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

DOI: 10.1016/j.ehb.2018.12.006

Do you have to be tall and educated to be a migrant? Evidence from Spanish recruitment records, 1890–1950

Dácil Juif[a], Gloria Quiroga[b,*]

[a] Universidad Carlos III de Madrid, Spain
[b] Universidad Complutense de Madrid, Spain

Keywords:
Internal migration
Migrant self-selection
Spain
Early industrialization
Human stature

A B S T R A C T

We use Spanish military records stemming from the late-19th to the mid-20th century to assess internal migrants' self-selection. We find that migrants were, on average over the whole period, around one centimeter taller than non-migrants, and in the booming 1920s, the height advantage of movers reached three centimeters. The positive self-selection was larger for migrants originating in poorer provinces and traveling longer distances. A further finding is that migrants were positively selected in terms of literacy and socio-economic status according to their occupation. Professionals were most likely to have migrated internally and farmers least.

1. Introduction

The selectivity of migrants in early stages of industrialization and the effect that selective migration may have had on hosting and sending areas has received considerable attention in the recent economic history literature.

Most of these studies have focused on inter-continental migration during the so-called Age of Mass Migration in the late 19th to early 20th century.[1] The selectivity and the contribution of migrants are mostly assessed in terms of the migrants' education (literacy or numeracy) or occupations relative to the population that stays behind or to the receiving population.

A number of scholars have claimed that European immigration contributed significantly to economic development in the receiving nations. Stolz et al. (2013) argue that European immigration raised numeracy levels in the Brazilian provinces in which migrants settled, not only directly but also through spill-over effects on the native population by the creation of schools and hospitals. These areas grew faster than the non-immigration areas and the consequences persist until today. In a similar vein, Droller (2018) finds that counties in the Argentinean Pampa that received more European immigrants in the late 19th century entered a steeper GDP growth path and display higher levels of education even today. Sánchez Alonso (2007), her focus being on the River Plate as well, finds that immigrants displayed higher literacy levels than natives in receiving countries, and thus European immigration contributed to raise human capital. However, there are also cases in which particular groups of migrants did not contribute to raise education levels, i.e. their education level was lower than in the receiving country population (Juif, 2015).

Evidence on historical selectivity of European migrants with respect to the population at risk of migrating is mixed. Sánchez Alonso (2000) holds that migrants to the former Spanish colonies, who left Spain and Italy from 1870–1930, were positively selected in terms of literacy with respect to the population of their home country and province. Wegge (2002) finds that 19th century emigrants to the US, born in the German province of Hesse-Cassel, belonged mostly to the relatively skilled occupational group of artisans. According to her, the poorest and least educated did not have the possibility to migrate and the very wealthy lacked incentives to do so. Artisans could afford the passage costs more easily than wage laborers and faced lower transaction costs in converting their wealth into a portable form of capital than farmers. On the other hand, also using education-based proxies, Abramitzky et al. (2012) hold that Norwegian migrants to the US in the period of 1850–1914 were negatively selected in terms of their occupations and wealth from the sending population. Juif (2015), using age heaping proxies of numeracy, also finds a negative selection of emigrants from the Canary Islands to Cuba. She argues that migration was mainly driven by push-factors, particularly crises of the agricultural export sector that affected poor farmhands most.

Using height proxies of income, the results of Spitzer and Zimran (2018) suggest that migrants from Italy to the US came

* Corresponding author.
E-mail address: mariagloria.quiroga@pdi.ucm.es (G. Quiroga).

[1] The drivers of the Mass Migrations relate to the conjunction of events such as the transport cost revolution, a sharp rise in population growth in the old world and labor scarcity and land abundance in the New World (Hatton and Williamson, 1998).

from the poorer regions of Italy but were positively selected when compared to the potential migrants in their province of origin. Blum and Rei (2018) assess the selectivity of European Jewish migrants to the US during World War II and find that both refugee and nonrefugee passengers were positively selected, but non-refugees were even more so, suggesting that it was predominantly the European elite who escaped the Holocaust.

A few studies have analyzed the selectivity of internal national migrants in history and they mostly focus on migrants to the capital city. Humphries and Leunig (2009) analyze the effect of height on the probability to migrate to London among British seamen in the mid-19th century (controlling for the distance to their birth place and their age) and they find that the seamen who migrated were taller than those who did not. Their interpretation is that those who had better life chances were pulled to London and thus life there must not have been so bad. Beltrán Tapia and De Miguel Salanova (2017) find, for the same period that we study, that migrants to Madrid were positively selected in terms of their literacy. Silvestre (2005) has assessed several aspects of internal migration in Spain and he finds that internal migrants in the 1920s - other than international migrants - came mostly from poorer rural regions but the most impoverished areas in the South released very few workers. However, he does not use individual level information on migrants and stayers, and thus he cannot assess the selectivity of migrants from the province of origin.

Internal migration is mostly related to the process of urbanization and the exodus of rural workers from areas with high population growth where labour cannot be absorbed by the agricultural economy (also due to the rise of productivity in agriculture). Most of the mentioned studies on internal migration focus on the capital city, largely because of data constraints. They use a source containing detailed individual level information on residents of the city, including their birth place, and compare it with aggregate data concerning the places of origin. According to Beltrán Tapia and De Miguel Salanova (2017), almost 50 percent of Spanish internal migrants went to Madrid or Barcelona, but this still leaves a very large share of migration streams unexplored. For instance, migration to mining areas and other industrial places in Spain should not be neglected.

Unveiling the nature of internal migrant selectivity is important. The “brain drain”, i.e. the exodus of skilled physicians and other professionals, is considered an important impediment for economic development today. Furthermore, skill selective migration is probably one of the largely unexplored factors that contributed to regional convergence or divergence. It has been argued that income inequality between regions of Spain increased in the late 19th to early 20th century, in the early stages of industrialization, when industries became concentrated in Catalonia and Basque Country (Martínez-Galarraga et al., 2015). This was also the time when the railway network expanded and barriers to internal trade and labour movement were reduced. Migration movements continued to expand until the Civil War in 1936. Possibly, selective migration from largely rural areas to places where industries emerged and government service jobs were created, contributed to diverging development paths between regions. Furthermore, as mentioned by Beltrán Tapia and De Miguel Salanova (2017), studies that aim at assessing the ability of different regions to promote human capital acquisition could be actually capturing the effect of skill selective immigration or emigration. Lastly, internal migration does not face any legal restrictions and thus, other than studying current international migration, the setting is representative for the latent supply, i.e. the self-selection of those willing to move.

