The comovement between height and some economic development indicators in Spain

Ramón Maria-Dolores
José Miguel Martínez Carrión

(Universidad de Murcia)
The Co-movement between Height and Some Economic Development indicators in Spain, 1850-1978

Ramón María-Dolores and José Miguel Martínez-Carrión
Universidad de Murcia
February, 2012

Keywords: Height, physical stature, anthropometrics, education, economic development, VAR forecast errors

JEL classification: I1, I3, N3, N9

Abstract

This paper investigates the relationship between height and some measures of human welfare in Spain for the period 1850-1978. For that purpose, we employ several filtering methods to measure the correlation between variables such as first order differences, deterministic trends, the Hodrick and Prescott filter, the band-pass filter or den Haan (2000)’s methodology which uses a new set of statistics to characterize the co-movement between variables to capture the dynamic between variables. We always find a strong and positive correlation between height and GDP per cápita, height and the weight of health services in total consumption, and height and openness. By contrast, we have a negative correlation between height and the mortality rate and height and the ratio between the deflator of private consumption and the GDP deflator. By applying den Haan (2000)’s method, we find that the co-movement between height and GDP per cápita is always positive, increasing in the medium and long-run. This correlation is higher after the Spanish Civil War (1936-1939). We also observe that height and mortality rate have a negative correlation in the medium-run. Other health indicators such as the weight of health services in total consumption and the ratio between the deflator of private consumption and the GDP deflator show a positive and negative co-movement in the short-run, respectively. Nevertheless, they change their sign of correlation in the long-run. Finally, we observe a positive co-movement between height and illiteracy rate in the short-run, a negative one in the long-run, and a strong and positive co-movement between height and the grade of openness.
1. Introduction

Anthropometric history has improved our knowledge about the quality of life of populations during the last few decades. Physical stature and other anthropometric indicators are commonly accepted by economic historians, and recently by economists, in the measurement of aspects of human well-being and in exploring the impact of socioeconomic processes on biological welfare and health outcomes (Fogel, 1994; Komlos, 1994, 1995a; Steckel 1995, 2009; Steckel and Floud, 1997). One of the main topics debated are the determinants of height and relationships between stature and the indicators of economic development. A high degree of correlation between height and economic development was observed in developing countries in the second half of the 20th century (Steckel, 1979, 1983). Later contributions have clarified the existing relations between height, income per capita, and income inequality in the long term (Brinkman et al, 1988; Drukker and Meerten, 1995; Coll, 1998; Craig and Weiss, 1998; Haines, 1998; Jacobs and Tassenaar, 2004; Moradi, 2010; Peracchi, 2008). Recently, a wide battery of indicators on standard of living was used to measure their relationships to height, resulting in important connections between health, mortality, and economic development (Easterlin, 2000; Arora, 2001; Deaton, 2003, 2007; Fogel, 2004; López-Casanovas et al, 2005; Persico et al, 2004; Steckel, 2008).

In cross-sectional analyses, the correlations between height and levels of economic development are positive and almost perfect, at least in the populations of the pre-industrial past (Komlos, 1995b, 2003). However, in longitudinal analyses, the correlations are more controversial. The first contributions established that the temporary series of income per capita could explain the variations in the changes of heights by age and that the stature can be used as a proxy of income in information absence on material well-being (Brinkman and Drukker, 1988), although it is well known that both indicators do not necessarily evolve in parallel. Many papers have studied this relationship, in world economic history. Several have found that height tends to increase in developing countries in Europe, North America, and Japan in the late 19th and early 20th centuries, regardless of the level of industrialization (Sandberg and Steckel, 1997, Weir, 1997, Shay, 1994, Honda, 1997, Baten, 2000,
Federico, 2003, Vecchi and Coppola, 2006, Arcaleni, 2006). Height declined, however, in the United States, the UK, and the Netherlands while industrialization was proceeding at the end of the 18th and the first half of the 19th centuries (Margo and Steckel, 1983; Floud et. al, 1990; Komlos, 1998, Haines, 2004). Nevertheless, there is a consensus about an international convergence in biological welfare and other non-income indicators of the standard of living during the 20th century (see Kenny, 2005, Deaton 2007, 2008). The main reason for this is that stature is also influenced -among other variables- by the conditions of health in childhood and adolescence, maternal education, public policies of social welfare, child labour, location, and cultural values (Komlos and Baten 1998; Schutkowski, 2008, Steckel, 2009).

More recent contributions, as that of Baten et al. (2010), point out that human stature in adulthood captures an accumulation of factors that contribute to changes in the standard of living of a generation. This implies that GDP per head could also explain height level.

Relationship between height and economic development also found echoes in Spain. The first studies of anthropometric history were made by economic historians who explored this issue with panels of data from military sources and by different methods. Thus, Gómez-Mendoza and Pérez-Moreda (1995) examined the relationship between height and educational attainment and infant mortality in the early 20th century. These authors compared the average height of recruits by province and published their findings in Anuarios Estadísticos de España, with economic performance and infant mortality of the region (Gómez-Mendoza and Pérez-Moreda, 1995). Martínez-Carrión explored the trends of height and income between 1850 and 1990 from local recruitment data of conscripts (Martínez-Carrión, 1994, Martínez-Carrión and Pérez-Castejón 2000); and Quiroga and Coll (2000) discussed height differences as a proxy for income inequality, and that changes within the differences in height among social groups could indicate shifts in income inequality. Given the limitations posed by some conventional indicators of welfare, the possibilities opened up by anthropometric studies are wide-ranging and relevant to the economic and social history of contemporary Spain.
Recently, using both Spaniard height data panels from the Encuesta Nacional de Salud (Spanish National Health Survey) and dataset of the European Community Household Panel (Eurostat), economists and demographers have presented new evidence on the evolution of adult height and weight (García and Quintana-Domeque 2007, 2009). On the one hand, the influence of generational or environmental effects on adult height has been evaluated and the mechanisms through which socio-economic position may influence individual height have been discussed (Costa-Font and Gil, 2008; Spijker, Pérez and Cámara, 2008). On the other hand, the determinant of height during a period of significant social-economic transformation (1960-2000) has been explored. Results suggest that the epidemiological transition before Spanish entry into the European Union led to improvements in adult height (Bosch, Bozzoli and Quintana-Domeque, 2011).

