characteristics of the household like economic sector or place of residence. The estimated α would be the income level when wealth is equal to zero (labor income in our framework above) and β would be the rate of return of wealth; in this case, given that inventories incorporate total wealth, β would be a downward biased estimation of the rate of return of productive wealth.

As we have already discussed above in this section, the available PIs in archives (which are the base for our DS1) are not a random sample of the households in a particular geographic area because more affluent households are over-represented. In order to correct this selection bias we have constructed a distribution of households' incomes for the 11 GUs in our DS2[26] and, whence we had the whole distribution of incomes for our DS2, we have calculated the percentage of households in our DS1 that belong to each income group in DS2 (see Table 2). For instance, if we consider the 194 observations of the three regions together, there are only 10 households (out of the 194 with inventories) belonging to the first quintile. It implies that the poorest 20% of households left only the 2.1% of inventories. In the other extreme, the richest 10% of households left 39.70 % of the inventories. It means that households in the first decile are 37.8 times more probable of being included in the collection of probate inventories than people in the first quintile. In our econometric model, we will take into account this problem weighting each observation in DS1 with the ratio between the percentage of households in that quintile (20%) and the percentage of inventories in that quintile in DS2.[27]

[26] Given that the number of observations in each GU is not proportional to its population, we have weighted each observation in DS2 to make the relative size of each GU in our data set proportional to its relative population.

[27] This is the strategy suggested by Lindert (1981) and applied by previous researchers using the information form PIs to calculate average wealth or wealth inequality (Hanson Jones 1982, Lindert 1986, Roine and Waldenström 2015). In our case, weights are less obviously necessary because we are not trying to estimate

TABLE 2

WEIGHTS AND INCOME LIMITS IN PROBATE INVENTORIES, c. 1750-70

Weights	Inventories	% Inventories	% Population	Income limits	
9.70	4	2.1	20		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	1122.5
0.47	41	21.1	10	1123.0	1808.0
0.25	77	39.7	10	1810.0	61350.0
	194	100.0			

Source: Authors' calculations

The results of the weighted regression of equation 3 are presented in table 3. The value of the constant is 493 *reales* suggesting that those with wealth close to zero would earn that annual wage which is what many *jornaleros* (day-labourers) earn.[28]

TABLE 3

LINEAR REGRESSION ESTIMATES

	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

Note: Standard errors in parenthesis. Variables are significant at 10% (*), 5% (**) or 1% (***) level.
Source: Authors' calculations.

The value of beta of 0.029 (equivalent to r in equation 1), means that an increase of wealth of 100 monetary units is associated to an increase of income of 2.9 monetary units. Assuming that wealth is exogenous and causality runs from wealth to income, can be interpreted as an average rate of return of 2.9%. Rates of return of financial investment in

those indicators but to link wealth and income; selection bias in our case would be a problem if, for instance, there are influential variables –in the relationship between income and wealth- in the top or in the bottom of the distributions. Actually, the econometric results presented in the remaining part of this section do not change very much when the estimations are based on un-weighted regressions.

[28] The values more frequently imputed to the occupational category of *jornalero* were 360 and 480 *reales*.

this period in Castile were around 3%.[29] The low value can be linked to the fact that not all the wealth is productive as we explained above.[30] Specification described in equation 3 implies that elasticity is a function of the level of wealth.[31] The wealth elasticity of income at the means of the two variables (and assuming the dummies equal to zero) is 0.347.[32]

Even though the analysis of a linear relationship is useful, visual inspection of Graph 1 and the pattern of residuals suggest that a linear specification is probably not the best one because there is some concavity on the data. If a third order polynomial in wealth is used in the regression, the fit of the model improves and the implied function is clearly increasing and concave. In order to incorporate a non-linear relationship between the variables, the log-in-log specification has two theoretical appeals.[33] First, it puts incomes in logs which is consistent with the observed empirical fact that the income distribution is well approximated by a log normal distribution. Second, it based on the implicit assumption that the wealth elasticity of income is constant which allows to compare this variable across different estimations in a more straightforward form.[34] Additionally, if we accept the assumption that wages are well described by a log normal distribution, we can assess the relative merit of a semi-log approximation (with wages in logs and wealth in levels) vs. a log-log approximation (with both variables in logs) using the approach suggested by Davidson-McKinnon (1981); this test clearly rejects the semi-log in favour of the log-log approximation.[35] Accordingly, we estimated an alternative specification with the two main variables in logs

$$x_i = \alpha' + \beta' tw_i + \gamma' Z_i + e_i \qquad (4)$$

where x_i is the log of income of household i, tw_i is the log of total wealth of household i, and Z_i is a set of other household characteristics.