Using heights as our main indicator is an advantage, also because it has been employed much less frequently than literacy or occupations in studies on migrant selectivity until now. Average heights are an outcome of the state of nutrition, health, and hygienic and sanitary conditions in youth (see, for instance, Fogel, 1994; Steckel, 1995; Komlos and Baten, 2004; Komlos and Meermann, 2007).[2] Thus, human stature is indicative of income, wealth and life chances. Furthermore, individuals facing better conditions while growing up become taller but also develop better cognitive abilities, reach a higher level of education and thus higher incomes as adults (Case and Paxson, 2008). For some occupations there are returns to strength, which is correlated with height. Human stature is therefore an important determinant of labour productivity in a broader sense, and thus the effects of selective migration regarding stature will likely influence the growth path of sending and receiving areas. It has been argued that having a certain level of initial wealth and education are important preconditions for making the move but health is certainly another factor conditioning the possibility to migrate. Furthermore, as pointed out in Spitzer and Zimran (2018), the height variable is less coarse than occupations or literacy, which do not allow to reveal variations within a group.[3]

We add to the literature a study of internal migration in Spain. An important contribution is the use of an unusually complete database. The data consists of a random sample drawn from military records which capture the entire male population at the age of recruitment. It covers a long and decisive time period (1893 to 1954), and includes extensive information on each individual. Our database allows us to assess the characteristics of the individuals who moved (permanently) within Spain as well as those who stayed and were measured in their place of birth.

The article is structured as follows. In Section 2 we will describe the database, including the sampling method, as well as its potential biases. In Section 3 we will proceed to the analysis of migrant selectivity, first, on the national level and then disaggregating by region of origin and destination. Section 4 presents the results and interpretation. Section 5 concludes.

2. Data

The importance of heights as a measure of strength and probability of success in combat was understood by the armed forces in the 19th century. This is why they collected anthropometric information, and also introduced minimum height requirements for the military service.

Many anthropometric studies have used military conscript data because they concern a large number of individuals. A disadvantage of these data is that they include only males and just one age group. Furthermore, they are often truncated, i.e. upward biased, due to minimum height requirement, but this is not the case for our data, because every man of a certain age (except the very few deserters at this stage) was measured. Since the medical examination from which we draw the information took place one year before joining the military, potential deserters would probably still go through the medical test which could absolve them from conscription. Fig. 1 shows that our height data have a normal distribution and are not truncated.

For Spain, information on human stature derived from military records has been published in aggregate statistics for the period after 1954 and up to the end of the 20th century, when compulsory military service was abolished. Height data on the earlier period are much more difficult to retrieve. Ours is the only database that

[2] Genetics matter at the individual level, but genetic variation averages out when comparing large groups with each other (Steckel, 1995).

[3] For instance, the occupational group of farmers includes impoverished labourers, share-croppers and well-off landowners (see also Abramitzky et al., 2012)

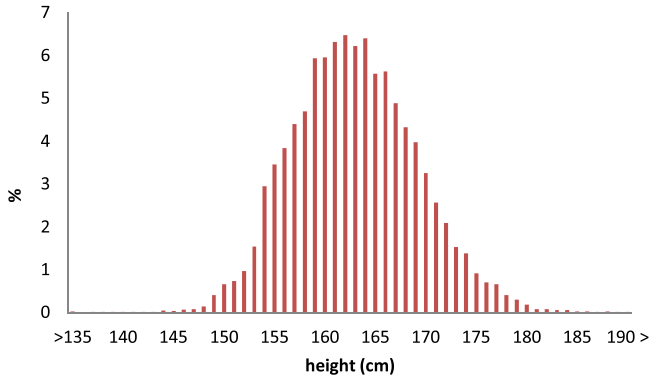

Fig. 1. Distribution of Spanish recruits' heights (1893–1954).

consists of a representative sample of male heights for the whole of Spain in the period before 1954.[4]

From around 1860, the municipal Enlistment Centres began to record detailed information on every young man called for a medical examination before recruitment (see Quiroga, 2003a). We go back to the Recruitment Sheet (*hojas de filiación*), a document filled out for every individual that contains very rich personal information: the names of his parents and his own, his birthplace and place of residence, anthropometric measures such as stature, weight and thoracic perimeter, profession and ability to read and write, as well as information relative to his military posting. Moreover, if the individual knew how to write, his signature would appear on the sheet and, if not, his fingerprint, a cross, or the signature of another person on his behalf. These forms were filled by every young man who had reached a certain age[5], on a particular day of the year (for many years this was the second Sunday of February), at the town hall corresponding to his residence place.

In theory, recruitment records were kept in the municipal archives. Unfortunately, very few archives conserve them at present and when they do, it is in a very fragmentary way. However, the military authorities of each province kept copies of the same information. These military files, once the recruit had completed his full military service, were remitted to a single centre, the Troop's branch of the Military Archive. There, the Ministry of Defence has conserved all the documents, both for men who completed their Military Service and for those who were excluded for some reason.

The Archive contains an enormous amount of documentation – around twenty million files – belonging almost entirely to the 20th century. As a general rule, the information for most provinces begins in the decade of the 1890s, although there are some provinces for which information begins somewhat later: Asturias, Cantabria and Murcia. In addition, some areas present gaps in the years of the Civil War, a time during which much information was destroyed or lost. The closing date of this paper is 1954, because aggregated information on the military recruits began to be published in this year.

To handle the enormous number of recruit files, it was necessary to use an annual random sampling representative of all Spanish regions, which allows us to perform inter-regional comparisons. The sampling for the period 1912 to 1954 was carried out the following way. For every singly year, from autonomous communities (of recruitment) that consist of a single province, 30 individuals were randomly selected. For autonomous communities with more than one province[6], 20 individuals were drawn from each province; for autonomous communities with more than two provinces, ten individuals were included for each province. This yields a total of 680 individuals for every year.

The information from 1893 to 1910 needs to be handled more carefully, but can still be used. Until 1911, there was no universal recruitment and buying exemption from service or hiring a substitute was widely practised by wealthy families. Since young men who were exempted in this manner did not have to go through the medical examination, the wealthy could be underrepresented in this source for the early time period. Furthermore, the dossiers of some provinces have not survived for this period. Nevertheless, a sample is carried out by including a statistically large enough number of individuals from each province available for every year (N = 40) and a sufficient number pertaining to low skilled professions (farmhands, etc.) and to higher skilled professions (artisans, students, industrial workers, etc.). In a second step the numbers were weighted by the share occupied by each group of professions in the actual population. Information on the distribution of occupations is derived from the national censuses. Furthermore, the age of recruitment was gradually raised from 19 years in 1899 up to 21 years in 1907.[7] This may affect the average stature, given late-growers, but it should not distort the results of our height comparison between migrants and non-migrants or between regions.[8]

The result is a sample of almost 50,000 individuals including data on stature, profession, town and province of birth and town and province of residence at the time of the medical examination, as well as self-reported literacy. Furthermore, an evaluation of their signatures has been carried out. The analysis of men's signatures might yield more accurate information on their functional literacy than the self-reported literacy. [9]

Among the advantages of this source we may underline its universality - since data exist for practically all Spanish males regardless of their physical condition - and the possibility to build annual series of stature, education and occupations, also at a small regional level.

Most important for this study, we have information on the migratory background of individuals. We know where they were born and where they were recruited, and thus, whether they moved at some point in between. There are no official statistics on interregional migratory movements in Spain until the 1960s, and thus this source is a valuable addition.