Finally, María-Dolores and Martínez-Carrión (2011) model human stature and the rest of standard of living indicators through a Vector Autoregressive Model (VAR) and estimated a Vector Autoregressive Equilibrium Correction Model (VECqM) to quantify the height and the GDP per head response to the different indicators of living standards. They observed that changes in average height for the period 1850-1958 could be explained by different indicators of economic development such as for the weight of health services in total consumption, for deflator of private consumption, schooling rate, infant mortality and trade.

In parallel, Spanish economic growth has been revised in the last few years. According to the estimates of Prados de la Escosura, Spain underperformed over the long-run mostly due to its sluggish growth in the hundred years up to 1950. Greater destruction of human rather than of physical capital during the Spanish Civil War (1936-1939) and its aftermath explain its performance during the 1940s and 1950s. The 1940s constituted a phase of delay in the Spanish economy. This economy has been catching up with advanced countries over the last fifty years, in which 1959-74 is marked out as a period of outstanding performance (Prados de la Escosura, 2003, 2007).
Comparison between the behaviour of the Spanish economy and the most advanced nations of Western Europe reveals that Spain was reaching nearly three-quarters up to the level of production in Europe and one-half up to the level of North America. The marked reduction of the Spanish relative position as a result of the civil conflict of 1936-39 would improve in the 1950; however, Spain did not reduce its distance from the developed nations. According to new estimates, Spain’s GDP per head stood at 50.4 percent of that of the United States in 1933 and decreased to 35 percent in the 1950s. The move towards a pro-market attitude with deregulation, and the gradual opening up of Spain to the international economy, resulted in sustained growth and a catching up with Western Europe during the second half of the twentieth century. A dramatic growth slowdown followed by a sustained catching up, separated by 1986 -the year of Spain’s accession to the European Union- characterized the last quarter of the twentieth century (Prados de la Escosura, 2007).

In this paper we attempt to observe the relationship between height and economic development in Spain, with new data of the height of conscripts through alternative methodologies to calculate correlations between variables. We employ several filtering methods to measure the correlation between variables such as first-order differences, deterministic trends, the Hodrick and Prescott filter or den Haan (2000)’s methodology. The methodology by den Haan (2000) uses a new set of statistics to characterize the co-movement between variables. These statistics are very effective in providing information to aid a researcher who wants to build and test a structural model, and it could be clearly applied to the relationship between height and human welfare indicators. In this study we focus on Spanish males because no long-run information is available on female height. It takes into account several development indicators, such as income per capita, the weight of health services in total consumption, the ratio between price of consumption goods and GDP deflator, schooling rate, mortality, and the degree of openness, which could potentially influence physical stature in the period 1850-1958. To our knowledge, this is the first paper in the anthropometric literature to use this methodology.
The rest of the paper is organized as follows. In Section 2 we describe our database; in Section 3 we introduce the different filtering methods and den Haan (2000)’s methodology; in Section 4 we apply them to study the relationship between height and the different potential determinants of the standard of living in Spain for the period 1850-1978; and in Section 5 we offer our main conclusions.

2. Data description

2.1. Height

For height, standardized time-cohort series can be constructed for Spain from the 1850s onwards\(^2\). Data for the early 19\(^{th}\) century are both scarce and fragmentary, and almost non-existent before the end of the 18\(^{th}\) century. The first recruitment is carried out in 1770 (with the Monarchy of Carlos IIII). During the last decades of the 18\(^{th}\) century, the replacements of the Spanish conscripts are formed by men aged between 16-40 years. In the first decades of the 19\(^{th}\) century the conscription happened to be formed by men of 16-25 years old. It is necessary to hope the decade 1850s recruitment to find conscripts of a single age: 20 years (1850-1885), 19 years (1885-1900), 20 years (1901-1906) and 21 years (1907-1970). The main sources for an anthropometric study of a long-run cohort series in Spain are the Local Military Recruitment Acts –LMRA- (the original record is Actas de clasificación de los mozos y declaración de soldados, Actas de Reclutamiento y Reemplazo). These data sources and records are preserved in the Sección de Quintas of local historical archives of the municipalities. Data samples reported in this paper were obtained from such records.

In Spain the height of every man was measured at the age of conscription, beginning in the 1850’s decade. The stature is expressed in millimetres. Height was measured by a tallador, administrative personnel of the City council. Data sources also contain information of occupation, place of birth and address (if they emigrated). Our data samples include records of all the young men who were

\(^2\) For recent up-to-date information on the sources of stature data and on anthropometric history in Spain, see the monographic number of the journal Agrarian History, 47, coordinated by Martinez-Carrión (2009).
inspected for conscription in 19 municipalities from three geographical regions (Andalusia, the Region of Murcia and the Valencian Community) or five provinces distributed by the Spanish Levant\(^3\). The municipalities are representatives of different socio-economic environments, some of which are formed by cities, towns and villages.