---

[29] In Castile official interest rates decreased to 3% in 1705. In the former Crown of Aragon decreased to 3% with the Royal Decree (Real Pragmática) of July 6, 1750. See Yun (1987), p. 357. Anyway, there was no integrated public debt market for all Castile. In fact, there was a great diversity in nominal interest rates (from 5 to 10%) offered by the *juros de alcabalas* in different cities as Burgos (e.g. this city was able to offer 11 types of different interest rates in the mid-eighteenth century), Cadiz and Murcia. See Álvarez-Nogal (2009), pp. 129-130.

[30] If the two outliers to the right of graph 1 are excluded the slope changes to 0.049.

[31] If $X_i = Y_i + rPW_i$ the elasticity is $\epsilon_i = \beta \frac{PW_i}{Y_i + rPW_i}$ which is an increasing and concave function of PW_i.

[32] If the two outliers are excluded the elasticity goes up to 0.551.

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

[34] If we assume that the correct econometric model is a log in logs $x_i = \propto + \beta w_i^{\beta}$ where x=log(X) and w=log(W), we are implicitly assuming that $X_i = \gamma r PW_i^{\beta}$. This specification has a problem: it implies that income is zero whenever productive wealth is zero, which is probably not very reasonable.

[35] Results available under request.

In this case the parameter β' can be interpreted as the percentage change in income associated to each percentage point increase in wealth. In this case, there is the underlying assumption that this elasticity is constant for the different levels of income. The results of the regression with the variables in logs is in table 4:

TABLE 4

LOG REGRESSION ESTIMATES

	Dependent variable: *Log Income*
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

Note: Standard errors in parenthesis. Variables are significant at 10% (*), 5% (**) or 1% (***) level.
Source: Authors' calculations.

The elasticity of .579 suggests that an increase in one percentage point in wealth is associated to an increase of slightly more than a half a percentage point in income. The dummies associated to tertiary and secondary sectors are positive suggesting that for a given level of wealth, income is larger for households whose head is working in those sectors. Although we do not have an empirically backed explanation for this finding, it would be consistent with the idea that a higher share of wealth is "productive" in the secondary and tertiary sectors (for instance because land has a value as status goods but it is not always associated to higher incomes).[36]

Given that we have observations from Palencia in the North, Guadalajara and Las Vegas in the Center and Granada in the South, we can check if there are systematic differences in the estimated relationship across regions. In table 5 we present the results of the regression of specification (2) for each region.[37]

[36] Another possibility, more related with the characteristics of the data, is that wealth in these sectors could easier to conceal and therefore is systematically under-reported in PIs and therefore it seems to be more productive. For a discussion of possible under-reporting in the process of elaboration of a Probate Inventory see Nicolini and Ramos (2010, p. 156-159).

[37] In these "regional" regressions, the weights are specific for each region.

TABLE 5
LOG REGRESSION ESTIMATES: REGIONAL ANALYSIS

	Dependent variable: *Log Income*		
	Region 1: Palencia	Region 2: Guadalajara city and Madrid	Region 3: rural Granada
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

Note: Standard errors in parenthesis. Variables are significant at 10% (*), 5% (**) or 1% (***) level.
Source: Authors' calculations.

In the case of Palencia, wealth elasticity is quite close to the one obtained with the whole sample, the dummy variable associated with urban households is not significant (while it is negative in the whole sample), and the dummies associated to tertiary and secondary sectors are positive and significant.

In region 2, the elasticity is smaller than in region 1 and the parameters associated to both tertiary and secondary are positive; all the estimated parameters are statistically significant at 5 %. The effect of being located in an urban context is negative and statistically significant in Region 2 but non-significant in region 1, suggesting that what happens in the city of Guadalajara (the only city in Region 2) is different of what happens in Palencia (the only city in Region 1).