[4] Martínez-Carrión and Castejón (1998) use data from municipal archives only for the municipality of Elche. Martínez-Carrión and Moreno-Lázaro (2007) have recruitment data on heights for five provinces (two in the South-East and three in the North-West). María-Dolores and Martínez-Carrión (2011) use data on heights in a sample of five Spanish provinces. Martínez-Carrión et al. (2016) use provincial averages (not individual level data) for two years, 1858 and 1913. Other studies on heights cover Spain in this period equally fragmentarily.

[5] Age of recruitment underwent some variations during this time period. Before 1899, the age of recruitment was fixed at 19 years of age. From 1900 to 1906, recruitment age was raised by one year, and from 1907 onwards, it was fixed at 21. The period of the Civil War and the first years of the Second World War were exceptions to this rule. The age was raised in 1936 and 1937 to 22 and lowered to 19 and 20 during WWII (see Quiroga 1998, 2003a). In our sample, 70 percent were aged 21, 14 percent aged 20 and 16 percent aged 19.

[6] Autonomous community corresponds to the European NUTS 2 level and province to the NUTS 3 level.

[7] According to the 1912 Law, conscription took place in the year in which the person reached 21, but during the Civil War and until 1947, there were some disturbances in the age of recruitment. The 1938 reserve and the 1941 to 1947 reserves were called up almost one year in advance, while the 1939 conscripts passed their medical examination one year later than they ought to.

[8] However, it is possible that the share who had ever migrated is affected, as the older the individual the higher the probability that he ever moved permanently during his life.

[9] First, the signatures are split into five different categories: very good, good, average, bad and very bad. We then proceed to split them into above and below average. The idea is that we can detect whether someone has practice in signing or writing. The below average signatures include a child-like writing and sometimes even spelling mistakes. For more information, see Quiroga (2003b).

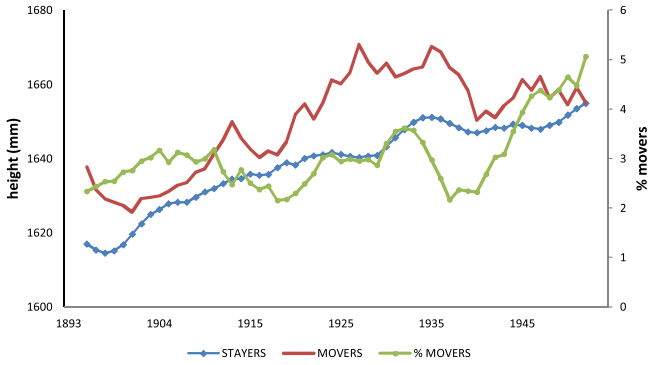

Fig. 2. Average height of Spanish conscripts and percentage of movers in the sample, 1893–1954 (five-year moving averages, by recruitment year).

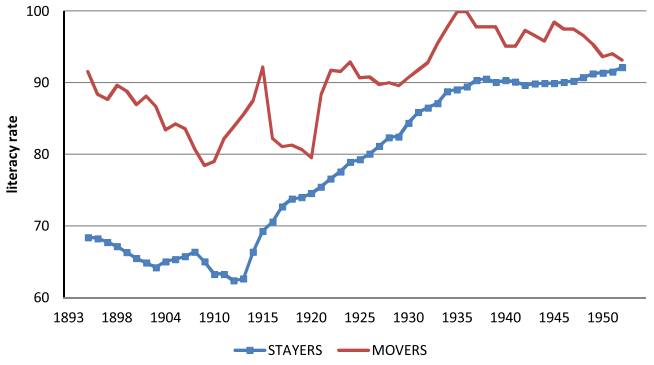

Fig. 3. Literacy rates of Spanish recruits (1893–1954), five-year moving averages.

For the first part of the analysis, we will divide the sample into those who moved and those who did not move, to assess the selectivity of migrants at the national level. In the second part of the analysis, we assess the selectivity of migrants from their province of origin.

3. Evolution of heights, literacy, and migrant selectivity

Fig. 2 presents the average height of the Spanish recruits, divided into "stayers", i.e. those who were born in the same province where they were measured, and "movers", conscripts who were examined in a different province than their province of birth. The graph also shows the evolution of the share of individuals who had moved before the time of measurement. We observe heights at the moment when men went through their medical examination, i.e. at age 19–21. At this stage, almost everyone had reached their completed height. However, the completed heights reflect the socioeconomic and health environment at an earlier stage of life. It is usually argued that the first three years of life and particularly adolescence are decisive for the final height reached. During adolescence, the height lost earlier due to adverse conditions can be caught up. Thus, we can assume that the final average height can to a large degree be attributed to socio-economic conditions around three years prior to measurement, but the 19–21 years before may have mattered as well.

We can derive three observations from this graph.

First, the evolution of the stature reflects different periods of Spanish economic and political development. We can distinguish a) A difficult end of the century, reflected in a decline of heights of both movers and stayers (although more so of movers). As Prados de la Escosura (2017) observes, this period is characterized by "sluggish growth"; b) A steady rise in the first thirty years of the 20th century –particularly in the golden 1920s, reflected in heights of 1930–1935 - that was abruptly interrupted by the Civil War; c) Civil War (1936–1939), World War II and a post-war period, characterized by adverse socioeconomic conditions, are marked by a slight descent and then stagnation of heights until 1954. A decline is observable even though in 1936 and 1937 the age of measurement was one year later than in the preceding years. It will not be until the late 1950s (not shown) that the situation of the Spanish economy begins to improve.

Second, the share of movers, i.e. internal mobility, goes through different phases. The share of recruits living in a different province than the province of birth is between 2.5 and 3 percent from 1893 and 1914. Disruptions during the First World War cause a decline in mobility to 2 percent. Afterwards, internal migration rises again to reach around 3.5 percent in the mid-1930s. In the 1920s, GDP growth in Spain accelerated and the pull of industry and services became stronger. Furthermore, permanent national mobility probably replaced international migration after 1914 (Sánchez Alonso, 1995). The rise in migration rates coincides with Silvestre's (2005) estimates. According to his calculations[10], around 4.3 percent of the population consisted of permanent internal migrants between 1921 and 1930. Given the young age of individuals included in our sample, it is not surprising that the share is lower here. During the years of the Civil War (1936 to 1939), the migrant rate is at its lowest point: 2 percent. Afterwards, its rise accelerates reaching up to 5 percent at the end of the period of observation.

Third, the height gap between movers and stayers varies over the period of analysis. Movers always have a higher mean height and, on average over the whole period, the difference between both series is over one centimeter. But the advantage of movers is largest in the 1920s to early 1930s (reaching almost three centimeters), coinciding with a period of economic bonanza and relatively high internal migration. The gap is reduced during the Civil war and almost closes at the end of the period.

We complement this analysis with the information available on recruits' literacy, quality of signatures and occupations.