Our data series presented here is a good proxy for height at the national level, because their economic and environmental characteristics are considered representative of the Spanish economy of the 19th and 20\(^{th}\) centuries. Several studies on total productivity, labour productivity, income per capita, well-being and inequality, located Levant populations in the average threshold of Spain (Domínguez, 2002; Germán, Llopis, Maluquer, Zapata, 2001; Núñez, 2005). Our sample is representative of the Spanish average. Several anthropometric studies find that the stature averages of Valencia and Murcia regions are nearer to height average of the Spaniards: Estimates of average height in the replacements of 1915-29 and 1965-80 reflect the height average of Spain in these analyzed periods (Gómez-Mendoza and Pérez-Moreda, 1995: 85; Martínez-Belmonte 1983; Martínez-Carrión, 1994: 697; Rebato, 1998).

We construct average height for each cohort by year of birth. The new series of the height of Spanish men that we present in this paper contains data of 328,248 conscripts, who aged 19-21 were listed and measured between 1857 and 1969. The backgrounds of these series are in previous works elaborated by Martinez-Carrión and Pérez-Castejón (1998, 2000) and the recent dataset of Puche-Gil (2009) presented in his doctoral dissertation\(^4\). The new series on human height in Spain links with the national series elaborated by the Statistics of Recruitment and Replacement of the Armies (Estadísticas del Reclutamiento y Reemplazo de los Ejércitos) from 1954 and estimated by Quiroga and Coll (2000) and Quiroga (2003a).

---

\(^3\) Vera (province of Almería); Region of Murcia: Cartagena, Cieza, Murcia, Mazarrón, Totana, Torre-Pacheco; Yecla (province of Murcia); Alcoi, Elche, Orihuela, Pego, Villena (province of Alicante); Alzira, Gandía, Requena, Sueca (province of Valencia); Castellón y Villareal (province of Castellón).

\(^4\) Dataset is composed of 141,861 observations with height data. See doctoral thesis of Puche-Gil (2009), to whom we are grateful for providing us with the data.
The most significant problems facing the literature on height using military information are: (i) changes in the recruitment age and, (ii) the minimum height requirements for military service. In Spain, however, all potential recruits were measured before they were declared fit or not for military service or not as there was a universal draft, so there is no selective bias from a truncated distribution. With reference to the first problem in height series posed by the introduction of changes affecting conscription age and by the rounding of height data, this has been addressed by using standardized heights at age 21 taking into account estimations of growth. Data here are presented by cohorts of births. We built a new height series using five-year moving averages standardized at the age of twenty-one years.

How representative is our evidence? In other papers the reliability of data has been ascertained using the Kolmogorov-Smirnow and similar tests. On this occasion, data are normally distributed or Gaussian and do not suffer from typical truncation problems (Figure 1).

**Figure 1: Distribution of heights by year of birth (1837-1948)**

Figure 1.a Distribution of heights by year of birth, 1837-1865

---

5 On height standardized at the age of 21 years, see Martínez-Carrión and Moreno-Lázaro (2007).
6 Martínez Carrión and Pérez Castejón (2000) use a sample of the height of 127,310 conscripts out of a total of 141,911 men (89.7 per cent) called up for service. This height series has a high correlation with Quiroga’s height series for the same sample period (0.93); see Quiroga (2003a). On height standardized at the age of 21 years, see Martínez-Carrión and Moreno-Lázaro (2007).
Figure 1.b. Distribution of heights by year of birth, 1866-1885

Figure 1.c. Distribution of heights by year of birth, 1886-1915
1.2. Income and other measures of welfare

For real income, we use the annual series of Spanish GDP per head at constant 1995 prices in pesetas provided by Prados de la Escosura (2003). His recent estimates on the economic growth show that in Spain there were three main phases in the long-run economic development: 1850-1950, 1951-1974 and 1975-2000, with a shift to a lower level during the first period as a consequence of the Civil War of 1936-1939. Phases or long swings in which growth rates differ from the long-run trend as a result of economic policies, access to international markets, and technological change can be distinguished. During the first phase, 1850-1883, the growth rate of GDP per capita was well above the 19th century’s average. A slowdown in growth took place between the mid 1880s and 1920.

The most intense growth of the period achieved in the 1920s coincided with the Primo de Rivera Dictatorship (1923-1929). This was a period of institutional stability that provided a favourable environment for investment and business. Lastly, a fourth long swing took place between 1929 and 1952. The Civil War (1936-39) had a negative impact on the later growth, reducing the product per capita to nearly 15% of the secular
growth rate (Prados de la Escosura, 2003: 148). The weak recovery of the period 1944 - 1952 stands out in the international context. In spite of World War II, European economies achieved an average growth rate of 1.4 compared with Spain’s 0.6 during the 1940s. The Spanish economy did not recover pre-war GDP levels until 1951 (absolute terms), and 1955 (per capita terms). The change in trend beginning in 1951 ushered in an exceptional phase of rapid growth which lasted until 1974. As in other countries in the European Periphery during the Golden Age (1950-1973), the main spurt of economic growth in Spain was delayed until the 1960s. Catching up took place in the late 20th century, in which the years 1959-74 are marked out as a period of outstanding performance (Prados de la Escosura, 2007).

Figure 2 compares the evolution of GDP per head and height for the period 1850-1978. GDP per head refers to the year in which recruits were measured. Our results do not reveal a clear relationship between both series until the end of the 19th century, when similarities between the two series increased. Nevertheless, that relationship also seems to be high for the period 1900-1920. We obtain a correlation of 0.91 between both series for the whole sample period. Table 1 shows the average height and GDP per cápita for each decade, together with their variations. The lack of correlation between income and height during the initial stages of modern economic growth may well be related to Kuznets’ inverted U hypothesis, namely that income inequality rises and then falls with the level of economic development. Recent research observes a long-term rise in the inequality index during the early phase of globalization that had peaked by World War I (Prados de la Escosura, 2008). With other databases of Spanish recruits, Quiroga and Coll (2000) show a long-term increase in height inequality among socio-professional groups between the turn of the century and World War I. Martínez-Carrión and Moreno-Lazarro (2007) demonstrate that inequality trends between rural and urban areas increased during the late 19th century. The Spanish case could support this hypothesis already verified in other countries during early industrialization and the first stages of economic growth, although this is not the main focus of the paper.