In Region 3 the only variable included in the regression is the log of wealth. This is because the localities included in our data set from Region 3 are mostly rural and with most of the head of the households engaged in activities in the primary sector (only 1 head of the household is engaged in secondary activities- a blacksmith- and 3 heads of the household are engaged in tertiary activities –an apothecary, a carriage-driver and a *Regidor Perpetuo*-); there are no urban centers in this region. The estimation of equation (2) in this case generates a wealth elasticity that is larger than in the other two regions: the point estimate is

smaller than one but the confidence interval includes values larger than one.[38] In this region, we cannot reject the hypothesis that income and wealth increase proportionally. A quite remarkable feature of the results in table 5 is that in Granada, with only 29 observations, the association between the two variables is very high with an R^2 of 0.71 (and adjusted R^2 of 0.70).

Several features of the results deserve additional comments: first, it is remarkable the close association between two variables that, collected from completely independent sources, provide alternative perspectives on the affluence of the households. This association suggests that, although it is tempting to be suspicious about the precision with which some historical sources provide an accurate picture of the material well-being of the population in the past, the two variables are quite good at capturing some underlying dimension of that material well-being.

A second implication of the results is that in average and in the two regions with more observations, the income of a household increases more slowly than its wealth. The limitations of our data renders impossible to provide a completely satisfactory explanation of this fact but it provides the possibility to discuss some hypotheses. The first one is that labor income (or more generally income non-related to physical assets) is relatively larger in the bottom part of the distribution; if this would be the case, as long as the households have larger wealth, their labor income decreases in relative terms producing a less-than proportional increase in income (vis-a-vis wealth). Figure 5 (graphs from 1 to 3) shows the relationship between the share of personal income in total income and total income. In addition to some households for which the share is zero or one (represented by the dots in the horizontal lines 0 and 1) it is clearly observable that the larger total income is, the smaller is the share of labor income.[39] Graphs 2 and 3 show that this pattern is much more detectable in those households whose head is mainly engaged in agricultural activities than in households linked to activities in the secondary and tertiary sectors.

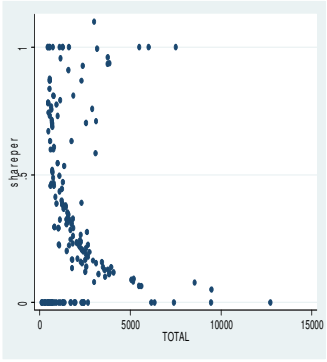
FIGURE 4

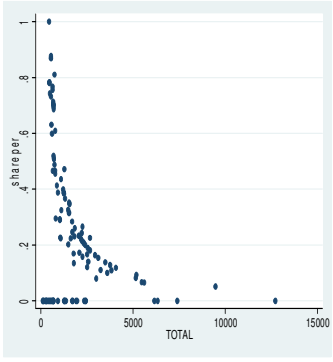
PERSONAL INCOME AS SHARE OF TOTAL INCOME

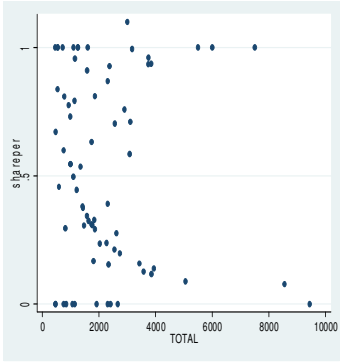
[38] The difference of the results in Region 3 is not driven by the fact that almost all the households in Region 3 are rural. If the equation (2) is estimated only for the rural households of the other regions the elasticity increases slightly (compared with the one obtained with both rural and urban households) but it is still smaller than one.

[39] The household with a share larger than one corresponds to Antonio de Laya, a lawyer with a personal income of 3,300 *reales* but a negative income because he had 297 *reales* from *censos en contra*. These "*censos en contra*" were the annual interest rate paid for financial liabilities. More details in Nicolini and Ramos (2015, p. 13).

Graph 1: Total	Graph 2: Primary sector	Graph 3: Secondary and Tertiary sectors







Source: Authors's calculations.

Another possibility is that wealth is more concentrated in income-producing assets in the bottom part of the distribution and more concentrated in other kind of assets (status goods, luxury consumption items, etc.) in the top of the distribution. Preliminary analysis of the participation of different kind of goods in the inventories suggests that this is the case: it is possible to see in Table 6 that the share of real estate (land and buildings), presumably income generating assets, in total wealth is clearly larger in the first quintiles while the shares of perishable goods and money in cash (probably not producing income) seem to be larger in Quintiles 4 and 5. The top quintile shows larger shares of debts in favor and financial assets (presumably producing some income in earned interests) and shop assets (presumably related to income generating activities in the tertiary sector) than the other parts of the distribution but these differences are not really large.