Fig. 3 presents the self-reported average literacy of Spanish recruits between 1893 and 1954. We can observe stagnation or even a slight decrease in literacy at the end of the century, and up to 1913 (at a level of between 62 and 68 percent for the large majority, the "stayers"). Then there is a steady increase in literacy up to 1935, when stayers reach 90 percent self-reported literacy. Lastly, the data hint at stagnation at this level up until the end of the period of observation.

Using information reported in the national censuses, Nuñez (1992, p. 94) calculates that 63 percent of the adult male population (older than 10) was literate in 1910, 70 percent in 1920 and 81 percent in 1930. This almost coincides with our self-reported literacy rates for stayers. Ours are slightly higher - by 2–4 percent - which is plausible given that the younger population has a higher probability of being literate in a society in which school enrolment is increasing.

The graph also shows the positive selection of migrants regarding literacy. The fact that Spanish migrants were generally more literate than those who stayed behind has been pointed out

[10] Silvestre's calculations of internal migration within an intercensal period are made by subtracting the number of people born in another province (BAP) in the previous census - adjusting for the census specific mortality - from the number of people born in a different province in the later census: Internal migrations$_{t-1,t}$ =BAP$_t$ −(S×BAP$_{t-1}$). A problem of this methodology is the Survival Rate (S), which is different for the migrant population than for the whole census, given that migrants typically have a special age distribution and thus a different mortality rate than the whole population.

in the literature before. Sánchez Alonso (2000) finds that Spanish emigrants to Latin America had higher levels of literacy compared to the population of their birth province, and Beltrán and Salanova (2017) find this for internal migrants to Madrid. The reason is that education is important for migrants to reach the necessary information about potential destinations and that the potential returns to migration are higher for skilled workers. The difference between movers and stayers is largest in the early period, up to 1915, when movers display 15–20 percent higher literacy. Afterwards, the gap is significantly reduced to around 5–10 percent, with a slight increase in the mid-1920s, and it closes towards the end of the period of analysis. The latter could be a threshold effect, as the population is reaching 90 percent literacy, but it could also be related to friends and family effects, i.e. to a rising stock of previous migration, and to the declining cost of transport, which the literature has detected as factors reducing the positive selectivity of migrants.

A large number of historians have also used the share of people able to sign their name as a measure of literacy (Schofield, 1968, 1973). The UNESCO's definition of literacy is "the ability to read and write, as well as understand, a brief sentence related to his/her daily life" (see also Quiroga, 2003b). As writing is generally learnt at school after reading, the ability to sign would imply that a person has gone through the process of learning how to read and is, at least at a very basic level, able to write. Other than self-reported literacy (the answer to the question is either yes or no), someone's signature can give us some information about the mastery of the ability to write. Admittedly, our categorization is subjective - and a low signature quality can also be the consequence of physical inability-, but self-reported literacy is also subject to the recruit's criterion and honesty. As an additional exercise, we have classified the signatures of the recruits into six categories according to their quality (see Table A1). When comparing the quality of signature of stayers and movers we see that the share of signatures that we have categorized as "poor", "very poor" or "illiterate" is seventy-two percent for the stayers versus only forty-five percent for the movers. A disadvantage is that we only have signatures for around 50 percent of the sample. What this complementary analysis suggests is that beyond the most basic literacy, the migrants' mastery of writing (at least of their signature) was also superior to that of stayers.

The recruitment files source allows for another way of classifying the conscripts to look at differences between stayers and movers, namely by occupation[11]. We grouped the reported professions into nine major socio-professional categories: agricultural wage worker (including shepherds, woodcutters and similar jobs); farmer (comprising tenants, share-croppers and agricultural foremen); traditional service worker (includes barber, cook, miller, porter, waiter, coal seller); artisan (shoemaker, basket maker, silversmith, carpenter); industrial manufacturer (smelter, blacksmith, miner, forger); other modern service worker (railroad driver, postman, printer); white collar (clerk, telegraphist, accountant); professional (teacher, professor, medical doctor, lawyer, priest) and student. The category "other"[12] is a rest category that includes the unclassifiable entries.

Table 1 shows the average height for each occupational category, as well as the share of movers and stayers who belong to each group. Except for the category of "other modern services", which is relatively heterogenous, and contains many movers, the average height of individuals belonging to each group rises according to the socio-economic level and the level of instruction needed.

Table 1
Percentage of stayers and movers and average literacy rate by occupation (1893–1954).

Occupational category	Class	Average Height (in mm)	Literacy %	Stayers %	Movers %
Agricultural wage workers	1	1,629	0.60	33.10	19.57
Farmers	2	1,634	0.80	26.99	5.46
Traditional services	3	1,635	0.87	8.56	12.17
Artisans	4	1,635	0.87	9.75	14.11
Industrial manufacture	5	1,641	0.82	7.57	11.48
Other modern services	6	1,656	0.95	3.36	10.65
White Collar workers	7	1,653	0.99	4.71	10.24
Professionals and students	8	1,663	1.00	4.77	14.32
others	9	1,645	0.90	1.18	2.01
Total		1,636		100.00	100.00

A glance at the two last columns of Table 1[13] shows that the occupational categories of higher socio-economic status and higher skill requirements are overrepresented among movers. The share of students and professionals as well as of modern services is three times as high among those who have migrated internally than among those who have not; the share of white-collar workers is double. Artisans, traditional service workers and industrial manufacturers are overrepresented among the movers as well (at a ratio of movers to stayers of 1.5 to 1). On the other hand, those who worked in agriculture were much more likely to remain in their regions of origin. The least mobile group are the farmers (the ratio of stayers to movers is 5 to 1), probably because land property is more difficult to turn into movable capital than other types of wealth. The low mobility of farmers has been found by other scholars, for instance by Wegge (2002) for Germany. In our case, their literacy was also relatively low, which probably made it more difficult for them to find a job in an urban setting.

This analysis seems to confirm that movers were better off in terms of height, literacy and socio-occupational status than their counterparts who stayed in their province of birth.

In a second step it is important to assess which provinces sent and received most migrants and whether migrants are also positively self-selected from their province of origin. Given large geographical differences in average heights[14] (see Quiroga, 1998), we might just be capturing the geographical composition of migrants, i.e. it could be for some reason that most movers originated from richer (taller) regions. Furthermore, given that the individuals in our sample must have made the move at a quite young age, it is also possible that the movers grew and became literate or learnt a high-skill profession *because* they went to a place that offered better opportunities.

As the literature has claimed, if the pull effect of the destination was stronger, then those with better life chances at home might have migrated (movers would be positively self-selected) from their province of origin, and they might even surpass the average height in the province of destination. If the push effect was stronger, then the internal migrants could be negatively selected

[11] It is unclear in some cases whether the profession reported is that of the recruit or that of his father. However, we do not consider this an impediment, since we are also interested in the socio-professional status of the family amidst the conscript grew up.

[12] The height in this category has a much larger standard error than in the other categories.

[13] This table shows the average across the whole period of observation.