---

7 We have also examined the relationship between the detrended GDP per head and height series and we found a positive value.
Figure 2: The relationship between height (measurement year) in centimetres (left vertical axis) and GDP per capita in pesetas 1995 (right vertical axis), 1850-1978
Table 1: Average height and GDP per head (1850-1978)

<table>
<thead>
<tr>
<th>Average Height (measurement year in centimeters)</th>
<th>Average GDP pc (pesetas1995)</th>
<th>Height variation (centimeters)</th>
<th>Annual GDP pc growth rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1850-1879</td>
<td>162.07</td>
<td>163.90</td>
<td>0.3029</td>
</tr>
<tr>
<td>1880-1919</td>
<td>163.92</td>
<td>229.53</td>
<td>1.9099</td>
</tr>
<tr>
<td>1920-1939</td>
<td>164.94</td>
<td>297.95</td>
<td>1.5289</td>
</tr>
<tr>
<td>1940-1959</td>
<td>167.75</td>
<td>311.91</td>
<td>4.22</td>
</tr>
<tr>
<td>1960-1978</td>
<td>173.05</td>
<td>777.70</td>
<td>3.945</td>
</tr>
</tbody>
</table>


The literature on height and economic growth has shown the importance of using anthropometric measures linked to different dimensions of health and standard of living. Recently, some studies have shown the beneficial effects of nutritional improvements, comprehensive health care, health awareness and suitable housing conditions on height (Komlos and Lauderale 2007). With reference to these lines of research, we use two different measures of consumption directly related to health: consumption of health services and the price level of consumption goods. These series are only available up to 1958 and are provided by Prados de la Escosura (2003) at constant 1995 prices. Following the interest in the relationship between mortality and height (Fogel 2004), we employ the mortality rate from 1858 to 1980 (Nicolau, 2005: 124-6; Carreras and Tafunell, 2005).
Figure 3: The relationship between height (measurement year) in centimetres (left vertical axis) and Mortality rate (right vertical axis).

Figure 4: The relationship between height (measurement year) in centimetres (left vertical axis) and the weight of health services in total consumption (right vertical axis).
Figure 5: The relationship between height (measurement year) in centimetres (left vertical axis) and the ratio of private consumption deflator divided by GDP deflator (right vertical axis).

Table 2: Average height and Standard of living Indicators (1850-1978)

<table>
<thead>
<tr>
<th>Period</th>
<th>Average Height (measurement year in centimeters)</th>
<th>Mortality Rate (%)</th>
<th>Weight of Consumption of hygiene products / Total Consumption</th>
<th>Ratio of Deflator for private consumption/ GDP deflator (1995=100)</th>
<th>Population no schooling rate (%)</th>
<th>Openness (exports +imports/GDP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1850-1879</td>
<td>162.07</td>
<td>30.19</td>
<td>0.47</td>
<td>1.41</td>
<td>55.74</td>
<td>0.96</td>
</tr>
<tr>
<td>1880-1919</td>
<td>163.92</td>
<td>27.64</td>
<td>0.78</td>
<td>1.47</td>
<td>53.00</td>
<td>3.22</td>
</tr>
<tr>
<td>1920-1939</td>
<td>164.94</td>
<td>18.45</td>
<td>1.46</td>
<td>1.40</td>
<td>46.77</td>
<td>3.24</td>
</tr>
<tr>
<td>1940-1959</td>
<td>167.60</td>
<td>11.75</td>
<td>3.04</td>
<td>1.73</td>
<td>33.29</td>
<td>2.86</td>
</tr>
<tr>
<td>1960-1978</td>
<td>173.05</td>
<td>8.45</td>
<td>-</td>
<td>1.72</td>
<td>21.35</td>
<td>17.63</td>
</tr>
</tbody>
</table>

Sources: Height, table 1; mortality, Nicolau (2005); consumption of hygiene products, deflator for private consumption, GDP deflator and consumption per head, Prados de la Escosura (2003); population no schooling rates, Núñez (2005: 232-236); openness, (Tena, 2005: 628-30)
Figures 3-5 plot the relationship between stature, mortality rate, the weight of health services in total consumption and the ratio of private consumption deflator divided by GDP deflator, respectively. We observe how the relationship between GDP per head and mortality rate is clearer after the Civil War, although it is also high during the rest of the sample period (see Table 2). The effects of the ‘flu’ epidemic in 1918 and the Civil War 1936-39 were immediately shown in the mortality series and also influenced the height of following generations of conscripts, measured during the 1933-1959 period. With regard to the weight of health services we observe a high value of the correlation coefficient (0.83).

The relationships between stature and human capital measured by the levels of education, schooling of children, and literacy rates has also received the attention of specialists (Meyer and Selmer, 1999; Schultz 2003). Recently, Case and Paxson (2008) have recognized that height is associated with better physical and mental health and cognitive ability. We have also included some education variables in our study such as the schooling rate (Núñez, 2005). Spanish historians have indicated that education and the human accumulation of capital in Spain have been delayed processes, very unequal in men and women at least until 1960. This constitutes a factor that explains part of the historical delay, relatively compared with the development in other European countries. Although studies demonstrate that the literacy process spread widely and average height increased in the course of the first half of 20th century, there was inequality increase between literate and illiterate conscripts after the Civil War, which is confirmed by the divergence in stature in both access to education and in deterioration of signature quality of Spanish recruits (Martínez-Carrión and Puche-Gil, 2009; Quiroga, 2001: 616-617).