TABLE 6

PERCENTAGE OF ASSETS IN PROBATE INVENTORIES BY QUINTILES

Concept	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

Note: these shares are calculated only by probate inventories (116) from Palencia province. See Ramos Palencia (2010).
Source: Authors' calculations of probate inventories from Archivo Histórico Provincial of Palencia.

A third implication is that the relation between wealth and income is not the same for household whose head is engaged in agriculture than for households whose head works in the secondary or tertiary sectors. In addition, while in some regions like Palencia urban households seem to have a similar relation between income and wealth than rural households, it is not the case in Guadalajara where urban dwellers have smaller income for a given wealth than rural dwellers. This could be relevant when there are comparisons of levels of economic inequality across sectors or regions and some proxies like real estate or some subsets of assets are used to proxy for wealth or income (see Van Zanden 1995 and Alfani 2010).

6. Conclusions

The estimation of economic inequality in pre-industrial economies is forced to use creatively indirect data. The income distribution, usually reconstructed using social tables, tend to be imprecise in the top part of the distribution given that the variability of income within an occupational category is large. Distribution of wealth, usually using fiscal sources or probate inventories is badly estimated in the bottom part of the distribution because of the important selection bias that arises from the fact that in general poor people is not included in this kind of records. In this paper we have combined information of both income and wealth form the same set of households which helps to shed some light in the relation between the two distributions in a context of a pre-industrial, medium size, average income, semi-urbanized Spanish population.

The possibility to identify the position of each households in our data set of surviving probate inventories in the whole distribution of households (reconstructed using the Ensenada Cadaster), opens the possibility to analyze the magnitudes of the selection bias.[40] We can confirm that survival inventories are a biased slice of the society and that poorer households are clearly under-represented; however, some households in the bottom of the distribution are represented in our data set and we suggest that the widespread idea that households below the median income or wealth are completely absent from the probate inventories samples is not correct.

Another finding of this paper is that the income assigned by the Cadastre and the wealth registered in the inventories are closely associated not only in the whole sample of

[40] In this case, the "selection" is influenced by the choices and possibilities related to the process of the elaboration of the inventory in the 18th century plus the fact that some inventories have not survived until the present.

194 PIs but also when the three regions considered in this study (Palencia, Guadalajara-Madrid and Granada) are analyzed separately. This result suggest that even though the levels of income inequality seem to be consistently smaller than the levels of wealth inequality the particular location of one household in one distribution is very dependent of the location in the other distribution.

The data set used in this paper suggest that the best econometric specification to estimate the association between the variables is when both income and wealth are in logarithms and the results with this specification suggest the elasticity of income with respect to wealth is between 0.4 and 0.9 depending on the area of Spain. It implies that if a household has a 10 % higher wealth, its income would be between 4 % and 9 % larger.

An elasticity smaller than one is consistent with the usually observed fact that wealth inequality is larger than income inequality. The reasons behind this particular feature of the relationship between the two distributions in Modern Spain are not fully established in this paper but some hypotheses have been advanced: on the one hand, it seems that labor incomes are slightly larger relative to other sources of income in the bottom part of the distribution. On the other hand, the changes in the relative importance of the different components of wealth (land, livestock, buildings and urban properties, financial assets, money and consumption goods) across the different parts of the distribution suggest that those assets more clearly related with the generation of income (land) are more important in the bottom of the distribution.

Finally, the parameters associated with the dummy variables *SECONDARY* and *TERTIARY* are positive and significant in the whole sample and in two of the three regions suggesting that for a given level of wealth, households whose head works in those sectors tend to have higher income than those engaged in activities in the primary sector. These systematic differences in the relationship between income and wealth across urbanization levels and economic sectors could be a warning about using land or real estate as a proxy for wealth and wealth as a proxy for income: those households whose head is engaged in trade or manufacturing will have more wealth than the predicted by the value of their real estate and more income than the predicted by their wealth.

7. 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)." *Dondena Working Paper* 61, March, 2014.

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.

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

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

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

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).

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.

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. "Income Inequality in France, 1901-1998." *Journal of Political Economy* 111, no. 5 (2003): 1004-1042.

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

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.

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. (2016): "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, (1955): 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.