[14] The tallest young men resided in the Canary Islands, Catalonia and the Basque Country in the late 19th century. The latter two are the paramount industrial regions and the archipelago is said to have a proliferate climate. At the turn of the century, Madrid catches up and joins the leading group. The autonomous communities that display the shortest statures are Extremadura, Castilla la Mancha, Galicia and Castilla-León, and Murcia. Differences between regions reach up to 6 cm in the initial years (Quiroga 1998; Quiroga and Coll, 2000).

from their province of origin, and certainly shorter than the average population in the province of destination.

4. Origin, destination and self-selection of migrants

4.1. Migration movements

As mentioned earlier, it is useful to assess which areas were senders and which ones were receivers of migrants, given the regional differences in heights. This exercise has been carried out before by Silvestre (2005) using census data. Table 2 reports the net migration (the difference between the number of men recruited in an autonomous community and the number of men born in the same autonomous community) in absolute numbers for three different periods: the first period is the one for which our sample is possibly biased (1893–1910); the second period is a peaceful period form 1911 until 1935, and includes the booming 1920s; the third period includes the Civil War, the Second World War, but also part of the post war period. For an overview, we have divided the autonomous communities into larger regions: a) Southern and central Spain, with large negative net migration rates; b) the capital city, Madrid, with by far the largest positive net migration rates; c) Northern Spain with mostly positive net migration (apart from Asturias in the period of 1911 to 1935). There was probably considerable migration within this region, between provinces or to the industrial centres; d) In the Mediterranean region we include Cataluna and Comunidad Valenciana. Cataluna was, together with the Basque Country, the most industrialized autonomous community in Spain and one of those that received most economic migrants. In this overview table, migration taking place within the same autonomous community is not taken into account, however.

For a more detailed analysis, in the Appendix (Table A2a and A2b) we show the net migration by province, and provinces can be located in a map of Spain. The northern provinces of Santander, Vizcaya and Guipuzcoa display high positive net migration, i.e. more in-migration than out-migration. These were among the most industrialized provinces of Spain. Adjacent provinces like Oviedo, Burgos and Palencia have relatively high rates of emigration, because of their proximity to an important pole of attraction, and low immigration. Madrid has by far the highest positive net migration rates, followed by Barcelona and Seville. The provinces surrounding these urbanized places have relatively large negative net migration rates (Toledo, Cuenca, Guadalajara; Huelva, Córdoba, Badajoz, Malaga).

This analysis shows also that most internal migrants did not originate in the "tallest" regions but rather the contrary. This hints to the fact that the height advantage of migrants is not the result of being born in well-off provinces.

4.2. Personal traits of migrants

In the following we will analyze the characteristics of those who migrate at the individual level. Logistic regression analysis allows us to assess if height and literacy increased the probability of migrating, controlling for the average height in their province of origin, and holding the time of measurement constant.

The dependent variable is a binary variable which takes the value of one if someone moved between birth and measurement and zero otherwise. The main independent variable is individual height. In the appendix (Table A4) we also show the results for literacy, signature quality and occupation group. Other control variables are: recruitment year (as migration largely increased over time), age (since we have small variations in age at measurement), mean height in the province of birth in the corresponding period of recruitment. The regressions in Table 3 perform as expected. Results suggest that the individual's stature increased the odds of migrating (column 1). After controlling for the year of recruitment, a one unit increase in height raises the probability of migrating by 0.0001.This does not seem much, but it is considerable given that the mean of the "migrant" variable is only 0.03, i.e. three percent

Table 2
Absolute net migration by autonomous community.

Region and autonomous community	Net migration		
	1893-1910	1911-1935	1936-1954
South and Centre			
Andalucia	−21	−5	−41
Castilla la mancha	−43	−44	−36
Castilla leon	−48	−43	−27
Extremadura	−1	−5	−16
Murcia		−9	−11
Capital city			
Madrid	114	103	53
Northern Spain			
Aragon	−27	−12	7
Asturias		−8	2
Cantabria	1	1	9
Galicia	17	0	4
La Rioja	−11	−2	6
Navarra	−1	3	5
Pais Vasco	65	19	21
Mediterranean			
Cataluna	−1	34	37
Comunidad Valenciana	−14	−10	6
Islands			
Baleares	1	5	24
Canarias	1	5	9

Source: Authors own calculation. Data: Recruitment files.

Table 3
Logistic regressions of migration on other individual characteristics.

	(1)	(2)	(3)
Height (indiv.)	0.0001***	0.0001***	0.0000**
	(6.20)	(6.13)	(2.31)
Recr. year	0.0002***	0.0003***	0.0003***
	(4.81)	(4.58)	(3.47)
Height birth prov.		−0.0001	−0.0001*
		(-1.15)	(-1.81)
Age		−0.0027	−0.0001
		(-1.62)	(-0.08)
Read			0.0151***
			(2.31)
Occupations			
agricultural worker			−0.0251***
			(-8.41)
farmer			−0.0586***
			(-13.73)
traditional services			−0.0021
			(-0.66)
artisan			**Ref.cat.**
industrial manuf.			−0.0010
			(-0.33)
modern services			0.0177***
			(5.10)
white collar			0.0091**
			(2.68)
professional			0.0184***
			(5.84)
other			0.0049
			(0.79)
Obs.	47,132	47,132	43,900
Pseudo-R2	0.0063	0.0066	0.0772

Notes: Table shows the average marginal effects, p-values as stars, and z-scores in parentheses. * = significant at 10%, **significant at 5%, ***significant at 1%.

had moved internally (see Table A3). The significant positive coefficient of height remains after including additional controls (column 2 and 3). The coefficient of *height birth prov.* has the expected negative sign, suggesting that migrants originated mostly in provinces where the mean height of men was low, particularly in the south and in the centre. The coefficient for the age of measurement is insignificant. In column 3 we add other personal characteristics. Being literate significantly increases the probability of having migrated (by 0.02). Furthermore, relative to artisans, agricultural wage workers were less likely to move. Even less mobile were farmers. Those belonging to the occupational categories of "white collar workers", "other modern services" and "professional and students" were significantly more likely to have moved, in that order. "Traditional services" and "industrial manufacture" workers were not significantly different from artisans in their likelihood to emigrate.

4.3. Explaining the migrants' height advantage

Finally, we want to assess which factors have an impact on the height difference between movers and stayers, i.e. what affects the self-selection of migrants. The economic history literature has emphasized migration chains (see, for instance, Wegge, 1998) and geographical and cultural proximity between home and host area (Stolz and Baten, 2012) as factors reducing the costs of migration and thus making it easier for the less well-off to make the move.