The new estimate confirms that most of the human capital embodied in the Spanish population in the second half of the 20th century was due to expanded primary schooling rather than to secondary or university studies. Likewise, it identifies the Civil War of 1936-39 as one of the most serious setbacks during two centuries of slow and irregular capital human accumulation. The Civil War and the early years of the Franco regime contributed to the depletion of the stock of human capital and had negative effects on welfare population (Núñez, 2005). Our results show that there to be a
negative correlation of -0.77 between height and schooling rate. Figure 6 offers the relationship between height and illiteracy rate.

![Graph showing the relationship between height and illiteracy](image)

**Figure 6:** The relationship between height in centimetres (left vertical axis) and the population no schooling rates (right vertical axis)

Finally, we have also included other important development indicators such as the *grade of openness* of the country, measured as the ratio export plus imports over GDP, and built using Prados de la Escosura (2003: 188). There are very influential papers in the literature that claim to find a negative association between barriers to trade and economic growth (Ben-David, 1993, Sacks and Warner, 1995, Edwards, 1998 and Frankel and Romer, 1999). The association between health, welfare state, and openness continues to be little investigated in the anthropometric history. Such a subject would be interesting in Spain in order to evaluate the effects of the autarkic and economic liberalization policies during the long Dictatorship of Franco (1939-1975). Studies made on the Nazi autarky and old communist countries in Europe show the negative effects
on well-being, as a result of the strong controls and the restrictions of the market: fixed prices, production contingents, import restrictions, food rationing and limitations on certain types of foods that people could eat (Baten and Wagner, 2003; Laska-Mierzejewska, T. and Olszewska, E. 2007). The relation between height and openness could also take into consideration the integration of international trade and the increase in the size of the state with the introduction of the welfare state. The influence of the welfare state in the physical growth of population has been made in the Norwegian case (Sunder, 2003). Nevertheless, this association could consider integration with international trade and increase in the size of the state as coming about with the introduction of the welfare state. The annual series of the grade of openness we used came from recent estimations made by Tena (2005: 628-631).

Figure 7: The relationship between height in centimetres (left vertical axis) and the grade of openness (right vertical axis)

Figure 7 shows the relationship between height and openness. This relationship is shown to be positive (0.62), especially after 1959. The Stabilization Plan in 1959 ushered in a new era during the dictatorship of General Franco, when Spain experienced
larger growth rates than ever before. This relationship was also strong during the period 1880-89 and before the Civil War date. Our data reveal that the autarky had a negative effect on the stature of adolescents; stature diminished in the replacements of the 1940s, as appraised in cohorts of the 1920s. Other anthropometric studies made recently with data of the Spanish National Health Survey (2003-2006), for several generations, indicate that the economic liberalization of 1959 reduced the height difference between genders (Costa-Font and Gil, 2008; Spijker, J, Pérez, J., Cámara, A. 2008).

3. Filtering methods to measure the correlation between stature and economic development indicators

In this section we introduce some filtering methods to measure the correlation between stature and economic development indicators during our period of analysis. We propose different detrending filters that lump together the irregular component and the cyclical component of the series. We will use the following groups of filtering methods:

3.1. First group of filtering methods

First order differences.

This method takes the cycle to be the variable in first differences. It assumes that the trend is the lagged variable or, similarly, the series is a random walk with no drift. Therefore, a series, \( y_t \), can be represented as follows:

\[
y_t = y_{t-1} + C_t + \varepsilon_t
\]

where the trend is \( y_{t-1} \) and an estimate of the detrended component is obtained as \( y_t - y_{t-1} \).
Deterministic trends

The usual procedure in this case is to take the least squares residual after regressing the series on a constant and a polynomial function of time. The implicit assumption is that the trend and cyclical components are orthogonal, and the trend is a deterministic process which can be approximated with polynomial functions of time. These assumptions imply a model for \( y_t \) of the form:

\[
y_t = T_t + C_t + \varepsilon_t
\]

\[
T_t = a_0 + a_1t + a_2t^2
\]

We take as polynomial order two.

Hodrick and Prescott filter

The Hodrick and Prescott filter is perhaps the most popular one in applied macroeconomics. This filter extracts a stochastic trend that moves smoothly over time and is not correlated with the cycle component. An estimate of the trend is obtained by minimizing:

\[
\min \{ \theta \} \left[ \sum_{r=1}^{T} C_r^2 - \lambda \sum_{r=1}^{T} ((T_{r+1} - T_r) - (T_r - T_{r-1})) \right]^2, \lambda > 0
\]

The result depends on a smoothing parameter (\( \lambda \)) that penalizes large fluctuations. A large value implies a higher penalty and, therefore, a smoother cycle. For annual data the value of \( \lambda \) typically used is 100.
**Band pass filter**

This filter is a frequency domain based filter. It assumes that the trend component has the power at lower frequencies of the spectrum. In this case we define the limits of the frequency band, say $p_l$ and $p_u$, to isolate the cyclical component with a period of oscillation between $p_l$ and $p_u$. We make a choice for the cycle length between 2 and 8 years.

**3.2. Den Haan (2000)'s methodology**

Den Haan (2000) points out that an important source of disagreement in the literature is the focus on only one correlation coefficient. By focusing on only the unconditional correlation, one is losing valuable information about the dynamic aspects of the co-movement of variables. Following den Haan (2000), we estimate the correlation coefficients of VAR forecast errors by calculating the forecast errors for each horizon considered (from one year to 40 years) as the difference between the realizations and the corresponding forecasts, and then calculating the correlations of these forecast errors for each horizon. One of the advantages of this methodology is that we obtain dynamic correlations.

Next, we briefly describe how to use a VAR to study the correlation structure between two variables (for instance, height and some development indicators) at several forecast horizons.

Let us consider an N-vector of random variables, $X_t$. The vector $X_t$ may include any combination of stationary processes and integrated processes of arbitrary order. In order to characterize the co-movement, $X_t$ must contain at least (the log of) two variables. Consider the following VAR

$$X_t = \alpha + \beta t + \gamma t^2 + \sum_{i=1}^{\ell} A_i X_{t-i} + U_t(4)$$
where $\alpha$, $\beta$ and $\gamma$ denote fixed N-vectors of constants, $A_t$ represents fixed $N \times N$ coefficient matrices, $U_t$ is an N-dimensional white noise process, that is $E(U_t)=0$, $E(U_tU_t')=\Omega$ and $E(U_tU_t')=0$ for $s \neq t$. $L$ is the total number of lags included. The $K$-period ahead forecast and the $K$-period ahead forecast error of the random variable $H_t$ (height) are denoted by $E_{t+K}H_t$ and $H_{t+K,t}^{ue}$, respectively. Similarly, we can define $E_{t+K}Y_t$ and $Y_{t+K,t}^{ue}$. Let us denote the correlation coefficients between $H_{t+K,t}^{ue}$ and $Y_{t+K,t}^{ue}$ by $\text{COR}(K)$.

As pointed out by den Haan (2000), if all time series included in $X_t$ are stationary, then the correlation coefficient of the forecast errors will converge to the unconditional correlation coefficient between $H_t$ and $Y_t$ as $K$ goes to infinity. If $X_t$ includes integrated processes, then correlation coefficients may not converge but they can be estimated consistently for the fixed $K$.

Den Haan (2000) also shows the relationship between correlation coefficients and impulse response functions. Let us denote the covariance between $H_{t+K,t}^{ue}$ and $Y_{t+K,t}^{ue}$ by $\text{COV}(K)$ and, with no loss of generality, let us assume that there are $M$ structural shocks driving height and the other variable. Den Haan (2000) shows that

$$\text{COV}(K) = \sum_{k=1}^{K} \text{COV}^\Lambda(k)$$

$$\text{COV}^\Lambda(k) = \sum_{m=1}^{M} H_{t+k}^{\text{imp},m} Y_{t+k}^{\text{imp},m}$$

where $z_{k}^{\text{imp},m}$ is the $k$-th period impulse response of variable $z$ to a one standard deviation disturbance of the $m$-th shock. Therefore, the covariance between height and the other development indicator variable is simply the sum of the products of height and the other variable impulses across the different structural shocks.

By observing the correlation coefficients of VAR forecast errors at different horizons, the researcher can obtain richer information about system dynamics than by looking at the unconditional correlation coefficient. As illustrated by den Haan (2000),
considering only one correlation coefficient might be misleading in some cases. Moreover, den Haan’s method avoids the type of ad-hoc assumptions necessary to compute impulse response functions. There is a shortcoming, however. This procedure does not identify all the different impulse response functions (that is, it does not identify the response to all the different structural shocks).

Note that the application of den Haan (2000)’s methodology is in accordance with Jacobs and Tassenaar (2004) and María-Dolores and Martínez-Carrión (2011) where human stature is modelled as a Vector Autoregressive Model (VAR). They assume that the (unobserved) increments in height depend linearly on income and other explanatory variables, such as described in section 2.

We study the co-movement by considering annual data. Using this annual data, we first study the co-movement between height and GDP per cápita. Second, we analyze the co-movement between height and mortality rate, height and the weight of health services, and height and the ratio of deflator of private consumption to GDP deflator. All these are some health indicators. Finally, we observe the co-movement between height and illiteracy rate and height and openness.

4. Main results.

4.1. First group of filtering methods.

Table 3 offers the main results of applying the various filtering methods introduced in Section 3. We appreciate a strong and positive correlation between height and GDP per head with all the filtering methods. This correlation is larger by using deterministic trends filtering methods than the rest.

---

8 According to the anthropometricians, the average height of a given cohort is a function of their nutritional intake during infancy and childhood, setting aside genetic factors. Net nutrition is defined as the gross food intake less losses due to maintenance necessities, disease and work exertion. So, the influence of income on height is indirect. Income affects height because is one of the determinants of food and hygiene products consumption and through its relation with child labour and disease environment.
Table 3: Main correlation between height and Standard of living Indicators
(1850-1978)