For this purpose, we construct a panel dataset at the province level and for three periods of time (1893–1910; 1911–1935; 1936–1954). We then proceed to estimate the following regression model using random effects:

$$\text{HEIGHT GAPit} = \alpha + \beta_1 \text{ DISTANCE}_{it} + \beta_2 \text{ MIG STOCK}_{it} + \beta_3 \text{ HEIGHT BIRTH PROV}_{it} + \alpha t + uit$$

HEIGHT GAPi t is the difference in heights between the stayers and all the movers (irrespective of their destination) in each province *i* and for time period *t*. The independent variables include:

a) $DISTANCE_{it}$ is the average distance migrants travelled to their destinations. We first constructed a variable denoting the distance between birth and destination province for the individual migrants (using the centroid of each province). Then we built the average of this variable for all emigrants from each birth province and for each time period. We assume that distance raises the costs of migration and thus it deters the less well-off from migrating. The coefficient is thus expected to be positive, i.e. if on average migrants move to the adjacent province, we expect a lower positive selection than if migrants head far-away destinations. In column 2 we include a distance variable which takes the value of 1 if the average distance that migrants moved away was more than 100 km.
b) $MIG\ STOCK_{it}$ is the net migration rate in the province of origin *i* in period *t*. With this variable we control for the "friends and family" effect. By sending information and remittances, as well as providing temporary accommodation and help, these social networks encourage chain migrations and allow less well-off individuals to make the move. Thus, the larger the diaspora, the lower are the costs of migrating and the lower the impediments for the less well-off. The assumption is here that the direction of migration remained broadly constant.
c) $HEIGHT_{it}$ (inverse poverty constraint) is the average height in the province of origin *i* in the corresponding period *t*. We expect a negative sign for this variable. One possible reason is that more people were able to bear the initial monetary costs of migration in the richer areas, and thus potential migrants did not necessarily belong to the best-off section of the population. This variable would thus represent the inverse of the "poverty constraint" to emigration, which earlier literature has pointed out (e.g. Hatton and Williamson, 2005; Stolz and Baten, 2012). Another argument could be that in richer provinces more positively selected potential migrants left for the Americas, whereas in the poor provinces the majority of - positively self-selected - migrants headed to Spanish destinations. According to Sánchez Alonso (2000)[15], around 1910 migrants to Latin America came from relatively rich and literate regions of Spain, because intercontinental migration was income constrained, and Silvestre argued that internal migration could have served as a substitute to international migration, which was more costly and risky.

Table 4
Regressions of height difference between movers and stayers on distance, previous migration and average height.

Dep. Var.	(1) heightdiff	(2) heightdiff
Distance	0.468	
	(1.781)	
Net Migration Rate	−0.460	−0.491
	(0.359)	(0.362)
Height Birth Prov. (inverse poverty constraint)	−0.359**	−0.338**
	(0.128)	(0.112)
Distance > 100km		14.684*
		(13.958)
Constant	600.112***	553.133***
	(208.069)	(183.509)
Observations	112	112
Number of birth prov.	45	45
R-squared overall	0.107	0.118

Notes: Table reports the coefficients, p-values as stars, and robust standard errors in parenthesis. We excluded the period-province combination if the observation number was smaller than 4. * = significant at 10%, **significant at 5%, ***significant at 1%.

In Table 4 we report the results of the regression. The coefficient for distance in specification 1 has the expected positive sign but is insignificant. However, in specification 2 we use instead of a continuous distance variable a dummy which takes the value of 1 if the average distance migrated was higher than 100 km. The coefficient of this variable is positive and significant. It seems that, if the distance travelled is above a threshold, it affects the positive selection of migrants more significantly. The net migration in the province of origin, which aims to capture migration streams in place has the expected negative sign but is insignificant. The average height in the province of birth is negative and significant in both specifications, meaning that the selectivity of migrants was

[15] According to Sánchez Alonso (2000), the largest emigration rates around 1910 were in the North-West of Spain, but also in Andalucía and the Canary Islands. Our database allows us to shed some light on the origin of most international migrants from Spain as well. It includes "foreign" recruits, whereby the majority is born in Latin America (61%), then Europe (30%), and most probably consists of (second and later generation) return migrants. If we assume that most of them go back to their ancestors' place of origin, which is plausible, international migrants largely come from the North and North-Western Castile (40%), Southern Castile and Andalucía (19%), as well as from the Canary Islands (10%). The population of the North of Spain was comparatively rich and literate. Galicians were not particularly tall, but the rest of Northern provinces had taller than average populations (Quiroga, 1998). Canary Islanders seem to be exceptional, as they were very tall, despite their very low income.

higher for those who came from poorer areas (with on average lower heights).

5. Conclusions

Using an unusually complete database, including individual level information regarding heights, literacy and occupations, reaching from the late 19th to the mid-20th century, we have analyzed the characteristics of internal migrants in Spain. Migrants were taller, more likely to be literate and to enjoy a higher socio-economic status according to their occupation than those who stayed behind. For some occupations like professionals, white collar or skilled industrial workers it may have been easier to make the move because their skills were more demanded in the industrial and urban centres. The positive selection of internal migrants may indicate that the pull effect was stronger than the push-effect, because those who had the better life chances decided to move to more attractive places.

Our analysis confirms the earlier finding that emigration rates were higher in the poorer southern and central part of Spain (i.e. migrants were more likely to originate from provinces with lower average height), but also in the provinces surrounding industrial and urban poles of attraction, particularly the Basque Country, Barcelona and Madrid.

Migrants were not only positively selected at the national level, but also from their province of birth, and the advantage in height and literacy of emigrant was probably larger in poorer provinces. This suggests that there may have been literacy or wealth constraints to migration at play.

The positive selection of those who left, particularly in poorer areas, suggests that migration may have acted as a "brain drain" and contributed to regional inequality within Spain. The results call for attention to the quality of migrants relative to their local environment, and not only in absolute terms or relative to a larger national pool.

Appendix A

See Table A4
Source: Silvestre (2005).

Table A1
Percentage of recruits according to the quality of their signatures.

	STAYERS (%)	MOVERS (%)
Illiterate	29.74	12.59
Very poor	9.83	7.75
Poor	32.25	24.47
Average	15.57	20.54
Good	10.19	24.97
Very Good	2.42	9.67
Total	100.00	100.00

Table A2a
Panel A. Net migration (total number measured minus total number born in a province).

Region	ID	Province	1893-1910	1911-1935	1936-1954	Region	ID	province	1893-1910	1911-1935	1936-1954
North	**1**	La Coruna	−4	1	3	Medi-terranean	**25**	Gerona	−1	−4	8
	2	Lugo	**7**	**−1**	**1**		**26**	Barcelona	**−6**	**37**	**13**
	3	Ponteveda	14	1	2		**27**	Tarragona	0	−4	9
	4	Orense	0	−1	−2		**28**	Castelló	−5	−3	3
	5	Oviedo	−11	−8	2		**29**	Valencia	0	0	5
	6	Santander	**1**	**1**	**9**		**30**	Alicante	−9	−7	−1
	7	Vizcaya	**39**	**17**	**10**		**31**	Murcia	−2	−9	−11
	8	Guipuzcoa	**6**	**3**	**4**		**32**	Baleares	1	7	24
North Castile	**9**	Leon	−9	0	−3	South Castile	**33**	Madrid	**114**	**103**	**53**
	10	Palencia	−11	−2	−3		**34**	Guadalajara	−2	−11	−5
	11	Burgos	−20	−13	−8		**35**	Cáceres	6	3	−8
	12	Zamora	1	−2	1		**36**	Toledo	−12	−13	−9
	13	Valladolid	3	−11	−13		**37**	Cuenca	−7	−5	−11
	14	Soria	1	−5	0		**38**	Badajoz	−7	−10	−8
	15	Salamanca	−8	−3	0		**39**	Ciudad Real	−10	−11	−7
	16	Avila	0	−1	−1		**40**	Albacete	−12	−4	−5
	17	Segovia	−5	−6	0	Andalusia	**41**	Huelva	3	−4	0
	18	Alava	20	−1	7		**42**	Sevilla	**12**	**14**	**2**
	19	Navarra	−1	3	5		**43**	Córdoba	−3	9	−6
Ebro Valley	**20**	Logroño	−11	−2	6		**44**	Jaén	1	−15	−15
	21	Huesca	−8	−2	6		**45**	Cádiz	−13	−1	1
	22	Zaragoza	−15	−7	4		**46**	Málaga	−15	−2	−6
	23	Teruel	−4	−3	−3		**47**	Granada	−3	−7	−6
	24	Lerida	6	5	7		**48**	Almería	−3	1	−11