<table>
<thead>
<tr>
<th></th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Height, GDP per head (i))</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. First order differences</td>
<td>0.78</td>
<td>0.82</td>
<td>0.86</td>
<td>0.81</td>
<td>0.77</td>
</tr>
<tr>
<td>2. Deterministic trends</td>
<td>0.93</td>
<td>0.95</td>
<td>0.97</td>
<td>0.89</td>
<td>0.86</td>
</tr>
<tr>
<td>3. Hodrick and Prescott filter</td>
<td>0.79</td>
<td>0.83</td>
<td>0.88</td>
<td>0.85</td>
<td>0.82</td>
</tr>
<tr>
<td>4. Band pass filter</td>
<td>0.76</td>
<td>0.80</td>
<td>0.85</td>
<td>0.81</td>
<td>0.78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>-0.84</th>
<th>-0.87</th>
<th>-0.89</th>
<th>-0.85</th>
<th>-0.81</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Height, mortality rate (i))</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. First order differences</td>
<td>-0.84</td>
<td>-0.87</td>
<td>-0.89</td>
<td>-0.85</td>
<td>-0.81</td>
</tr>
<tr>
<td>2. Deterministic trends</td>
<td>-0.93</td>
<td>-0.96</td>
<td>-0.99</td>
<td>-0.97</td>
<td>-0.94</td>
</tr>
<tr>
<td>3. Hodrick and Prescott filter</td>
<td>-0.89</td>
<td>-0.93</td>
<td>-0.97</td>
<td>-0.91</td>
<td>-0.85</td>
</tr>
<tr>
<td>4. Band pass filter</td>
<td>-0.80</td>
<td>-0.87</td>
<td>-0.93</td>
<td>-0.90</td>
<td>-0.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0.84</th>
<th>0.86</th>
<th>0.87</th>
<th>0.83</th>
<th>0.79</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Height, weight of health services (i))</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. First order differences</td>
<td>0.84</td>
<td>0.86</td>
<td>0.87</td>
<td>0.83</td>
<td>0.79</td>
</tr>
<tr>
<td>2. Deterministic trends</td>
<td>0.91</td>
<td>0.94</td>
<td>0.98</td>
<td>0.95</td>
<td>0.92</td>
</tr>
<tr>
<td>3. Hodrick and Prescott filter</td>
<td>0.87</td>
<td>0.89</td>
<td>0.90</td>
<td>0.86</td>
<td>0.82</td>
</tr>
<tr>
<td>4. Band pass filter</td>
<td>0.85</td>
<td>0.87</td>
<td>0.89</td>
<td>0.85</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>-0.38</th>
<th>-0.42</th>
<th>-0.48</th>
<th>-0.45</th>
<th>-0.42</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Height, relative deflator of consumption (i))</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. First order differences</td>
<td>-0.38</td>
<td>-0.42</td>
<td>-0.48</td>
<td>-0.45</td>
<td>-0.42</td>
</tr>
<tr>
<td>2. Deterministic trends</td>
<td>-0.92</td>
<td>-0.95</td>
<td>-0.98</td>
<td>-0.96</td>
<td>-0.94</td>
</tr>
<tr>
<td>3. Hodrick and Prescott filter</td>
<td>-0.57</td>
<td>-0.62</td>
<td>-0.66</td>
<td>-0.64</td>
<td>-0.62</td>
</tr>
<tr>
<td>4. Band pass filter</td>
<td>-0.46</td>
<td>-0.47</td>
<td>-0.48</td>
<td>-0.45</td>
<td>-0.42</td>
</tr>
</tbody>
</table>
Table 3: Main correlation between height and Standard of living Indicators (1850-1978) (cont.)

(Height, openness (i))

<table>
<thead>
<tr>
<th>Method</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. First order differences</td>
<td>0.60</td>
<td>0.61</td>
<td>0.59</td>
<td>0.55</td>
<td>0.52</td>
</tr>
<tr>
<td>2. Deterministic trends</td>
<td>0.71</td>
<td>0.71</td>
<td>0.69</td>
<td>0.66</td>
<td>0.61</td>
</tr>
<tr>
<td>3. Hodrick and Prescott filter</td>
<td>0.60</td>
<td>0.62</td>
<td>0.65</td>
<td>0.61</td>
<td>0.58</td>
</tr>
<tr>
<td>4. Band pass filter</td>
<td>0.52</td>
<td>0.57</td>
<td>0.60</td>
<td>0.57</td>
<td>0.53</td>
</tr>
</tbody>
</table>

(Height, population no schooling rates (i))

<table>
<thead>
<tr>
<th>Method</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. First order differences</td>
<td>-0.75</td>
<td>-0.78</td>
<td>-0.80</td>
<td>-0.77</td>
<td>-0.73</td>
</tr>
<tr>
<td>2. Deterministic trends</td>
<td>-0.93</td>
<td>-0.96</td>
<td>-0.99</td>
<td>-0.97</td>
<td>-0.94</td>
</tr>
<tr>
<td>3. Hodrick and Prescott filter</td>
<td>-0.77</td>
<td>-0.82</td>
<td>-0.86</td>
<td>-0.73</td>
<td>-0.70</td>
</tr>
<tr>
<td>4. Band pass filter</td>
<td>-0.74</td>
<td>-0.77</td>
<td>-0.81</td>
<td>-0.77</td>
<td>-0.74</td>
</tr>
</tbody>
</table>

The correlation between height and the mortality rate is strong and negative. We observe again a highest value by using deterministic trends filters. With reference to the relationship between height and the weight of health services in total consumption, its correlation is also positive and relevant.

The co-movement between the ratio of the deflator of private consumption to GDP deflator and height is, as expected, negative. We obtain that this correlation is not so high as the rest of the variables by using first order differences or band pass filters. With regard to the relationship between height and openness, we observe a positive relationship, although this relationship is greater with a lag in openness when we use first order differences or a deterministic trend method. Finally, the correlation between height and population no schooling rate is always negative.
4.2. De Haan (2000)'s methodology

We estimate correlation coefficients based on VARs that only include height and different economic development indicators. Another question that arises is whether or not to impose a unit root in the estimation of the VAR. We estimate each VAR twice; once using the first difference and once using levels. In both cases the forecast errors are calculated for the levels. The Akaike information criterion was used to determine the number of lags for each VAR system and whether a deterministic trend term should be included. Since the estimated correlation coefficients are subject to sampling variation, confidence bands are constructed using bootstrap methods. More specifically, the bootstrapped errors of each estimated VAR are used to generate 2500 simulated data sets. Then, the correlation coefficients at different horizons are estimated for each simulated data set and standard confidence bands are calculated.

As we mentioned before, one of the main advantages of this methodology is that we obtain dynamic correlations. The results are presented in Figs 8-23. Figure 8 plots the correlation coefficients of the K-period ahead of forecast errors when a unit root is imposed in the estimation of the VAR, and Figure 9 plots the correlation coefficients when no unit root is imposed. The results are very similar for the different VAR specifications. For the short-term forecast horizons the correlation coefficients are positive, although not very high, and for the long-term forecast horizons the correlation coefficients are stronger. Note that the average height of a given cohort is a function of their nutritional intake during infancy and childhood and this variable is related with the GDP per head. So, it is usual to find this relationship larger in the medium and long-run than in the short-run.
By considering the sample period 1850-1978, we obtain a low level of correlation between height and GDP per capita in differences. For an illustrative purpose, we re-calculate our correlations using two subsamples: 1850-1939 (pre Civil War period) and 1940-78. We find a negative correlation between height and GDP per
capita during the previous years to the Spanish Civil War and a high and positive correlation for the post-war period (with a maximum of almost 0.65).

Figures 10-13 plot the correlation coefficients of the K-period ahead of forecast errors whether or not a unit root is imposed in the estimation of the VAR for height and GDP per head. We observe a negative correlation between height and GDP per head during the years previous to the Spanish Civil War. This correlation increases again in the long-run. If we centre on the period 1940-78, this correlation is higher and more positive than by considering the whole sample period.

Figure 10: Co-movement between height and GDP per capita in differences 1850-1939
Figure 11: Co-movement between height and GDP per capita in differences 1940-1978

Figure 12: Co-movement between height and GDP per capita 1850-1939
Figures 14 and 15 plot the correlation coefficients of the K-period ahead of forecast errors whether or not a unit root is imposed in the estimation of the VAR for height and the mortality rate. We observe a negative correlation between height and infant mortality. This correlation moderately increases in the long-run.
Figures 14 and 15 plot the correlation coefficients of the K-period ahead of forecast errors whether or not a unit root is imposed in the estimation of the VAR with height and the weight of health services in total consumption. We observe a negative correlation between height and real consumption of hygiene products. This correlation is positive in the short-run and turns out to be larger and negative in the long-run.
Figures 18 and 19 plot the correlation coefficients of the K-period ahead of forecast errors whether or not a unit root is imposed in the estimation of the VAR with height and the ratio of deflator of private consumption to GDP deflator. We observe a negative correlation between height and this rate in the short-run. This correlation is positive in the long-run. How can we interpret these results? We think that health
indicators such as weight of health services and deflator of private consumption could clearly influence height in the short-run more than in the long-run.

![Figure 18: Co-movement between height and the ratio of private consumption deflator divided by GDP deflator 1850-1958](image1)

![Figure 19: Co-movement between height and the ratio of private consumption deflator divided by GDP deflator in differences 1850-1958](image2)

Next, we proceed to check how other human development indicators, which influence the level of income, are related to height. Figures 20 and 21 plot the correlation coefficients of the K-period ahead of forecast errors whether or not a unit root is imposed in the estimation of the VAR with height and the population no
schooling rates. We observe a positive correlation between height and the illiteracy rate in the short-run and a negative correlation in the long-run. This result is counter-intuitive in the short-run and as expected in the long-run. The less the population no schooling rates the greater the Spaniards’ height as was indicated in some papers above-mentioned.

Figure 20: Co-movement between height and population no schooling rates 1850-1978

Figure 21: Co-movement between height and population no schooling rates in differences 1850-1978
Finally, Figures 22 and 23 plot the correlation coefficients of the K-period ahead of forecast errors whether or not a unit root is imposed in the estimation of the VAR with height and the grade of openness. We always observe a positive and strong correlation between height and grade of openness increasing with the forecast horizon as expected.

Figure 22: Co-movement between height and openness 1850-1978

Figure 23: Co-movement between height and openness in differences 1850-1978
5. Conclusions

This paper has investigated the relationship between height and some economic development indicators in Spain for the period 1850-1978. For that purpose, we have proceeded to apply several filtering methods.

First, we apply filtering methods based on first order differences, deterministic trends, the Hodrick and Prescott filter and the band pass filter. We obtain a strong and positive relationship between height and GDP per capita and height and the weight of health services in total consumption. We also observe positive correlation between height and openness, although with a certain lag. With reference to mortality rate, the ratio of deflator of private consumption to GDP deflator and population no schooling rate we observe a negative co-movement.

By applying de Haan (2000)'s methodology to observe the relationship between height and GDP per cápita we derive a positive correlation between them, increasing considerably in the medium run up and reaching a maximum after nine years. This correlation is more pronounced after the Spanish Civil War. Other health indicators such as mortality rate, the weight of health services in total consumption, and the ratio of deflator of private consumption to GDP deflator offer expected results. There is a negative correlation between height and infant mortality and height and the deflator of private consumption. This correlation is negative in the short-term but positive in the long-run in the case of the deflator of private consumption. We observe a positive correlation between height and the weight of health services in the short-term which turns out to be negative in the long-run. These latter variables offer expected results in the short-run.

Finally, other indicators such as illiteracy rate or the grade of openness in the economy also offer interesting results. There is a positive correlation between height and illiteracy rate in the short-run but turns negative and larger in the medium-run and long-run (as expected). We observe a large correlation between height and the grade of openness in the economy, increasing considerably in the long-run.
References


Bosch, M., Bozzoli, C. and Quintana-Domeque, C. 2011. Infant Mortality and Adult Stature in Spain, Social Sciences and Medicine, 72, 11, 1893-1903.