Table A2b
Panel B: Net migration per 1000 population.

Region	ID	Province	1893-1910	1911-1935	1936-1954	Region	ID	province	1893-1910	1911-1935	1936-1954
North	**1**	La Coruna	−0.83	0.40	1.56	Medi-terranean	**25**	Gerona		−3.31	4.00
	2	Lugo	1.09	−0.40	0.53		**26**	Barcelona	−2.04	**14.74**	**6.47**
	3	Ponteveda	2.33	0.40	1.04		**27**	Tarragona	0.00	−1.55	4.74
	4	Orense	0.00	−0.40	−1.04		**28**	Castelló	−0.78	−1.26	1.58
	5	Oviedo		−3.98	0.36		**29**	Valencia	0.00	0.00	2.63
	6	Santander	0.25	0.75	1.55		**30**	Alicante		−3.06	−0.53
	7	Vizcaya	12.19	6.59	5.05		**31**	Murcia		−3.69	−1.92
	8	Guipuzcoa	2.50	3.16	1.89		**32**	Baleares	0.23	0.93	4.15
North Castile	**9**	Leon		0.00	−1.58	South Castile	**33**	Madrid	14.11	13.43	9.27
	10	Palencia		−1.42	−1.60		**34**	Guadalajara	−0.33	−4.25	−2.58
	11	Burgos	−3.13	−4.26	−3.69		**35**	Cáceres	0.94	0.58	−2.06
	12	Zamora	0.16	−0.83	0.51		**36**	Toledo		−5.99	−4.74
	13	Valladolid	1.03	−4.15	−6.74		**37**	Cuenca	−2.35	−1.89	−5.76
	14	Soria	0.15	−1.71	0.00		**38**	Badajoz		−2.14	−2.11
	15	Salamanca	−2.87	−1.16	0.00		**39**	Ciudad Real		−12.22	−3.68
	16	Avila	0.00	−0.39	−0.53		**40**	Albacete		−8.89	−2.65
	17	Segovia	−0.83	−1.97	0.00	Andalusia	**41**	Huelva	0.77	−1.34	0.00
	18	Alava	3.13	−0.40	3.66		**42**	Sevilla	1.86	2.75	1.00
	19	Navarra	−0.16	0.40	0.92		**43**	Córdoba	−0.73	3.57	−3.16
Ebro Valley	**20**	Logroño	−1.57	−0.26	1.05		**44**	Jaén	0.20	−6.33	−7.77
	21	Huesca		−1.17	3.16		**45**	Cádiz		−0.69	0.52
	22	Zaragoza	−3.35	−2.71	2.20		**46**	Málaga		−0.84	−3.16
	23	Teruel	−0.63	−1.07	−1.58		**47**	Granada	−0.69	−2.80	−3.16
	24	Lerida	1.36	2.66	3.68		**48**	Almería	−1.55	0.40	−5.76

Table A3
Summary Statistics of variables.

Variable	Obs	Mean	Std. Dev.	Min	Max
Height (mm)	47,132	1636.3	6.2	1007.0	1994.0
Low sign quality	33,699	0.57	0.49	0.00	1.00
Occupation class	47,132	2.96	2.19	1	9
Read	43,900	0.78	0.42	0.00	1.00
Province mover	47132	0.03	0.17	0.00	1.00
Recruitment year	47132	1920.7	1.9	1893.0	1954.0
Age	47132	20.54	0.76	19	20
Period	47,132	1.87	0.82	1	3
Average height birth prov-period	47132	1636.3	2.1	1596.6	1685.6

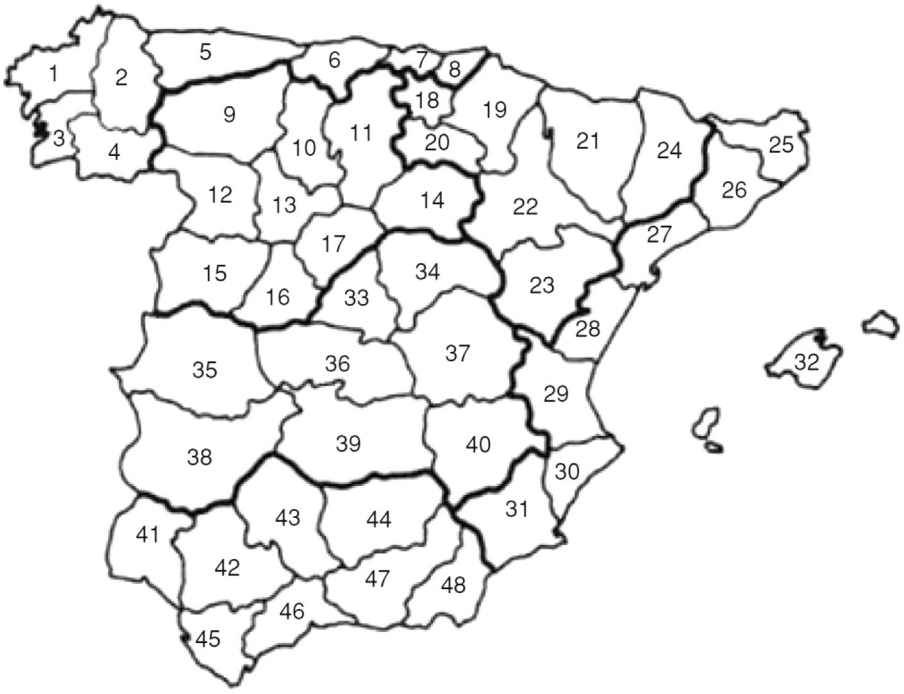

Table A4
Regressions of migration on literacy, signature quality and occupations.

Main expl. var.	(1) Literacy	(2) Low sign quality	(3) occupations
Read	0.0297***		
	(9.95)		
low sign quality		−0.0095***	
		(-4.64)	
Occupations:			
agricultural worker			−0.0278***
			(-9.88)
farmer			−0.0592***
			(-14.38)
traditional services			−0.0017
			(-0.55)
artisan			ref. cat.
industrial manuf.			−0.0017
			(-0.53)
modern services			0.0185***
			(5.51)
white collar			0.0106***
			(3.24)
professional			0.209***
			(6.95)
other			0.0008
			(0.13)
Recruitment year	0.0002*	0.0005***	0.0003***
	(2.33)	(5.28)	(4.46)
age	−0.0026	−0.0032	0.0003
	(-1.47)	(-1.54)	(0.17)
Height birth prov.	0.0001	−0.0001	−0.0001
	(1.49)	(-1.00)	(-1.50)
Obs.	43,900	33,699	47,132
Pseudo-R2	0.0147	0.0062	0.0732

Notes: Table shows the average marginal effects, p-values as stars, and z-scores in parentheses. * = significant at 10%, **significant at 5%, ***significant at 1%.

References

Abramitzky, R., Boustan, L.P., Eriksson, K., 2012. Europe's tired, poor, huddled masses: self-selection and economic outcomes in the age of mass migration. Am. Econ. Rev. 102 (5), 1832–1856.

Beltrán Tapia, F.J., de Miguel Salanova, S., 2017. Migrants' self-selection in the early stages of modern economic growth, Spain (1880–1930). Econ. Hist. Rev. 70 (1), 101–121.

Blum, M., Rei, C., 2018. Escaping Europe: health and human capital of Holocaust refugees. Eur. Rev. Econ. Hist. 22 (1), 1–27.

Case, A., Paxson, C., 2008. Stature and status: height, ability, and labor market outcomes. J. Polit. Econ. 116 (3), 499–532.

Droller, F., 2018. Migration, population composition and long run economic development: evidence from settlements in the pampas. Econ. J. 128 (614), 2321–2352.

Fogel, R.W., 1994. Economic growth, population theory, and physiology: the bearing of long-term processes on the making of economic policy. Am. Econ. Rev. 84 (3), 369–395.

Hatton, T.J., Williamson, J.G., 1998. The age of mass migration: Causes and economic impact. Oxford University Press.

Hatton, T.J., Williamson, J.G., 2005. What fundamentals drive world migration? In: Poverty, international migration and asylum. Palgrave Macmillan, London, pp. 15–38.

Humphries, J., Leunig, T., 2009. Was Dick Whittington taller than those he left behind? Anthropometric measures, migration and the quality of life in early nineteenth century London?. Explorations Econ. Hist. 46 (1), 120–131.

Juif, D., 2015. Skill selectivity in transatlantic migration: the case of Canary islanders in Cuba. Revista de Historia Económica/Journal of Iberian and Latin American Economic History 33, 189–222.

Komlos, John, Baten, J.örg, 2004. Looking backward and looking forward: anthropometric research and the development of social science history. Soc. Sci. Hist. 28 (2), 191–210.

Komlos, John, Meermann, Lukas, 2007. The introduction of anthropometrics into development and economics. Hist. Soc. Res. 32 (1), 260–270.

María-Dolores, R., Martínez-Carrión, J.M., 2011. The relationship between height and economic development in Spain, 1850–1958. Econ. Hum. Biol. 9 (1), 30–44.

Martínez-Carrión, J.M., Castejón, J.J.P., 1998. Height and standards of living during the industrialisation of Spain: The case of Elche. Eur. Rev. Econ. Hist. 2 (2), 201–230.

Martínez-Carrión, J.M., Moreno-Lázaro, J., 2007. Was there an urban height penalty in Spain, 1840–1913? Econ. Hum. Biol. 5 (1), 144–164.

Martínez-Carrión, J.M., Cámara, A.D., Pérez-Castroviejo, P.M., 2016. Parámetros antropométricos de los reclutas españoles antes de la transición nutricional: análisis de las desigualdades territoriales (1858-1913). Nutr. Hosp. 33 (6), 1477–1486.

Martínez-Galarraga, J., Rosés, J.R., Tirado, D.A., 2015. The long-term patterns of regional income inequality in Spain, 1860–2000. Reg. Stud. 49 (4), 502–517.

Nuñez, C.E., 1992. La fuente de la riqueza: Educación y desarrollo económico en la España contemporánea. Alianza Universidad., Madrid.

Prados de la Escosura, L., 2017. Spanish economic growth, 1850–2015. Palgrave Studies in Economic History, .

Quiroga, G., 1998. Height evolution in Spain, 1893-1954: an analysis by regions and professions. In: Komlos, enJ., Baten, J. (Eds.), Studies on Biological Standard of Living in Comparative Perspective. Franz Steiner Verlag, Stuttgart, pp. 359–383.

Quiroga, G., 2003a. Medidas antropométricas y condiciones de vida en la España del siglo XX, Tesis Doctoral. Universidad de Alcalá de Henares.

Quiroga, G., 2003b. Literacy, Education and Welfare in Spai (1893-1954). Paedagog. Hist. 39 (5), 599–619.

Quiroga, G., Coll, S., 2000. Height Inequality as a Proxy for Income Inequality. The Spanish case, 1895-1950. More Reflexions. J. Income Distrib. 9, 1–24.

Sánchez Alonso, B., 1995. Las causas de la emigración espanola 1880-1930. Alianza Editorial., Madrid.

Sánchez Alonso, B., 2000. Those who left and those who stayed behind: explaining emigration from the regions of Spain, 1880–1914. J. Econ. Hist. 60 (03), 730–755.

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

Schofield, R.D., 1968. The Measurement of Literacy In Pre-Industrial England. In: Goody, Jack (Ed.), Literacy In Traditional Societies. Cambridge, pp. 311–325.

Schofield, R.D., 1973. Some dimensions of illiteracy. Explorations in Economic History, , pp. 437–454 10, 4.

Silvestre, J., 2005. Internal migrations in Spain, 1877–1930. Eur. Rev. Econ. Hist. 9 (2), 233–265.

Spitzer, Y., Zimran, A., 2018. Migrant self-selection: anthropometric evidence from the mass migration of Italians to the United States, 1907–1925. J. Dev. Econ. 134, 226–247.

Steckel, Richard H., 1995. Stature and the standard of living. J. Econ. Lit. 33 (4), 1903–1940.

Stolz, Y., Baten, J., 2012. Brain drain in the age of mass migration: does relative inequality explain migrant selectivity? Explor. Econ. Hist. 49 (2), 205–220 2012.

Stolz, Y., Baten, J., Botelho, T., 2013. Growth Effects of Nineteenth-century Mass Migrations:"Fome Zero" for Brazil? Eur. Rev. Econ. Hist. 17 (1), 95–121.

Wegge, S.A., 1998. Chain migration and information networks: evidence from nineteenth-century Hesse-Cassel. J. Econ. Hist. 58 (4), 957–986.

Wegge, S.A., 2002. Occupational self-selection of European emigrants: evidence from nineteenth-century Hesse-Cassel. Eur. Rev. Econ. Hist. 6 (3), 365–394.